DETECTION OF FRAUDULENCE IN FISH MARKETING USING DNA BARCODING AS A MOLECULAR TECHNIQUE

A thesis submitted in partial fulfillment for the degree of Master of Science (MS) in Fisheries, University of Dhaka

Department of Fisheries
University of Dhaka
Dhaka-1000
Bangladesh

Submitted By
Roll No. : RH 706
MS Session : 2014-2015
Registration No. 2010-113-002
Session: 2010-11

December 2015

Dedicated

to

My Father

(Late Jagodish Chandra Mistry)

CERTIFICATE

We certify that the research work embodied in this thesis entitled ***"Detection of fraudulence in fish marketing using DNA Barcoding as a molecular technique"*** *submitted by* ***Sagorika Smriti,*** *Roll No.: Curzon-706, Session: 2014-15, Registration No.: 2010-113-002/2010-11, has been carried out under our 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.

We wish every success in her life.

Dr. Rakha Hari Sarker
Professor
Department of Botany
University of Dhaka
Dhaka-1000, Bangladesh

Dr Mohammad Shamsur Rahman
Associate Professor
Department of Fisheries
University of Dhaka
Dhaka-1000, Bangladesh

ACKNOWLEDGEMENT

My heartfelt gratitude to **Dr Mohammad Shamsur Rahman**, Associate Professor, Department of Fisheries, University of Dhaka for his incessant supervision, designing the study, guidance and constructive criticism throughout the work. I am also beholden to him for his boundless patience for careful reading and reviewing the write-up.

My heartfelt gratitude to **Dr Rakha Hari Sarker**, Professor, Department of Botany, University of Dhaka, for his intellectual guidance, constructive criticism, constant supervision and patience with me. My special thanks to him for his valuable advice and incessant encouragement during the present research work.

I am deeply indebted to **Professor Dr M Imdadul Hoque**, Dean of the Faculty of Biological Sciences, University of Dhaka, for giving kind permission to use 'Plant Breeding and Biotechnology Laboratory' for conducting this research work.

I express my sincere and heartiest gratitude to to **Mr Anwar Hossain**, Assistant Professor, Department of Fisheries and **Dr Tahmina Islam**, Lecturer, Department of Botany for their valuable suggestion, support and encouragement during my research work.

I am also grateful to Ministry of Science and Technology for funding for conducting my Thesis work. My especial thanks to all the members of Plant Breeding and Biotechnology Laboratory especially Shawan Dada, Sujay Da, Nita Apu, Eva Apu, Aneesa Apu, Yunus, Easha, Madhury, Tantri, Kona, Kaniz, Ruman, Rita didi and Salauddin vai for their untiring help and inspiration. I would like to express my appreciation to Halima, Masrupa, Abdur Razzak, Arif for their friendly company during the research work. I am deeply indebted to my Sisters Swarna and Ayshe for their incessant inspiration and friendly company during the preparation of this thesis.

I am grateful to my husband **Mr Sushanto Roy** for his friendly company, inspiration and help during the preparation of this thesis.

Finally, I would like to express my gratitude to my family member especially Mother-in-law for their love, moral support and co-operation throughout the period.

ABSTRACT

Fraudulence in fish market has emerged as a global problem. According to, The Consumer Rights Protection Act 2009 consumers are more and more demanding about composition and provenance of processed, unprocessed of edible food products. Recently, DNA barcoding has achieved support as a rapid, cost-effective and broadly applicable molecular diagnostic technique for this purpose.

However, the maturity of the barcode database as a tool for any kind of food authentication has yet to be authenticated using real market samples. The present case study was undertaken for this reason. This study was to conduct a vigorous, repeatable species substantiation protocol that could be used to benchmark the current and future incidences of mislabeling in Bangladesh fish market. In this study, we used a DNA barcoding approach to identify species substitutions cases in different fish species sold in Super shop and local fish market in Bangladesh.

We amplified the cytochrome oxidase c subunit 1(COI) barcode sequence (656 bp long) for all the analyzed specimens, and we compared them with reference sequences from different databases (GenBank and BOLD). Though the database is undergoing continual development, it was able to provide species matches of >81.81% sequence similarity for 10 samples tested. The overall fraudulence was 60% but for specific fraudulence for whole fish is 71.42% and fraudulence for fillet fish was 33.33%.

These results demonstrate that DNA barcoding is a reliable tool for detecting fish products adulteration in Bangladesh. We recommend its use for control and law enforcement to get rid of from fish market fraudulence.

TABLE OF CONTENTS

Title	Page No
Title page	I
Certificate	iii
Acknowledgement	iv
Abstract	v
Table of Contents	vi
List of Tables	Viii
List of Figures	ix
List of Flow chart	x
Chapter 1: Introduction	**1**
1.1 Background	1
1.2 Problem Statement	7
1.3 Rationale	9
1.4 Objectives	11
Chapter 2: Materials & Methods	**12**
2.1 Experimental fish	12
2.2 Sampling	13
2.3 Processing of samples	13
2.4 Isolation of DNA from fish muscle tissue	13
2.4.1Chemicals used for isolation of muscle DNA	13
2.4.2 Required materials for agarose gel electrophoresis	14
2.4.3 Preparation of Stock and Working Solutions used for genomic DNA Isolation	15
2.5 Genomic DNA isolation	17
2.6 Preparation of Stock Solutions Used for Gel Electrophoresis	18
2.7 Preparation of PCR Reaction Mixture	20
2.7.1 PCR Amplification	20
2.7.2 Electrophoresis of the amplified products and documentation	20

2.8 PCR product purification 21
2.9 Quantification of DNA concentration 22
2.10 Software use for sequence analysis 25
2.10. 1 NCBI BLAST 25
2.10.2 BOLD Database 25
2.10.3 MEGA6.06 26
2.10.4 CLUSTALW 26

Chapter 3: Results 27
3.1 The effect of species substitution and mislabeling in fish market 29
3.2 DNA sequence interpretation 29
3.3 DNA sequence alignment 34
3.3.1. Alignment between sn1 and *Sardinella longiceps* 34
3.3.2. Alignment between mm2 and *Rastrelliger kanagurta* 36
3.3.3 Alignment between cn3 and *Pseudocryptes elongates* 37
3.3.4 Alignment between lg4 and *Pampus chinesis* 38
3.3.5 Alignment of sequence between ma5 and *Clarius batrachus* 39
3.3.6 Alignment of sequence between cg7 and *Gudusi*a *chapra* 40
3.3..7 Alignment of Sequence of ta10 and *Tor putitora* 41
3.3.8 Multiple Sequence alignment 41
3.4 Phylogenetic tree analysis 44
3.5 Proportion (in percentage) of correctly labeled and potentially mislabeled 45
3.6 comprehensive labeling requirement 47
3.7 DNA Barcode analyze and optimizes threshold 47

Chapter 4: Discussion 50
Chapter 5: Conclusion and Recommendations 55
Literature cited 56
Appendices 61

LIST OF TABLES

Table	Captions	Page
1	List of fish species of different groups	14
2	Random primers used in the present study for screening	16
3	Volume of solutions to prepare extraction buffer	19
4	Component of PCR Reaction Mixture (for 17 reactions)	22
5	Measurement of DNA concentration Table 6: List of all identification results using the GENE Bank as search engines	25
6	Identification of Collected Samples using the Gene Bank and Bold Search Engines	29
7	Polymorphic sites observed between two sequences with the existing sequence in Bangladesh	32
8	polymorphic difference between sequence *Rastrelliger kanagurta* and *Megalaspis cordyla*	37
9	polymorphic sites observed between the two spp	39
10	Polymorphic sites found in different nucleotide	40
11	Identification of polymorphic sites between of ta10(collected) and *Tor putitora*	42
12	Visualization at a glance the polymorphic sites.	43
13	List of fish species of different groups	45

LIST OF FIGURES

Figure	Captions	Page
1	Flowchart for DNA concentration	24
2	PCR amplification of COI gene by using FishF1 and FishR1 primer	25
3	Neighbor joining phylogentic tree of cn3 (sample collected from a supershop, sold as chewa *(Pseudapocryptes elongotus)* with other sequences downloaded from Gene Bank and BOLD systems based on CO1 gene.	34
4	Multiple sequence alignment of CO1 gene fragment of 2 substitute fish spp. Some representative polymorphic sites are indicated by nucleotide position	36
5	Comparison between colleted fish Macerel and Gene Bank reference *Rastrelligir kanagurta*	38
6	Sequencing alingement between two species	39
7	Pairwiese sequence alignment of sample lg4 with the sequence of *Pampus chinensis*.	40
8	Sequence alignment and species confirmation	41
9	Sequence alignment of two spp.	42
10	Sequence alignment of two spp and confirmation of actual species.	43
11	Multiple Sequence alignment among th sample sequence , sequence of sold species and indentified species after submission the sequence at BOLD database	44
12	Neighbor joining phylogenetic tree of all sequences	46
13	Graphical Representation of the percentage of whole fish	47
14	Graphical Representation of the percentage of Fillet fish	47

LIST OF FLOW CHART

Flow Chart No	Captions	Page No
1	DNA bar-coding process. Adapted from Floydetal. (2010) and Yancy (2007)	15
2	Flow Chart 2 working flowchart for DNA isolation	22
3	Flow Chart 3 working schedule after visualization of genomic DNA	26
4	Stylized supply chain featuring technology adoption for supply chain monitoring	52

Chapter 1

Introduction

1.1 Background

DNA bar coding first came to the attention of the scientific community in 2003 when Paul Hebert's research group at the University of Guelph published a paper titled "Biological identifications through DNA barcodes". They proposed a new system of species identification and discovery using a short section of DNA from a standardized region of the genome. That DNA sequence can be used to identify different species, in the same way a supermarket scanner uses the familiar black stripes of the UPC barcode to identify your purchases.

DNA barcoding seeks to advance both species identification and discovery through the study of patterns of sequence divergence in a standardized gene region. A segment near the 5′-terminus of the mitochondrial cytochrome *c* oxidase subunit 1 (COI) gene has been selected as the barcode region for members of the animal kingdom (Hebert *et al*. 2003). Its effectiveness has been validated for various animal groups and most investigated species (> 94%) possess distinct barcode 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. Most prior barcode 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; Hajibabaei *et al*. 2006).

Earlier DNA barcode studies revealed of sequence variation in local faunas (Hogg *et al* 2004; Hebert *et al.* 2005), but these are now leading to continental or global barcode campaigns for a few groups such as birds, fish and Lepidoptera (Marshall 2005). Although the efficacy of DNA barcoding has gained validation, prior work on mammals has been restricted to two studies of primate species, most represented by a single individual (Nelson *et al*. 2005). Bats (order Chiroptera) are an obvious target for analysis as approximately 20% (1116 of 5416) of all mammal species belong to this order (Huck *et al*.2005). Moreover, although most mammal species are thought to have been

described, the incidence of overlooked taxa is likely to be high within bats due to their cryptic behavior and morphology.

DNA barcoding are already a powerful tool for the identification of seafood to the species level. We conclude that barcodes have broad applicability for authenticity testing and the phylogeographic patterning of genetic diversity. It can also inform aspects of traceability.

DNA barcoding aims to provide an efficient 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 efficiency of the method hinges on the degree of sequence divergence among species and species-level identifications are relatively straight forward when the average genetic distance among individuals within a species does not exceed the average genetic distance between sister species. Fishes constitute a highly diverse group of vertebrate development. In this context, the identification of fish species is challenging and DNA barcoding provide new perspectives in ecology and systematics 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 regions. While humanity is facing increasing evidence of the erosion of earth's biodiversity, this approach is proving its effectiveness in characterizing the complexity of the biodiversity realm at a pace unequalled by other characters. The primary goal of DNA barcoding focus on the assembly of reference libraries of barcode sequences for known species in order to develop reliable, molecular tools for species identification in nature. Current results suggest that, in a large array of organisms, species are generally well delineated by a particular sequence or by a tight cluster of very similar sequences that allow unambiguous identifications.

With the aim of assigning specimens to known species based on molecular tags, a 648-bp segment 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 make 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 processes. 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 processes such as pseudogenes ontogenesis, introgressive hybridisation, and retention of ancestral polymorphism pose potential difficulties in capturing species boundaries using mtDNA sequences. The detection of mixed genealogy between closely related species has been previously estimated to occur in nearly 20 percent of the cases in the wild. Recent barcoding studies emphasised 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 relies on the sequencing and comparison of a standardized portion of the genome to aid in specimen identification and species discovery. The DNA barcoding method now represents the largest effort to catalogue biodiversity using molecular approaches. Although initially regarded as controversial, numerous cases have been reported where the analysis of DNA sequence variation in the cytochrome c oxidase subunit 1 (COI) region of mtDNA has proven highly effective for the delineation and identification of animal species in general and fish in particular.

DNA barcoding, the analysis of sequence diversity in a standardized gene region, has gained considerable validation as a tool for species identification and discovery. Several studies have demonstrated its effectiveness for identifying both marine and freshwater fishes provoking an effort to build a barcode library for all fish species. Currently, records are available for 41771 fishes, representing 6566 fish species on the Barcode of Life Data System, BOLD. DNA barcoding also provides an independent means of testing the validity of existing taxonomic systems, revealing cases of inappropriate synonymy or overlooked taxa. These results suggest that the species boundaries need to be examined for the heavily exploited populations targeted by the aquarium trade, to properly inform conservation strategies and planning.

Advocating the use of an easily characterized 648 bp fragment from the mitochondrial 5′ region of the cytochrome *c* oxidase subunit I (COI) gene for animal identification, the primary goal of barcoding focuses on the assembly of reference sequence libraries derived from expert-identified voucher specimens in order to develop reliable molecular tools for species identification in nature. Barcoding has been mischaracterized as molecular taxonomy, although it is not intended to replace classical taxonomy. Its

purpose is to facilitate species identifications by non-experts and to do so in a rapid and cost-effective manner. The effectiveness of barcoding has been demonstrated in diverse taxa, including springtails, spiders, butterflies, flies, bivalves, fishes, birds and mammals. Barcoding systems also now established for plants, macroalgae and bacteria (Khallaf *et al*. 2014).

The Fish Barcode of Life campaign (FISH-BOL) seeks to establish a standard reference sequence library for the molecular identification of fishes worldwide.

Over the past decade, DNA barcoding has played a facilitatory role for accurate identification of marine ichthyiofauna, thanks to the integration of molecular and traditional taxonomic methods such DNA-based method provides a robust and standardized approach for marine species identification, as witnessed by the remarkable boost of species identified as well as its use for various applications as for example fisheries and conservation programs. DNA barcoding has been adopted in numerous studies illustrating its speed, reliability and accessibility.

DNA barcoding exhibits a sufficient level of variation to discriminate among species. The key advantage of DNA barcoding is that once a 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 generaly 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 marker region (i.e., the COI region of mitochondrial DNA) to identify plants and animal species through DNA sequencing (Hebert *et al*. 2003).

Recently, DNA barcoding has been employed as a species identification tool for food authentication and safety concerns, including incorrect product labeling, DNA-based

methods can also be used to monitor illegal trading involving protected or endangered speciesor to identify the species origin of commercially processed food.

However, some of the processing and preservation methods used with seafood products are not conducive to DNA barcoding with the full-length target gene region. DNA degradation has been recognized as a considerable limitation in DNA-based analyses of these samples and PCR amplification of full-length (i.e., ~650bp) barcodes from moderately or highly processed samples is significantly challenging. In addition, processed seafood products often contain multiple additives, preservatives and flavors that may affect the quantity and quality of DNA extracted from these products.

Alternatively, a mini-barcoding approach, which focuses the analysis on shorter DNA fragments within the full-length barcode, has been shown to be effective in obtaining DNA sequence information from specimens containing degraded DNA. The sequencing information generated from a small (≥100bp) mini-barcode fragment of COI within the full-length DNA barcode region can provide the information required for identification of individual species with more than 90% species resolution. However, extensive mini-barcode primer development species has not been carried out. DNA barcoding, based on the sequencing of a short, standardized region of the cytochrome c oxidase I (COI) gene, gathered increasing attention as a broadly applicable tool for identifying an array of animal species, including fishes (Hebert *et al* . 2003). The utility of the method for fish species identifications is grounded on the premise that the COI sequence shows considerably greater inter- than intra-species variation, allowing for the differentiation of ca. 97% of fish species (Ward 2009) and often being more discriminatory than alternative DNA markers used for this purpose (Cawthorn *et al*. 2011).

Although some potential limitations of DNA barcoding have previously been recognised the method has more recently been validated for use in forensic and regulatory fields (Dawnay *et al*. 2007). Momentum for the initiative has further been aided by, inter alia, the establishment of the Consortium for the Barcode of Life (CBOL) – an international alliance that promotes global standards for DNA barcoding, the development of the Barcode of Life Database (BOLD, www.barcodinglife.org) – an online data management system that serves as a global repository for barcode sequences (Ratnasingham *et al*. 2007), as well as the emergence of numerous campaigns seeking to barcode all life on earth. The Fish Barcode of Life Initiative (FISH-BOL, http://www.fishbol.org) is one

such campaign aiming to assemble a COI-reference library for all fishes (Ward *et al*.2009), with over 10 000 of the ca. 32 000 fish species being barcoded to date (2014).

COI barcoding (Cawthorn *et al*. 2012) and other DNA markers (Vinas *et al*.2009) have recently been used to reveal disturbing rates of fish mislabeling (21–50%) in South Africa, with both studies generating considerable media attention and likely leaving some industry role players infuriated and even humiliated. Such responses, however, typi those surrounding any major food scandal, where the immediate effects are often perceived as negative but the ensuing ones are largely positive. Research of this kind raises awareness around pertinent concerns, compelling the entire industry to resolve the issues. While weaknesses are exposed that are inherent to modern food supply chains (e.g. complexity, traceability), areas are highlighted that need improvement, prompting authorities to step up check and revise regulations.

Empirical support for the barcoding concept ranges from studies of invertebrates (e.g. springtails and butterflies) to birds (Hebert *et al.* 2004; Hogg and Hebert 2004). However, the approach is not beyond controversy. For a barcoding approach to species identification to succeed, within-species DNA sequences need to be more similar to one another than to sequences in different species. Recent studies show that this is generally the case, but there are exceptions. Hybridization among species would create taxonomic uncertainty: mitochondrial DNA is maternally inherited and any hybrid or subsequent generation would have the maternal species DNA only. Here we examine whether barcoding can be used to discriminate fish species. There are probably close to 30 000 fish species worldwide, constituting about 50% of all vertebrate species. They are systematically very diverse, ranging from ancient jawless species (Agnatha: hagfish and lampreys) through to cartilaginous fishes (Chondrichthyes: chimaeras, sharks and rays) and to old and modern bony fish (Osteichthyes: coelacanths, eels,carps, tunas, flatfishes, salmonids, seahorses, etc.). In 2000, fisheries provided more than 15% of total animal protein to the global food supply, employed about 35 million people, and had an estimated first sale value of about US$81 billion (FAO 2002). Fish and fish products are important contributors to human food security. Accurate and unambiguous identification of fish and fish products, from eggs to adults, is important in many areas. It would enable retail substitutions of species to be detected, assist in managing fisheries for long-term sustainability, and improve ecosystem research and conservation. Hitherto, a wide

variety of protein- and DNA-based methods have been used for the genetic identification of fish species (Ward *et al* .2003).

The scale of species coverage envisioned and the subsequent scope of potential applications to be supported distinguish DNA barcoding from previous molecular approaches. A growing body of literature on DNA barcoding demonstrates that the relatively short fragment of COI used for barcoding contains enough variation to accurately identify a large variety of animals to the species level (Waugh, 2007). This includes both freshwater (Hubert *et al*., 2008) and marine fishes (Rock *et al.,* 2008).

1.2 Problem Statement

In Bangladesh fishes (indigenous and exotic; freshwater and marine as well) are cultured to a large extent all over the country. Moreover, a large amount of fishes are also landed from inland open water and marine capture fisheries. But still there's need to produce more fish to fulfill the demand of growing population. Keep the growing demand in mind, several fraudulence in various form (species substitution/mislabeling/capture of endangered species) have been found in fisheries marketing chain of Bangladesh.

Reasons for these fraudulences include high demand with limited supply, high profit incentive, an increase in international trade of processed foods and lack of regulation enforcement and implementation. Therefore, the detection of species substitution has become an important topic within the food industry and there is a growing need for rapid, reliable, and reproducible authentication to verify species in commercial fish and other aquaculture products. To avoid mislabelling and commercial fraud, the US Food and Drug Administration has compiled an online Regulatory Fish Encyclopedia (http://www.fda.gov) listing acceptable market names, isoelectric focusing protein electrophoresis patterns and DNA barcoding data. Another similar initiative is the European FishTrace Consortium (http://www.fishtrace.org).

Therefore, to implement laws to prevent product substitution, there is a need for sensitive and reliable analytical methods that can be applied to determine the species of a fish, even when no detectable external features are present. The effectiveness of fish conservation and management programs can also be improved which aid in the protection of aquatic habitats and endangered species.

Molecular tools are advantageous for fish and fish products identification because of large number of fish species from distinct live history stages (eggs, fry and adults) can be examined; in addition, processed fish products lacking the morphological characteristics, such as frozen fillets and precooked fish, are also accessible (typically, these cannot be identified using the traditional identification procedure).

DNA Barcoding has the potential to be used with heavily processed food samples by using short mini barcode regions (Rasmussen *et al*. 2009). So, DNA barcoding has the potential to be a practical method for fish species identification to control of trade in endangered species, inspections of markets and products, improvements in traceability and identify undesirable animal or plant material in processed fish products. Assume there are two types of suppliers in a market for wild-caught salmon — those who supply genuine wild (Pacific) salmon, and those who supply farmed (Atlantic) salmon mislabelled as wild. It is assume also that consumers cannot differentiate between wild-caught and farm-raised prior to purchase.

With importation and consumer consumption of seafood increasing, a growing number of fish species are being encountered in the market as a result of increased demand and the globalization of the seafood industry. Subsequent economic deception and food safety concerns are pushing the need for accurately labeled food products and full disclosure of product composition. A dramatic increase in media coverage involving cases of market substitution demonstrates that high quality, nutritious and 'eco-friendly' food items are now a focal point for the educated consumer.

In this regard, the authenticity and certification of fish products is particularly important when fresh or frozen cuts of fish are encountered because misrepresentation of the actual product, whether through intentional or non-intentional mislabeling, is known to occur (Marko *et al*. 2004). Unfortunately, consumers are unable to detect these cases given that recognizable external morphological features are typically removed when the fish is filleted or otherwise processed. The lack of morphological features that are traditionally used to identify animal species is a common problem with food products, making authenticity tests impossible without alternative identification methods. Molecular diagnostic techniques have proven to be effective species identification tools, capable of bypassing the inherent problems of morphology-based identification methods.

However, early macromolecular techniques, such as electrophoretic and immunological identification (Rehbein *et al*. 1990), exhibited limitations of their own. For example, protein of interest often denature during heating and/or processing, are tissue- specific and are prone to contamination (Hofmann, 1987; Patterson & Jones, 1990), making these methods challenging to interpret and difficult to replicate. Today, DNA-based methods are more frequently employed for food authentication (Lockley *et al*. 2000). As with past electrophoretic and immunological methods, the use of DNA allows identification to proceed on samples lacking diagnostic morphological feature.

1.3 Rationale

Fraudulence in fish marketing is very common in Bangladesh. Several reports have been published on the daily newspapers about this matter, for example, piranha (*Pygocentrus nattereri*) is serving in restaurant as rupchanda (*Pampus argenteus*); hilsa/Jatka (*Tenualosa ilisha*) sold in the market as chapila (*Gudusia chapra*); 'barmis rui' found in the local market and sold as 'deshi rui' etc. The problem is more pronounced in case of processed or semi-processed fish products (fish fingers/fish sticks) as several super markets and fish shops now-a-days sell pellet fish rather whole fresh fish.

Unfortunately, consumers are unable to detect these cases given that recognizable external morphological features are typically removed when the fish is filleted or otherwise processed. The lack of morphological features that are traditionally used to identify animal species is a common problem with food products, making authenticity tests impossible without alternative identification methods. As per we know, in Bangladesh no works have been done on fraudulence detection in fish marketing using DNA Barcoding.

Therefore, the proposed work will provide a clear indication about the market substitution of fresh fish/fish products in aquaculture marketing of Bangladesh. Cases of fraud in fish markets have garnered increasing public attention recently, raising concerns for and is traded in high volumes across the world (Anderson 2003), with about one third of global production traded across national boundaries (FAO 2012). The growth in fish consumption and trade has been linked to increasing consumers' awareness of the health benefits associated with fish (Tveterås *et al*. 2012).

The increased consumption of fish and the economic gains accruing to sellers of high-valued species have served as incentives for some firms within the fish supply chain to consciously misrepresent their products and substitute substandard products for economic gain. Results of studies in various countries show that mislabeling and substitution in fish supply chains have been on the increase relative to other products (Pepe *et al*. 2007; Miller *et al*. 2010). Species substitution and mislabeling has a number of negative outcomes: sellers of high-valued fish products incur losses, while consumers pay a premium for low- valued fish as they cannot differentiate fillets or species with similar morphological features prior to or, in some cases, even after purchase. Alternatively, consumers may incur transaction (search) costs in seeking information about a product's true quality.

Although it is not always the case, substitution and mislabeling in fish markets can create food safety hazards. An example is the consumption of toxic puffer fish mislabeled as monk fish in the USA in 2007, which left many consumers sick (Leschin-Hoar 2011). Mislabeled fish also create potential health hazards for consumers with allergies to specific types of fish. The inability or low accuracy of the traditional methods of fish species identification (e.g., use of morphological features), especially in filleted or processed forms, potentially diminishes sellers' incentives to "play by the stipulated rules".

These problems have prompted firms to invest in developing traceability and authenticity technologies to exploit market opportunities for verifiable authenticity assurances (Kemp 1994), as well as to protect their reputations. Furthermore, some consumers and consumers' organizations are seeking authenticity assurances for products traded in local markets, includ- ing those that arrive as a result of international trade. This may motivate the uptake of authenticity technologies by a third party, such as an environmental or consumer group as a means of monitoring, to verify quality claims and authenticity within fish supply chains.

However, authenticity does not necessarily imply full traceability although the two concepts are related. Fish authentication processes may assist in reducing market failures due to information problems, fraud, health risks, and unfair market competition between genuine and fraudulent producers and (or) sellers. likely to offer their products for sale.

The effect of species substitution and mislabeling in fish markets is examined using insights from the economics of information asymmetry and market failure.

Using DNA barcoding technology to verify authenticity is a potential solution. To audit and prevent species fraud on the commercial market, a number of molecular methods have been developed, including use of a unique protein or DNA profiles found in different species DNA barcoding provides a rapid, cost-effective method for accurate identification at the species-level through comparative analysis of sequence variation in a short, standardized fragment of the genome.

Although 'food fraud'' has been carried out since antiquity, these practices seem to have escalated in recent years. High-value, protein-rich foods are especially prone to substitution or mislabeling, as exemplified by the Chinese melamine saga of 2008 (Jha *et al*. 2010), the 2013 meat scandals in South Africa and the EU (Cawthorn *et al*. 2013) and the many documented cases of seafood fraud. While the former two examples were generally sporadic, seafood mislabeling has been a persistent and widespread problem, apparently intensifying in synchrony with the ever-declining state of the world's fish stocks. The aim of this study was to assess the current extent of fish misnaming or mislabeling in Bangladesh at the final supply chain link (consumer level) and to reconcile the results with previous studies.

1.4 Objectives

The overall objective of the thesis work was to identify the level of fraudulence in fish marketing in Bangladesh.

The study has been conducted for achieving following specific objectives

1. To evaluate the ability of DNA Barcoding to identify the species of fish in fresh or in processed products acquired directly from commercial markets and restaurants.
2. Comparison of the BOLD and Gene Bank databases to evaluate their relative performance in generating positive matches for species identification.
3. DNA barcoding for the detection and quantification of mislabeling in commercial raw fish fillets purchased from Bangladeshi markets.

Chapter 2

Materials and Methods

2.1 Experimental fish

In the present study, 20 different species of fish were examined. The samples were collected from different supermarkets and local markets (Agora, Meena Bazar, Prince bazer, Swpana, Nandan, Mottjhel, Gopibag , Jatrabari) of Dhaka metropolis. Samples from muscle tissue from dorsal portion were examined for each specimen.

Table 1: List of fish species used in this stuy

No of fish	Scientific Name	English Name	Local Name	Collection Place
1	*Saddinella loniceps*	Sardin	Sardin	Nandan
2	*Rastrelliger kanagurta*	torpedo scad	Surma	Nandan
3	*Pseudocryperyptes elongates*		Chewaa	Nandan
4	*Pampus chiensis*	Robusyt	Lal Chanda	Gopibag
5	*Clarius batracus*	Walking catfish	Magur (Felet)	Agora
6	*Eutropichthys vacha*	ray-finned fish	Bacha Fish	Prince Bazar, Mirpur
7	*Gudusia chapra*	Indian river shad	Chapila	Gopibag
8	*Gudusia chapra*	Indian river shad	Chapila	Jatrabari
9	*Bagarius barius*	Giant catfish	Bagha Ayer	Jatrabari
10	*Tor putitora*	Putitor Mahasher	Mohashol	Jatrabari
11	*LAbeo bata*	Bata	Bata	Motijhil
12	*Mystus aor*	Gaint catfish	Ayer(fillet)	Agora(gulsan)
13	*Wallago attu*	Wallago	Boal	Jatrabari
14	*Rastrelliger kanagurta*	Scsad	Cube mackerel	Swapnna
15	*Anabas testudineus*	Climbing perch	Desi koi	Motijhil
16	*Anabas testudineus*	Climbing perch	Koi	Dilkusha
17	*Lates calcarifer*	Barramundi	Koral	Hatirpul
18	*Bagarius barius*	Giant cat fish	Bagha ayer	Prince bazzer
19			Kalom ilish	Palashi
20	*Parastromateus niger*	Black pomfret	Black rupchanda	Dhanmondi

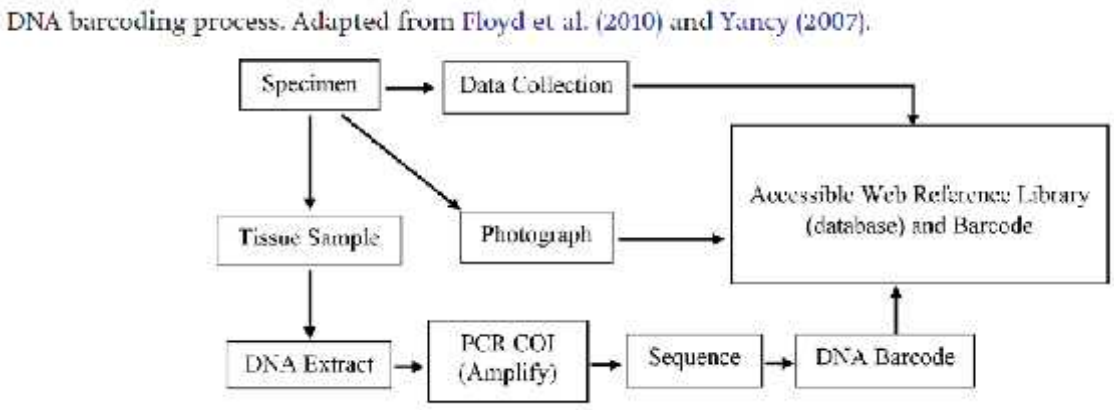
2.2 Sampling

Fish samples were collected in sterilized plastic bag from different fish markets of Dhaka metropolitan city early in the morning during the periods April 2015 to June 2015. Samples were then transported in the laboratory, using icebox within required time.

2.3 Processing of samples

The fish samples were processed within 2 h of collection following aseptic techniques. First, the samples were cut by scissor. Then the tissue samples were collected aseptically. The collected samples were separately packaged for further work.

Flow Chart 1: DNA barcoding process. Adapted from Floydetal *et al.* (2010)



2.4 Isolation of DNA from fish muscle tissue

2.4.1Chemicals used for isolation of muscle DNA

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

a) 1 M Tris-HCl; pH 8.0 (pH was adjusted with concentrated HCl)
b) 0.5 M EDTA; pH 8.0 (pH was adjusted with concentrated NaOH)
c) Phenol : Chloroform : Isoamyl alcohol (25:24:1 ; v/v)
d) TE buffer; pH 8.0
e) RNase solution (DNase free) 10 µg/ml (stock 10 mg/ml)
f) 70% Ethanol
g) Liquid Nitrogen
h) Isopropanol

i) DTT(dithiothreitol, clelands reagent)

j) SET isolation buffer (per 100 ml)

10% Sarcosyl	20 ml
5 M NaCl	2 ml
0.5 M EDTA	20 ml
50 mMTris-HCl	5ml

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

2.4.2 Required materials for agarose gel electrophoresis

The following equipment's and chemicals were used to conduct agarose gel electrophoresis:

a) A horizontal electrophoresis chamber and power supply

b) Gel casting tray and combs

c) Gel Documentation System

d) Gloves

e) Pipette and tips

f) DNA markers:

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

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

g) Electrophoresis buffer (TAE pH 8.0)

h) 6× sample loading buffer

i) Agarose

j) DNA stain (Ethidium bromide)

Table 2: Random primers used in the present study for screening

Primer code	Sequence (5′—3′)	G+C content (%)
Fish F1	TCAACCAACCACAAAGACATTGGCAC	46.154
Fish F2	TCGACTAATCATAAAGATATCGGCAC	38.462
Fish R1	TAGACTTCTGGGTGGCCAAAGAATCA	46.154
FishR2	ACTTCAGGGTGACCGAAGAATCAGAA	46.154

2.4.3 Preparation of Stock and Working Solutions used for genomic DNA Isolation:

For conducting the isolation procedures, the following stock solutions and working solutions were prepared.

2.4.3.1 1 M Stock Solution of Tris-HCl pH 8.0 (100 ml)

12.14 g of Trizma base (MW=121.14) was dissolved in 75 ml of distilled water. The pH of this solution was adjusted to 8.0 by adding about 5 ml of concentrated HCl. The volume of the solution was adjusted to a total of 100 ml with de-ionized distilled water. Then it was sterilized by autoclaving and stored at 4 ºC.

2.4.3.2 0.5 M Stock Solution of EDTA pH 8.0 (100 ml)

18.61 g of EDTA (EDTA.$2H_2O$, MW = 372.24) was added to 75 ml of distilled water and stirred vigorously with a magnetic stirrer. Approximately 2 g of NaOH pellets was added to adjust the final pH to 8.0. The final volume of the solution was adjusted to 100 ml by adding sterile de-ionized distilled water. The solution was sterilized by autoclaving and stored at 4 ºC.

2.4.3.3 5 M Stock Solution of NaCl (100 ml)

29.22 g of sodium chloride (NaCl, MW = 58.44) was dissolved in 75 ml of distilled water. The total volume of the solution was adjusted to 100 ml with distilled water. The solution was heated in an oven for 15 seconds and stirred vigorously on a magnetic stirrer to dissolve NaCl. It was then sterilized by autoclaving and stored at 4 ºC. NaCl was added in small amount at once to be dissolved in solution.

2.4.3.4 Ribonuclease A stock solution:

10 mg RNase A was dissolved in 1 ml of deionized distilled water and store in 20 ^{0}C.

2.5.3.5 Tris-HCl Saturated Phenol

It was prepared following procedure described below,

a) The crystal phenol was melted in a water bath at 65 ºC for 30 minutes.
b) Then 100 ml of melted phenol was taken and same volume of Tris-HCl (pH 8.0) was added.
c) It was mixed with a magnetic stirrer for 10 minutes and then was left in rest for 5 minutes.

d) At this stage, two distinct phases were visible, colorless upper phase and colored lower phase.
e) The upper phase was removed with the help of a dropper.

This step was performed for six times which took about 3.5 hours to obtain pH 7.75. After saturation, the phenol became the half of the initial volume. As phenol is very much corrosive and highly toxic, protective measures like wearing of Apron, Gloves and Musk were adopted during the whole process.

2.4.3.6 Phenol: Chloroform: Isoamyl alcohol (25: 24: 1) (100 ml)

50 ml of Phenol, 48 ml of Chloroform and 2 ml of Isoamyl alcohol were added and mixed properly using vortex mixture. Mixing was done under fume hood for ensuring safety. The solution was then stored at 4 ^{0}C and was shaken before every use. The Phenol:Chloroform:Isoamyl alcohol mixture is caustic and produces fumes. So, this solution was used only within the fume hood wearing gloves and eye protection.

2.4.3.7 70% Ethanol (100 ml)

30 ml double distilled water was added in 70 ml absolute ethanol to prepare 100 ml 70 % ethanol.

2.4.3.8 Stock Solution of TE (Tris-HCl EDTA) Buffer pH 8.0 (100 ml)

1 ml of 1 M Tris-HCl was added with 0.2 ml (200 µl) of 0.5 M EDTA. The final volume was adjusted to 100 ml with sterile de-ionized distilled water. The solution was sterilized by autoclaving and stored at 4 ^{0}C.

2.4.3.9 3 M Sodium acetate pH 5.2 (100 ml)

40.824 g of sodium acetate with 70 ml of ddH_2O and adjusted the final volume to 100 ml with ddH_2O and p^H was adjusted to 5.2. Then it was sterilized by autoclaving.

2.4.3.10 Preparation of Extraction Buffer

To prepare extraction buffer the following components and concentrations were used. Considering the economic use of chemicals, different volume of solutions were prepared as mentioned in the following table 2.

Steps of Extraction Buffer Preparation (100 ml)

Extraction Buffer was Prepared following the steps below (100 ml):

1. 5 ml of 1 M Tris-HCl (pH 8.0) was taken in a 250 ml conical flask.
2. 2 ml of 5 M NaCl was added to it.
3. 20 ml of 0.5 M EDTA (pH 8.0) was added next.
4. 20ml of sarcocyl (10%) was added.
5. The pH of the solution was adjusted to pH 5.0 with HCl and was made up to 100 ml by adding sterile de-ionized distilled water.

Table 3: Volume of solutions to prepare extraction buffer

Chemical Names	Molecular Weight	Stock Conc.	Reference Con./working conc.	Working Volume (100 ml)
NaCl	58.44	5 M	100mM	2 ml
EDTA (pH 8)	372.24	0.5 M	100 mM	20 ml
Tris-Base (pH 8)	121.1	01 M	50 mM	5 ml
Sarcosyl		10 %	10 %	20ml

2.5 Genomic DNA isolation

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

1. 250mg freshly collected skeletal 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 Eppendorf tube (2.0 ml) and incubated at 60 ^{0}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 DTT and followed by incubation at least 6 hours to dissolve the muscle tissue.
5. The samples were centrifuged at 4000 rpm for 20 min at room temperature to remove non-soluble debris.

6. The supernatants were transferred to fresh Eppendorf tube and an equal volume of Phenol:Chloroform:Isoamyl alcohol (25:24:1) was added and mixed well by slow inversion, then centrifuged the tubes at 6500 rpm for 10 minutes. This step was repeated for two times.
7. The supernatant was collected and 1/10 vol of 3M sodium acetate (ph 8.0) and .6 vol of chilled Isopropanol were added and mixed slowly to precipitate DNA (Sambrook *et al*.,1987) . Mixed the contents and kept into refrigaretion 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 °C for 2 hours.
11. The was prepicited with isopropanol and sodium acetate and further washed with 70% ethanol and pure DNA pellet was re-dissolved in 100µl of TE buffer.

2.5.1 Estimation of quality and quantity of isolated DNA samples

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 assay. Excessive genomic DNA may result smears and may not produce clearly defined bands in the gel. On the other hand, too little DNA will give non-reproducible patterns (Williams *et al*. 1993). Measurement of isolated DNA concentration can be done by comparing DNA with the standard DNA on agarose gel electrophoresis or by estimating the absorbance of DNA by spectrophotometer (SPECORD 50, Analytikjena, Germany) at 260 nm. Both the methods were carried out in this investigation.

2.6 Preparation of Stock Solutions Used for Gel Electrophoresis

For conducting the gel electrophoresis, the following stock solutions and other solutions were prepared.

2.6.1 50 × TAE Buffer (pH 8.3) (1 liter)

242 g Trizma base (MW=121.14) was dissolved into 900 ml of sterile de-ionized distilled water. Then 57 ml glacial acetic acid was added to the solution. Finally, 100 ml 0.5 EDTA (pH 8.0) was added in it. They were mixed well. The pH of the solution was adjusted by mixing concentrated HCl at pH 8.3. The final volume of the solution was adjusted to 1000 ml.

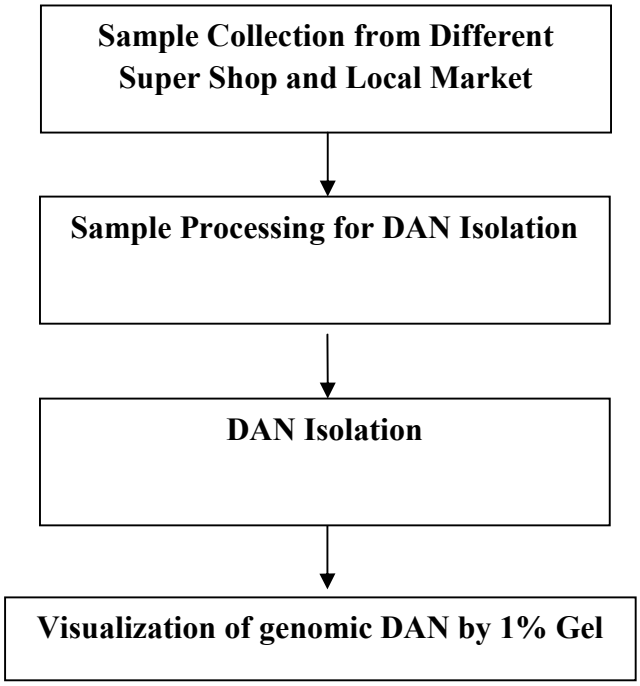
2.6.2 Ethidium Bromide Solution

For 1ml solution, Ethidium Bromide 10 mg was added to 1 ml of sterile de-ionized distilled water. It was then mixed by hand shaking. The solution was then transferred to a dark bottle and stored at room temperature. Stock solution of 10 mg/ml can also be purchased directly from companies.

2.6.3 Agarose Gel Electrophoresis

i. 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.
ii. 1.0 g of agarose was heated to melt into 100 ml of TAE buffer, ethidium bromide was added (10 μg/ml) and poured into gel casting tray fixed with appropriate combs.
iii. After the gel was solidified it was placed into gel-running kit containing 1xTAE buffer.
iv. Digested plant DNA solutions were loaded with 6x gel loading dye and electrophoresis was continued until DNA fragments were separated well.

Flow Chart 2 Working flowchart for DNA isolation



2.7 Preparation of PCR Reaction Mixture

The following components were used to prepare PCR reaction mixture (Table 4). The total volume of PCR reaction mixture was for 17 samples.

Table 4: Component of PCR Reaction Mixture (for 17 reactions)

Sl. No.	Reagents	Amount per sample	Total
1	Sterile de-ionized distilled water	18.7 µl	317.9 µl
2	*Taq* Buffer A 10X (Tris with 15 mM $MgCl_2$)	2.5 µl	42.5 µl
3	Primer	1.0 µl	17 µl
4	dNTPs 10 mM	0.25 µl	4.25µl
5	*Taq* DNA Polymerase 5U/µl	0.05 µl	.85 µl
6	Template DNA	2.5 µl	42.5µl
	Total	**25.0 µl**	**425 µl**

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

2.7.1 PCR Amplification

PCR amplification was done in an oil-free thermal cycler (Biometra, UNO II). The optimum amplification cycle was as follows:

	Initial denaturation	94 ºC	for	4 minutes
32 Cycles	Denaturation at	94 ºC	for	45 second
	Annealing at	54 ^{o}C	for	1 minutes
	Extension at	72 ºC	for	1 minutes
	Final extension at	72 ºC	for	5 minutes

After completion of cycling programme, the reactions were held at 4 ºC.

2.7.2 Electrophoresis of the amplified products and documentation

The amplified products were separated electrophoretically on 1 % agarose gel. The gel was prepared using 1.0 g agarose powder containing ethidium bromide and 100 ml 1×TAE buffer. Agarose gel electrophoresis was conducted in 1× TAE buffer at 90 Volts for 30 minutes. Molecular weight marker 1kb plus or 1Kb DNA ladder was electrophoresed alongside the reactions. DNA bands were observed on UV-transilluminator and photographed by a Gel Documentation system.

2.8 PCR product purification

1. PCR product has transferred to 1.5ml eppendrof.
2. 5 times FADF buffer added and centrifuged at 11000 rpm for 1 min
3. Flow through discarded and 750 ml was buffer added and centrifuged at 11000 rpm and liquid discarded.
4. again centrifuged at 14000 rpm for 10 minutes.
5. Then 40µl of Elution buffer added to the membrane centre of FSDF colum and waited for 3 minutes for fully absorption.
6. Centrifuged at full speed at 14000 rpm for 2 minutes to elote the DNA.
7. Then the liquid further placed in colum tube and centrifuged at 14000 rpm for 2 minutes.

2.9 Quantification of DNA concentration

1. Qubit™ working solution was prepared by diluting the Qubit™ reagent 1:200 in Qubit™ buffer and 200 µl of working solution prepared for each standard and sample
2. Assay tubes prepared according to the table below

	Standard Assay Tubes	**user sample Assay Tubes**
Volume of working solution (From Step 1) to add	190	180-190
Volume of standard (from kit) to add	10	
Volume of user Sample to add		1-20
Total Volume in each Assay tube	200	200

3. Vortexed all tubes for 2 to 3 seconds.
4. Incubated the tubes for 2 minutes at room temperature (15 minutes for the Qubit™ protein assay).
5. Inserted the tubes in the Qubit® 2.0 Fluorometer and taken readings.
6. Fluorometer, determine the stock concentration of the original sample

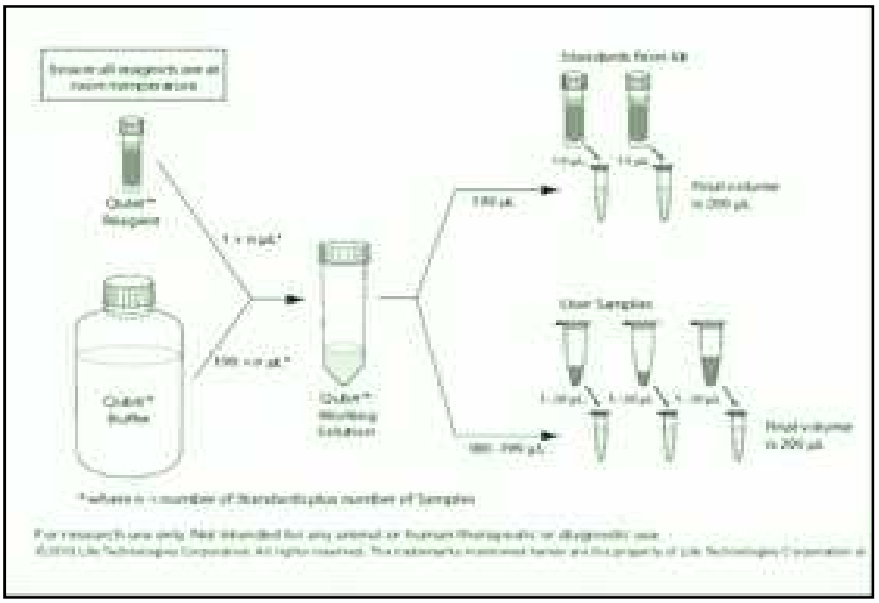


Figure 1. Flowchart for DNA concentration measurement

Table 5: Measurement of DNA concentration of 10 samples

Sample ID	Concentration of DNA (ng/µl)
sn1	19.9
mm2	15.9
cn3	11.3
lg4	19
ma5	13
bp6	8.78
cg7	7.76
cj8	9.18
bj9	7.56
ta10	22.4

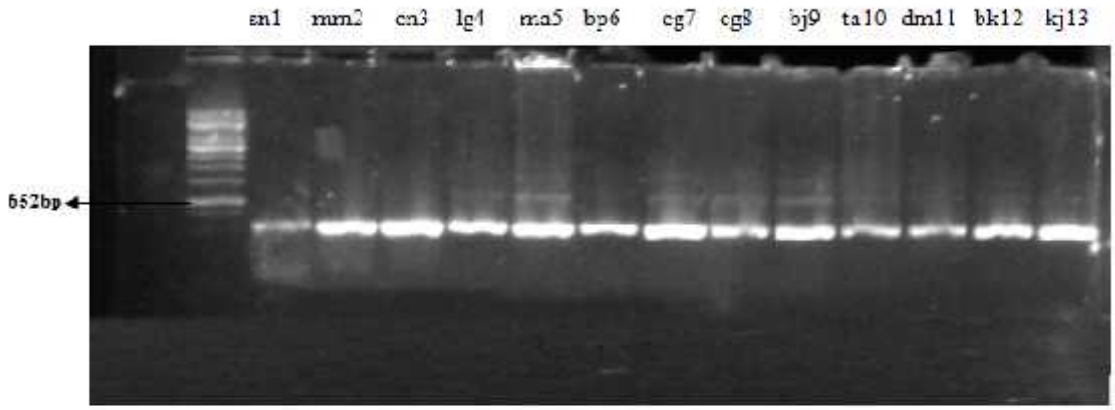
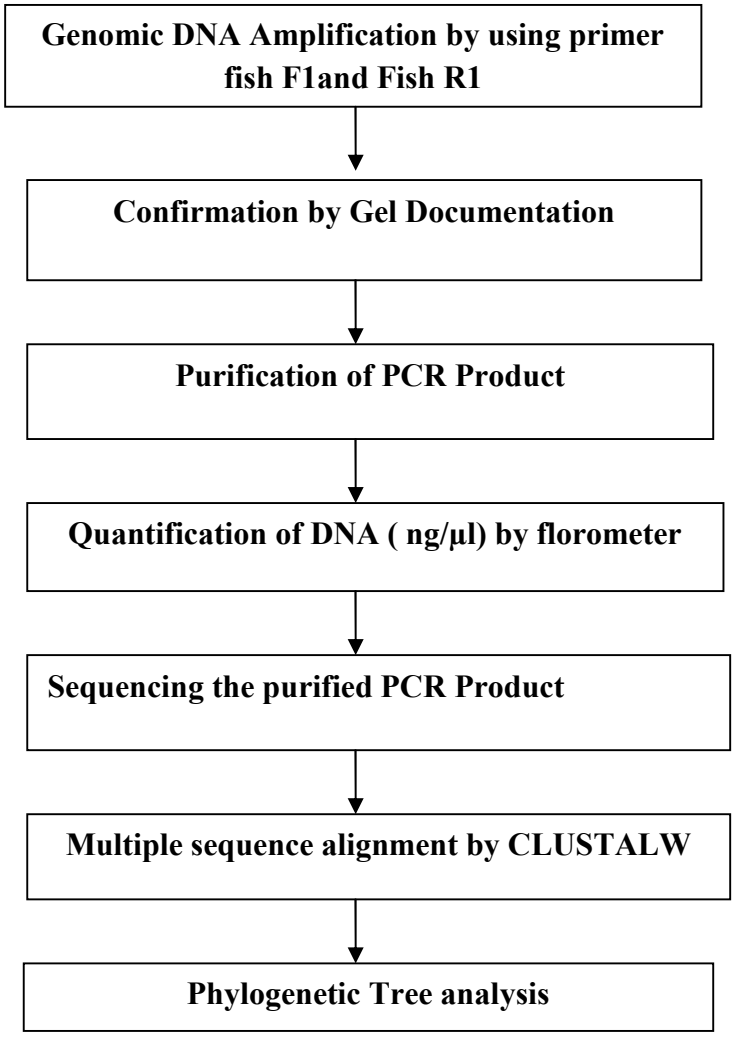


Figure 2. PCR amplification of COI gene by using FishF1 and FishR1 primer

Flow Chart 3 Working schedule after visualization of genomic DNA



2.10 Software use for sequence analysis

2.10. 1 NCBI BLAST

BLAST is like doing an experiment to get good, meaningful results, everyone needs to optimize the experimental conditions. BLAST on Windows Azure is a cloud-based implementation of the Basic Local Alignment Search Tool (BLAST) of the National Center for Biotechnology Information (NCBI). BLAST is a suite of programs that is designed to search all available sequence databases for similarities between a protein or DNA query and known sequences. BLAST allows quick matching of near and distant sequence relationships, providing scores that allow the user to distinguish real matches from background hits with a high degree of statistical accuracy. Scientists frequently use such searches to gain insight into the function and biological importance of gene products.

BLAST on Windows Azure extends the power of the BLAST suite of programs by allowing researchers to rent processing time on the Windows Azure cloud platform. The availability of these programs over the cloud allows laboratories, or even individuals, to have large-scale computational resources at their disposal at a very low cost per run. For researchers who don't have access to large computer resources, this greatly increases the options to analyze their data. They can now undertake more complex analyses or try different approaches that were simply not feasible before.

2.10.2 BOLD Database

Barcode of Life Data Systems (commonly known as **BOLD**) is a sequence database specifically devoted to DNA barcoding. It also provides an online platform for analyzing DNA sequences (Sujeevan *et al* 2007). As of 2013, BOLD included over 2.5 million DNA barcode sequences from over 190,000 species (Mark, 2013).

The Barcode of Life Data System (bold) is an informatics workbench aiding the acquisition, storage, analysis and publication of DNA barcode records. Bold is freely available to any researcher with interests in DNA barcoding. By providing specialized services, it aids the assembly of records that meet the standards needed to gain BARCODE designation in the global sequence databases. Because of its web-based delivery and flexible data security model, it is also well positioned to support projects that involve broad research alliances.

2.10.3 MEGA 6.06

The Molecular Evolutionary Genetics Analysis (MEGA) software is developed for comparative analyses of DNA and protein sequences that are aimed at inferring the molecular evolutionary patterns of genes, genomes, and species over time (Kumar *et a*l. 1994; Tamura *et al*. 2011). MEGA is currently distributed in two editions: a graphical user interface (GUI) edition with visual tools for exploration of data and analysis results (Tamura *et al*. 2011) and a command line edition (MEGA-CC), which is optimized for iterative and integrated pipeline analyses (Kumar *et al*. 2012).

In version 6.0, we have now added facilities for building molecular evolutionary trees scaled to time (timetrees), which are clearly needed by scientists as an increasing number of studies are reporting divergence times for species, strains, and duplicated genes (e.g., Kumar and Hedges 2011; Ward *et al*. 2013).

2.10.4 CLUSTALW

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

Chapter 3
Results

Full-length DNA barcodes (656 bp) were recovered using the Fish primers Fish F1 and FishR1. DNA of 17 out of 20 fish sample obtained from the Super Shop was amplified using the primer set and 10 were used for bi-directional sequencing (both forward and reverse). Peak intensities and sequencing qualities of the generated barcodes were compared to the sequences downloaded from NCBI GeneBank and BOLD databases.

Table 6: List of all identification results using the Gene Bank as search engine

Sample No	**Sold as**	**Type**	**Gene Bank Refference**	**Gene Bank accession no**	Mislabled
Sn1	Sardine	Whole Fish	*Megalaspis cordyla (97%)*	KM522836.1	Yes
			Megalaspis cordyla (95%)	HQ560952.1	
			Megalaspis cordyla (95%)	JX261015.1	
			Megalaspis cordyla (93%)	HQ149881.1	
mm2	Mackerel	Whole Fish	*Megalaspis cordyla (93%)*	KM522836.1	Yes
			Megalaspis cordyla (93%)	HQ560952.1	
			Megalaspis cordyla (93%)	JX261015.1	
			Megalaspis cordyla (93%)	HQ149881.1	
Cn3	Chewa	Whole Fish	*Scartelaos gigas (97%)*	KT277705.1	Yes
			Bolephthalmus (95%)	KP277118.1	
			Pseudocorymopoma doriae (94%)	JX983442.1	
			Pseudocorymopoma doriae (94%)	JX983441.1	
Lg4	Lal chanda	Whole Fish	*Piaractos brachypomus (96%)*	HQ420838.1	Yes
			Seriolella porossa (96%)	KM435146.1	
			Piaractos mesopotamicus (96%)	KM245046.1	
			Piaractos mesopotamicus (96%)	KM897305.1	
Ma5	Desi magur	Fillet	*Heteropneustes fossilis (98%)*	AP012013.1	
			Heteropneustes fossilis (98%)	KT001154.1	
			Heteropneustes fossilis (96%)	KT364787.1	
			Heteropneustes fossilis (94%)	JQ466398.1	
Bp6	Bacha	Whole	*Eutropichthys vacha (98%)*	AB919123.1	No

		Fish	*Eutropichthys vacha (91%)*	JN228951.1	
			Clupisoma prateri (94%)	JX983273.1	
			Clupisoma prateri (92%)	JN628921.1	
Cg7	Chapila	Whole Fish	*Seriolella porossa (97%)*	KM435146.1	Yes
			Piaractos mesopotamicus (96%)		
			Tenualosa ilisha (97%)	AP011611.1	
			Tenualosa toli (91%)	JX 983317.1	
Cj8	Chapila	Whole Fish	*Tenualosa toli (97%)*	AP00600.1	Yes
			Tenualosa toli (93%)	JX 98317.1	
			Tenualosa ilisha (97%)	AP011611.1	
			Tenualosa ilisha (97%)	AP011610.1	
Bj9	Bagha ayer	Fillet Fish	*Bagarius bagarius (94%)*	EU417762.1	N0
			Bagarius bagarius (92%)	FJ459434.1	
			Bgarius yarrelli (98%)	JQ026260.1	
			Bagarius bagarius (91%)	JN815268.1	
Ta10	Mohashol	Fillet Fish	*Elopichthys bambusa (96%)*	KM19612.1	Yes
			Squaliobarbus curriculus (96%)	KP731975.1	
			Hypothalmicthys molitrix (96%)	KP01119.1	
			Hypothalmicthys molitrix (96%)	KJ746961.1	

3.1 The effect of species substitution and mislabeling in fish market

The effect of species substitution and mislabeling in fish markets is examined using insights from the economics of information asymmetry and market. Information asymmetry arises when one party to a transaction has more information about the true quality of a good than another party. The price buyers are willing to pay is lower, reflecting this quality uncertainty. At this lower price, sellers of high quality goods are less likely to offer their products for sale; the market becomes dominated by "low quality" goods.

A number of solutions to information problems exist, including product warranties, as well as improved quality signalling prior to purchase such as through third party quality verification. Applying these insights to the context of fish substitution and mislabeling, in the absence of quality verification, a potential information asymmetry problem exists wherein sellers have more information about the true quality (species) than buyers. Using DNA barcoding technology to verify authenticity is a potential solution. Table7 illustrates these concepts in a situation where mislabelled Surma fish enter the market as Mackerel and Sardin. 8 different species with similar morphological features but which may differ in perceived quality. Quality in this context reflects differences in market prices.

3.2 DNA sequence interpretation

DNA sequencing results obtained from the fish samples collected from super shop and local fish market. Of the 20 acquired samples, the DNA from 19 (65%) was successfully amplified with the FishF1 and FishR1 primer. The resulting PCR products were sequenced to produce full length DNA barcodes averaging 656 base pairs (bp) in length, with no detectable insertions, deletions or stop codons (Fig. 2).

Among them 1 samples, constituting products marketed did not amplify with the primers. This amplification failure could most likely be attributed to DNA degradation or the presence of PCR inhibitors in the samples. There may be some concern as to whether the DNA was degraded and unrecoverable due to long term preservation.

Table 7: Identification of Collected Samples using the Gene Bank and Bold Search Engines

serial	sample ID	Sold as	Description	Max score	Total score	Query cover	E value	Identification	Accession	Bold reference database	Mislabeled
1	sn1	*Sardenella longiceps*	*Megalaspis cordyla*	1184	1184	100%	0	99%	KM522836.1	ABFJ246-07	**Yes**
2	mm2	*Rastreliger kanagurta*	*Megalaspis cordyla*	1184	1184	100%	0	99%	KM522836.1	ABFJ246-07	**Yes**
3	cn3	*Pseudapocryptes elongates*	*Scartelaos gigas*	686	686	100%	0	86%	KT277705.1	No sequence	**Yes**
			Taenioides nigrimarginatus	686	686	100%	0	86%	KJ865407.1	GBGCA11774-15	
4	lg4	*Pampus chinensis*	*Piaractus mesopotamicus*	1206	1206	100%	0	99%	HQ420833.1	ANGBF6848-12	**Yes**
5	ma5	*Clarius batrachus*	*Heteropneustes fossilis*	1184	1184	100%	0	99%	AP012013.1	ANGBF10687-12	**Yes**
6	bp6	*Clupisoma prateri*	*Clupisoma prateri*	1173	1173	99%	0	99%	JX983274.1	No sequence	No
7	cg7	*Gudusia chapra*	*Tenualosa ilisha*	1201	1201	100%	0	99%	AP011611.1	CYTC3774-12	**Yes**
8	cj8	*Gudusia chapra*	*Tenualosa ilisha*	1201	1201	100%	0	99%	AP011611.1	CYTC3774-12	**Yes**
9	bj9	*Bagarius bagarius*	*Bagarius bagarius*	1179	1179	97%	0	99%	EU417762.1	ANGBF5989-12	No
10	ta10	*tor putitora*	*Mylopharyngodon piceus*	1201	1201	100%	0	99%	HQ236003.1	GBGC6787-09	**Yes**

Note: 8 out of 10 samples were found as mislabeled, 71.42% mislabeled were found for whole fish and 33.33% mislabeled found for whole fish.

For the 19 samples yielding interpretable COI barcodes, maximum sequence similarity values of P94% were achieved in BOLD and/or GenBank and the top species identifications.

Specimen (sn1), sold as 'Sardin' (*Sardinella longiceps*) but identified in Gene Bank as (*Megalaspis cordyla)* which accssiovn no. KM522836.1 according to Gene Bank reference and accession no. ABFJ246-07 according to BOLD data base reference and the sold spp is 99%similiar with *Megalaspis cordyla* . *Sardinella longiceps* found in Indian ocean espiacally in the northern and southern parts . Highly school forming fish and feeds on phytoplankton and small crustracian. Marketed fresh, salted , dried salted.Also sold as smoked and canned fish. Sardine fishery is highly commercial. On the other hand *Megalaspis cordyla* pelagic species and generally occurs in inshore areas of the continental shelf and feeds mainly on fish.commercial value is lower than the (*Sardinella* longicep (Plate 1, FigA1) and it 99 % substitute by lower value fish *Megalaspis cordyla*.(FigA2) and Table 7.

Sample of Fish (mm2) which is sold as *Rastreliger kanagurta* but found as *Megalaspis cordyla* and 99 %similiar with it . Accession no of *Megalaspis cordyla* is KM522836 according to Gene Bank database and accession no is ABFJ246-07 according to BOLD database. *Rastreliger kanagurta* is 99 % substituted by the lower value fish *Megalaspis cordyla*.(Table7)

Sample cn3 sold as chewa (*Pseudapocryptes elongates*) but identified as (*Scartelaos gigas*) and similar with 86 % of *Scartelaos gigas* and 86 %of *Taeniodes nigrimarginatus* which accession no is GBGCA11774-15 according to bold reference , on the other hand no sequence in BOLD database of Scaretelaos gigas , the accession no of *Scaretelaos gigas* and *Taeniodes nigrimarginatus* according to Gene Bank is KT277705.1 and KJ865407.1 respectively.Scartelaos gigas sequenc unable to match in BOLd Database and for that resson we have to search NCBI Blast . Phylogenetic tree show that the differences among cn3, *Pseudapocryptes elongates, Scartelaos gigas* and *Taeniodes nigrimarginatus* and conclude that our sample sequence is 86 % similar with *Pseudapocryptes elongates, Scartelaos gigas* and *Taeniodes nigrimarginatus* (plate 1 and Fig.3).

Figure 3: Neighbor joining phylogentic tree of cn3 (sample collected from a supershop, sold as chewa *(Pseudapocryptes elongotus)* with other sequences downloaded from Gene Bank and BOLD systems based on CO1 gene.

Sample ma5 *(Clarius batrachus*) sold as desi magur (fillet) but identified as (*Heteropneustes fossilis*) and which is 99 % similar with Fish *Heteropneustes fossilis* and which accession no accoding to Gene Bank and BOLD database APO12013.1 and ANGBF10687-12 respectively.

Fish sample lg 4 sold as lal chanda *(Pmpus chinencis*) in local market but identified in GeneBank (*Piaractus mesopotamicus*) which is a non-native species in our country it is endemic to South American region and which is 99 % similiar with spesies of *Piaractus mesopotamicus* which accession no is HQ420833.1 according to Gene Bank and BOLD reference is ANGBF6848-12. *Pmpus chinencis* distribution aera mainly Indo-Pacific resigion. They feeds mainly on centophoes ,sals, medusase and other zooplankton. Harmless for human. But identified as *Piaractus mesopotamicus* which is harmful for human consumption and *Pmpus chinencisis* 99 % substute by *Piaractus mesopotamicus.* (Fig. D1 and D2 and Table 7).

Fish sample (bp6) sold as bacha *(Clupisoma prateri*) fish in super shop actually it is bacha fish according to GeneBank accession no is JX983274.1 and also found that no sequence present in BOLD database.

Fish sample cg7 and Fish cj8 which were collected from local market (Gopibag and Jatrabari) but found as (*Tenualosa ilisha*) and their accession no were APO11611.1 according to Gene Bank and accession no CYTC3774-12 according to BOLD Database, actually Jhatka which is forbidden to catch and sell.

Fish sample bj9 sold as Bagha Ayer fillet and actually found to Bagha Ayer according to Gene Bank reference and accession no EU417762.1 and ANGBF5989-12 according to Gene Bank and BOLD Database respectively .

Sample of Fish (ta10) *Tor putitora* sold as Mohashol fillet in Agora Super shop but identified as (*Mylopharyngodon piceus*) which accession no according to Gene Bank HQ236003.1 and GBGC6787-09. Accession No. of *Tor putitora* ANGBF325-12 according to BOLD database.

A total of 10 of the 19 (52.63%) sequenced samples could be readily discriminated at the species level (either matching the species under which they were sold or being assigned to an alternative species), with all showing >5% COI divergence from their nearest neighbouring species. Such results reiterate previous findings relating to the challenges of explicitly identifying closely-related (and potentially introgressed) members of this genus with COI barcoding and genetic-distance analyses (Viñas *et al*. 2009; Wong and Hanner, 2008).

Overall, taking all provinces and outlets into account, a total of 10 samples were sequenced , among them 8 samples were genetically identified as different species to those indicated at the point of sale or inferred from the names under which they were sold . In view of these results, cognisance should also be taken of the increasing complexity and obscurity of seafood supply chains, implying that fraud can manifest at any point from the fishing vessel to the consumer's plate. Since this study was limited to fish sold in restaurants and retail outlets, it cannot be categorically determined where the observed transgressions occurred.

Current cases where the BOLD identification engine reports multiple species with the same sequence similarity are records that are pending review and possible revision.

3.3 DNA sequence alignment

3.3.1. Alignment between sn1 and S*ardinella longiecs*

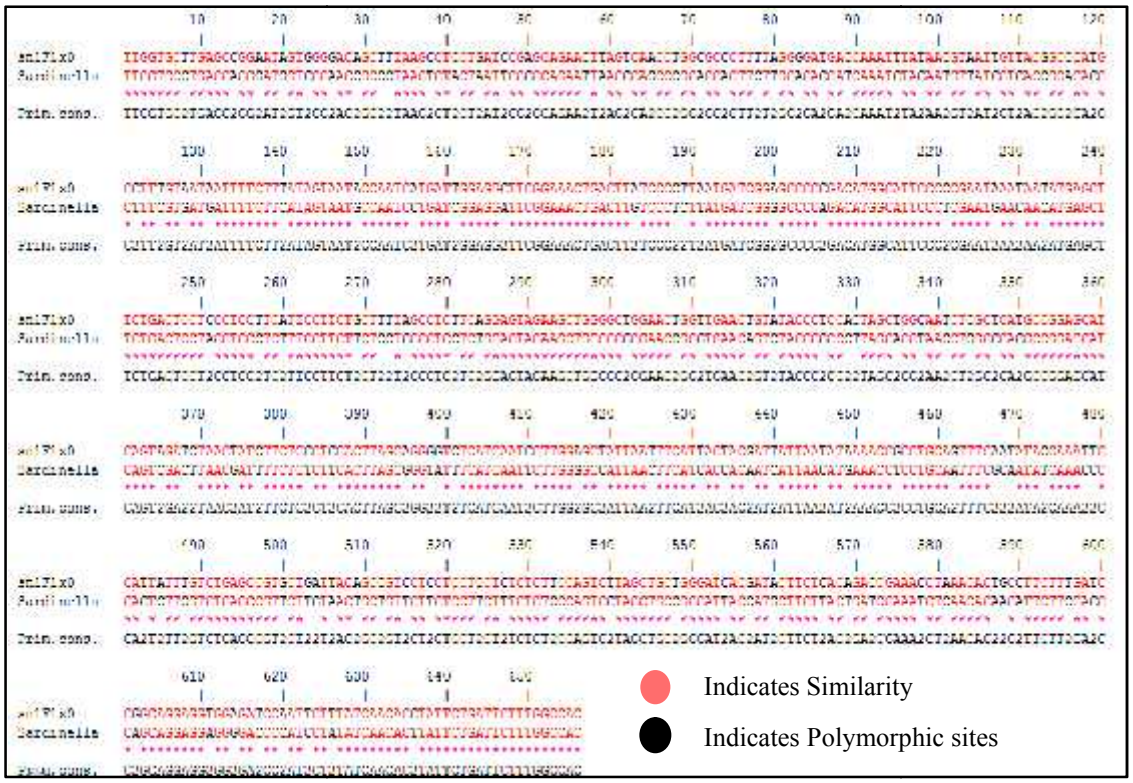


Figure 4: Multiple sequence alignment of CO1 gene fragment of 2 substitute fish spp. Some representative polymorphic sites are indicated by nucleotide position.

Comparison of CO1 gene sequence of *Sadinella longiceps* which is labbled in super shop is compared with the available sequence from Gene Bank database. It, it has been found that the fish is not the Sardin. After comparing these sequences it is different individual, 120 out of 650bp nucleotide bases of the sequence found polymorphic.

Alinment length 656bp and identity is 78.20%similiar with Sardinella and difference is 21.80 %.

Alignment data:

Alignment length : 656 bp

Identity (*) : 513 is 78.20 %

Different: 143 is 21.80 %

Table 8: Polymorphic sites observed between two sequences of sample sn1 and S*ardinella longiecs* sequence (BOLD database

Position (bp)	Sample sn1	S*ardinella longiecs*
20	A	G
50	A	G
80	A	T
110	T	C
140	T	C
200	A	G
230	T	C
260	A	T
320	A	C
330	C	T
350	T	C
410	T	C
500	G	T
540	T	C
560	A	G
590	C	A
620	A	C

So it can be concluded that the bought fish was mislabeled because 17 out of 656 nucleotide bases of the sequence were found polymorphic with the species the sold as in the market. Highly commercial value is substituted by lower commercial value fish. The sample sold in the Super shop is *Megalaspis cordyla.*

3.3.2. Alignment between mm2 and *Rastreliger kanagurta*

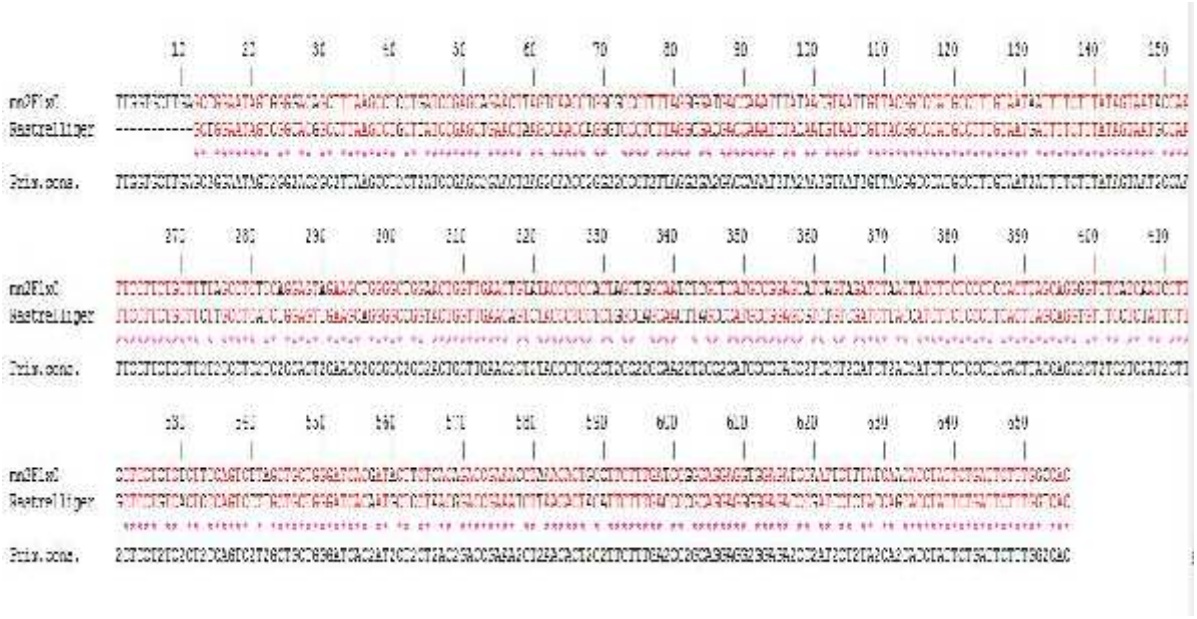


Figure 5: Comparison between collated fish Macerel and Gene Bank reference *Rastrelligir kanagurta* .

When the fish sample collected the sample labeling was *Rastrelliger kanagurta* but identified as *Magalaspis cordyla.* As they were morphologically different in no doubt it is taxonomically different. Polymorphism is visible out of 650bp. Alignment length 656bp and similarity is 79.20% with the sample which they sold.

Alignment length: 656
Identity (*): 519 is 79.12 %
Strongly similar (:) : 0 is 0.00 %
Weakly similar (.) : 0 is 0.00 %
Different : 137 is 20.88 %

Table 9: polymorphic difference between sequence *Rastrelliger kanagurta* and *Megalaspis cordyla.*

Position	sn1	*Rastrelliger kanagurta*
290	A	T
320	A	C
410	C	T
530	T	A
540	T	C
560	A	G
590	C	A
620	A	G

3.3.3 Alignment between cn3 and *Pseudocryptes elongatus*

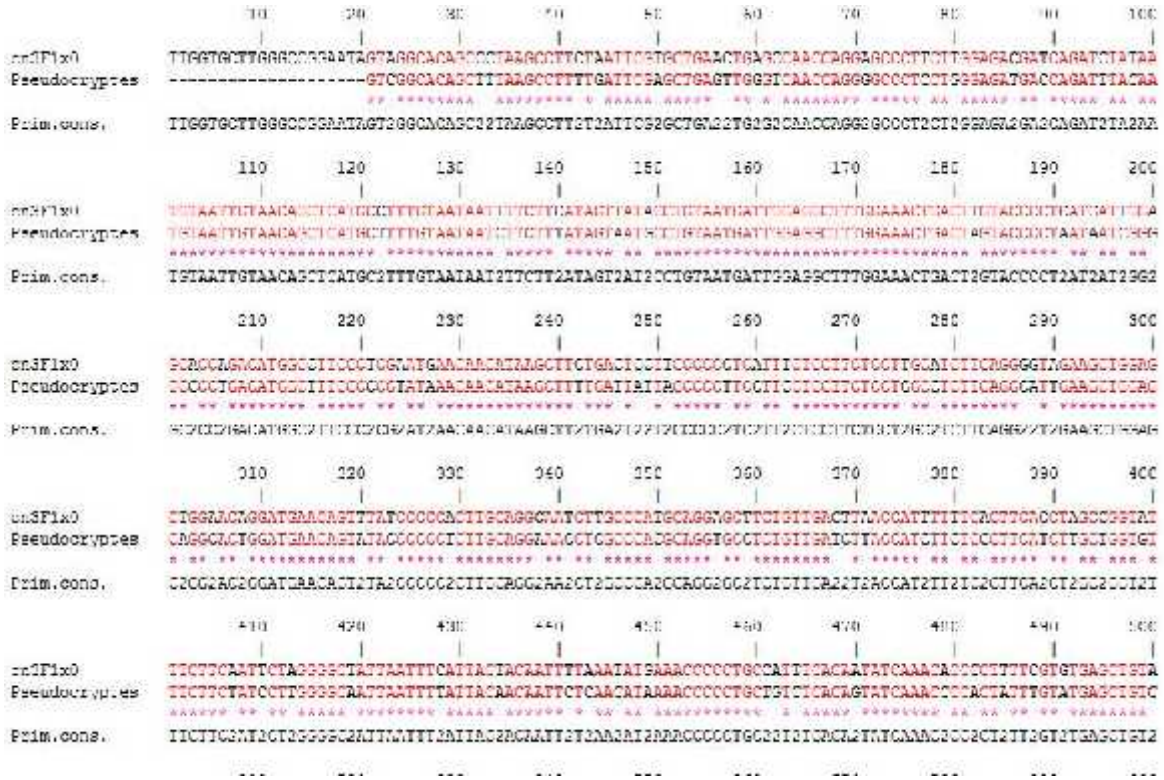


Figure 6: Sequence alignment between two species (cn3 and *Pseudocryptes elongatus*)

The collected sample was *Pseudapocryptes elongates* but find as according to Gene Bank reference it has found *Scartelaos gigas* which accession no is KT277705.1

3.3.4 Alignment between lg4 and *Pampus chinensis*

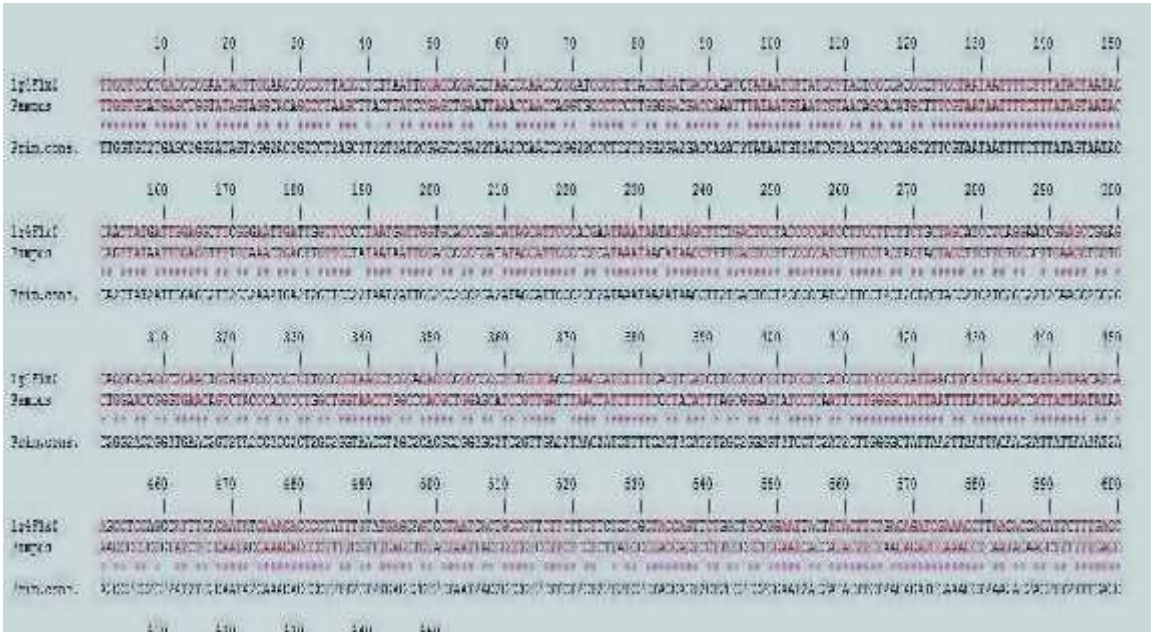


Figure 7. Pairwiese sequence alignment of sample lg4 with the sequence of *Pampus chinsensis*

In the market, the spp lg4 sold as Lanchanda but found as *Piaractus mesopotamicus* which is not rupchanda . To get the high price, it is sold as that Lal rupchanda . The accession number of the sample sequence is HQ420833.1 and similarity is 99% with *Piaractus mesopotamicus*. It is also found polymorphic nucleotide bases and 12 out of 650bp selected as positions to diagnostic value at genus level (Table 12)

Table 10. Polymorphic sites observed between the two sequence lg4 and *Pampus chinensis*

Position	lg4 sample	*Pampus chinensis*
80	A	T
110	T	A
170	C	T
180	T	C
200	T	A
260	C	T
290	C	T
370	A	C
390	C	T
410	C	T
460	C	G

According to the nucleotide difference, it can be indentified as different species found so the species are individual. Everyone should be very careful to buy fish. Consumer should be very careful to buy fish to get rid of fish market fraudulence.

3.3.5 Alignment of sequence between ma5 and *Clarius batrachus*

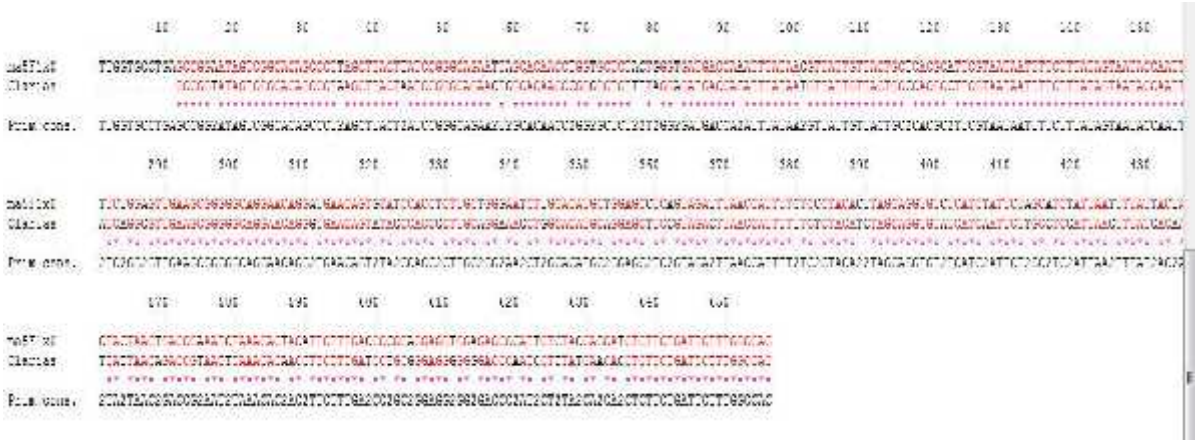


Figure 8 Sequence alignment and species confirmation between ma5 and *Clarius batrachus*

The sample ma5 sold as fillet of magur fish in Agora Super shop but identified as hybrid Shing which is low quality and hybrid Shing is available in the market and the production cost is lower than to collet Desi magur from rural area. Accession no of Magur in BOLD Database ANGBF2196-12 but identified as *Heteropneustes fossilis* which accession no is ANGBF10687-12. Polymorphism has found in different nucleotide bases (Table 11)

Table 11. Polymorphic sites found in different nucleotide between ma5 and *Clarius batrachus.*

Position	ma5 sample	*Clarius batrachus*
80	G	A
320	G	A
380	C	T
390	T	C
590	A	C
620	C	A

3.3.6 Alignment of sequence between cg7 and *Gudusia chapra*

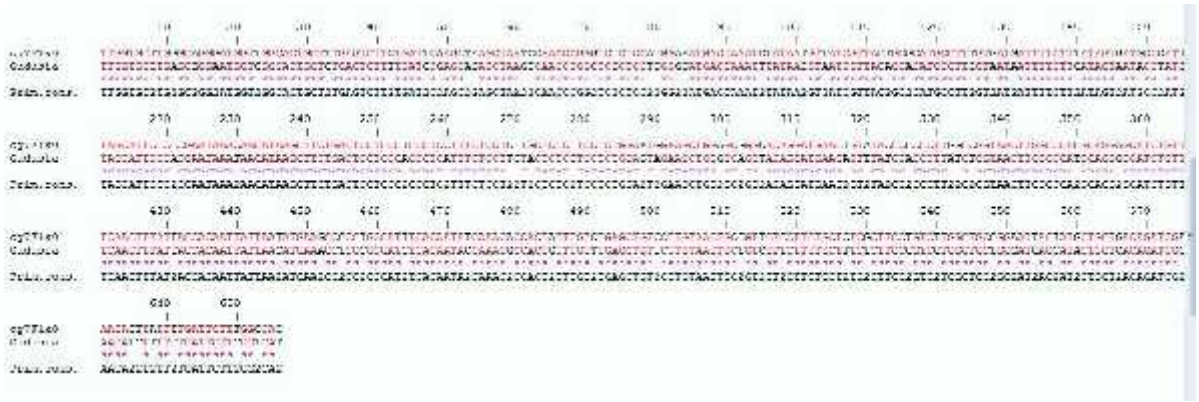


Figure 9 . Sequence alignment of two species (cg7 and *Gudusia chapra)*

Fish sample sold in the market as Chapila (*Gudusia chapra*) but indentified a*s Tenualosa ilisha* which accession no according to Gene Bank and Bold database APO11611.1 and CYTC3774-12 respectively on the other hand the accession no of Gudusia chapra according to Bold database is AP011603.1 (Plate 2 and FigF1 ,FigF2).

They sold fish the name as Chapila because the name of small size Hilsha is known as Jhatka (which lenth is <23cm). To protect the Hilsha Fishary it is forbidden by the Government to catch and sell Jhatka.

3.3..7 Alignment of Sequence of ta10 and *Tor putitora*

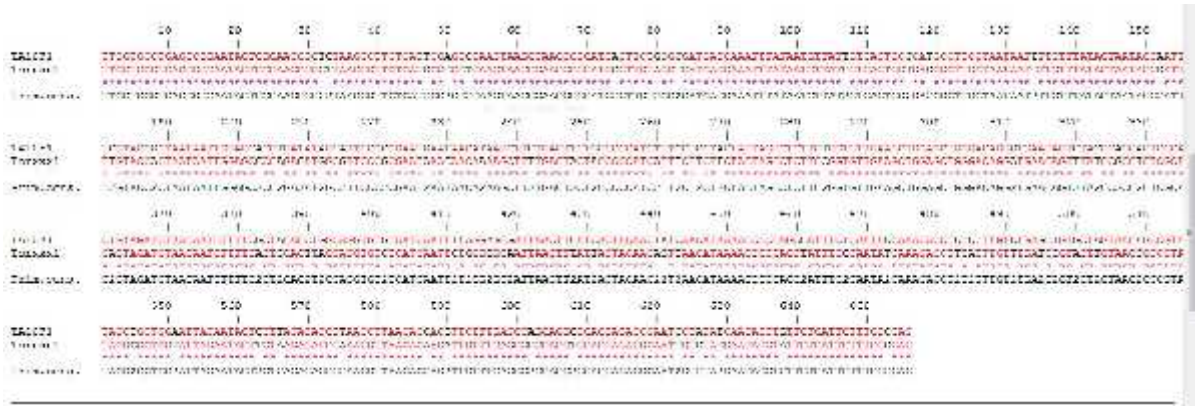


Figure 10. Sequence alignment of two spp and confirmation of actual species(ta10 and *Tor putitora* from Gene bank).

The sample ta10 sold as Mohashol (*Tor putitora*) but indentified as on *Mylopharyngon piceus* which accession no according to Gene Bank HQ236003.1 and GBGC67867-09 according to BOLD database. On the other hand the accession no *Tor putitora* ANGBF325-12 according to Bold database Polymorphic side were found at different point to detect the Genus of the species. Table 12 indicate the polymorphic sites

Table 12. Identification of polymorphic sites between of ta10 (collected) and *Tor putitora*

Position	ta10	original, *Tor putiora*
50	A	G
80	G	A
230	T	C
260	T	A
390	C	T
440	T	A
500	G	A

3.3.8 Multiple Sequence alignment among sn1, *Sardinella longiceps and Megalaspis*

The sequence aligned among three such as sn1, *Sardenella longiceps* and *Megalaspis* spp and it was observed found that there is relationship between sn1 and *Megalaspis* spp

but different from *Sardinella longiceps* sequence. Polymorphic sites has been found at sn1and *Sardinela longiceps* sequence, but no difference sn1 and *Megalaspis cordyla* spp. Significant difference and relationship has been shown (Table13).When the species will be same their neucleotide bases will be same . From plate 1 Fig,A1 and B1 both of the species is substituted by Megalaspis cordyla species which commercial value is lower than *Sardenella longiceps* and *Rastreligrr kanagurta.(*plate 1).

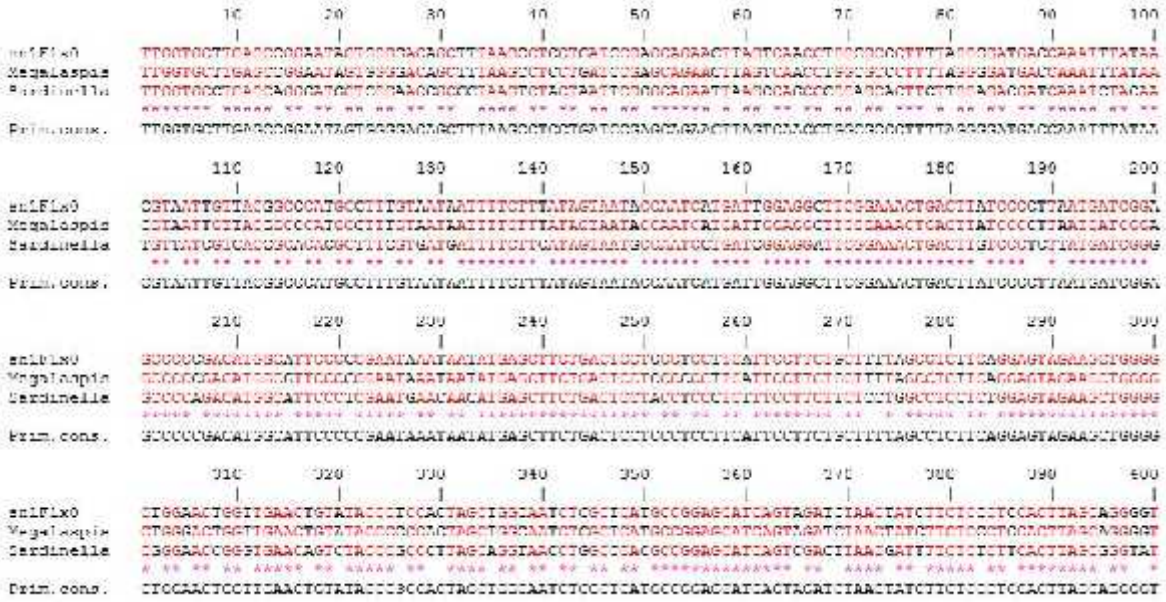


Figure 11. Multiple Sequence alignment among the sample sequence, sequence of sold species and indentified species after submission the sequence at BOLD database

After sequence alignment of different individuals, 10 out of 400 nucleotide bases of the sequence found similar in 10 points between the sample sequence and the Species found after submission the sequence to BOLD database (20,50, 80, 110, 140, 200, 230,260, 320, 350) and found as difference between the sequence (by which name they sold) and the seuquence done after collecting the sample from super shop.The sold sample sequences matches *Megalaspis cordyla* but different from from the *Sardinelle longiceps* , their labeled was incorrect.

Table 13 Visualization at a glance the polymorphic sites amomg sn1, *Megalaspis cordyla* and *Sardinella longiceps*.

Position	Sn1	*Megalaspis cordyla*	*Sardinella longiceps*
20	A	A	G
50	A	A	G
80	A	A	T
110	T	T	C
140	T	T	C
200	A	A	G
230	T	T	C
260	A	A	T
320	A	A	C
350	T	T	C

3.4 Phylogenetic tree analysis

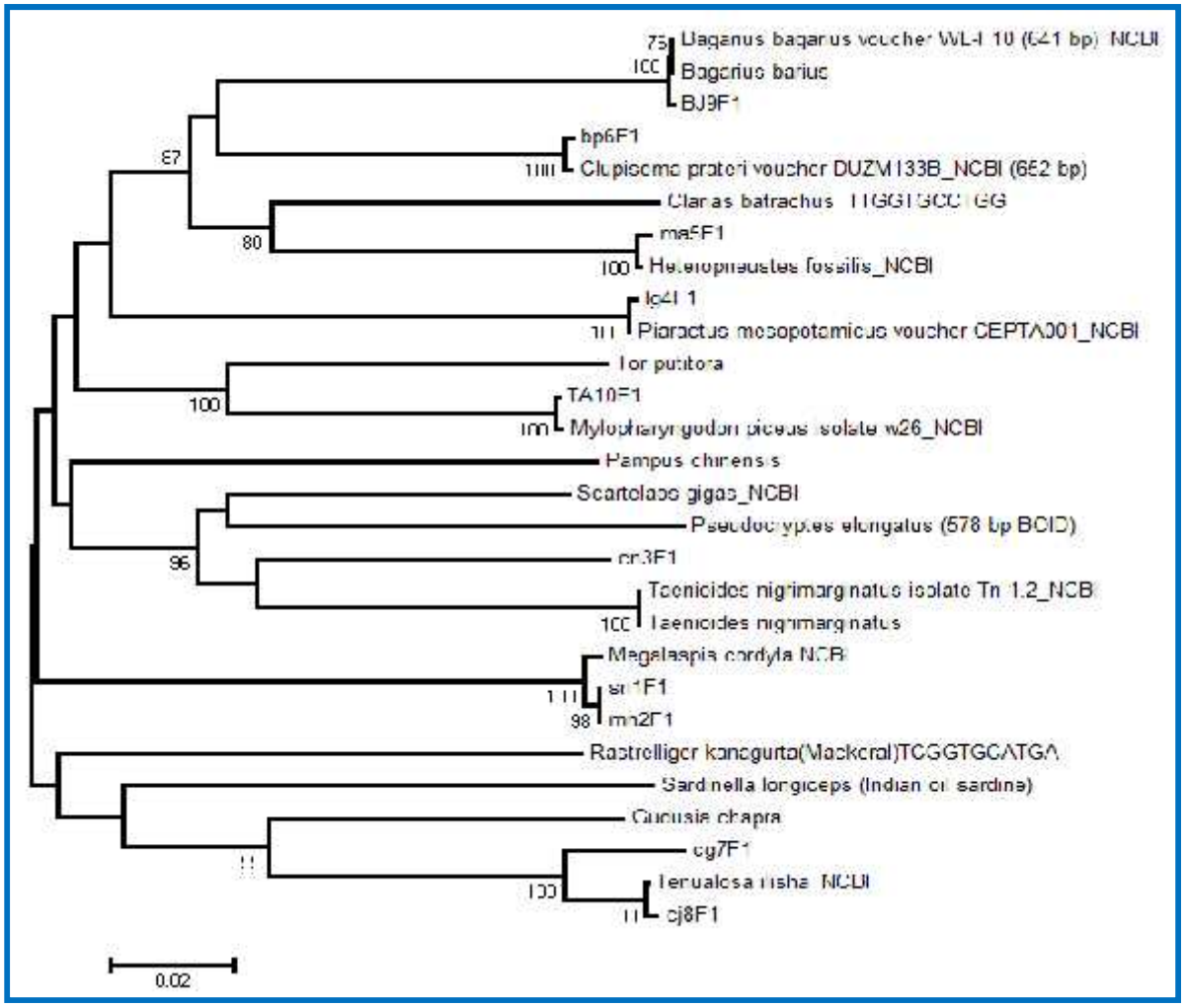


Figure 12. Neighbor joining Plylogenetic tree of all sequences (10 sample sequences, sold name sequences and real species sequence).

All amplified sequences that exceded 650 nucleotides in length with no insertions, deletions, or stop codons were observed, thus reducing the possibility of mtDNA copies in the nucleus. The sequences obtained from the samples were deposited on Gene Baank. Successful matches varied from 89 to 100 % pairwise sequence identity (Table7). Only one samples could not be identified in the BOLD species Reference database. Nonetheless, the BOLD Full database returned hits with a percentage of identity as high as GenBank, with the advantage of being a more reliable source of taxonomic identification. To better evaluate this difference, a neighbor-joining tree using the K2P evolutionary model was built for each group (fillets and whole fishes).

3.5 Proportion (in percentage) of correctly labeled and potentially mislabeled

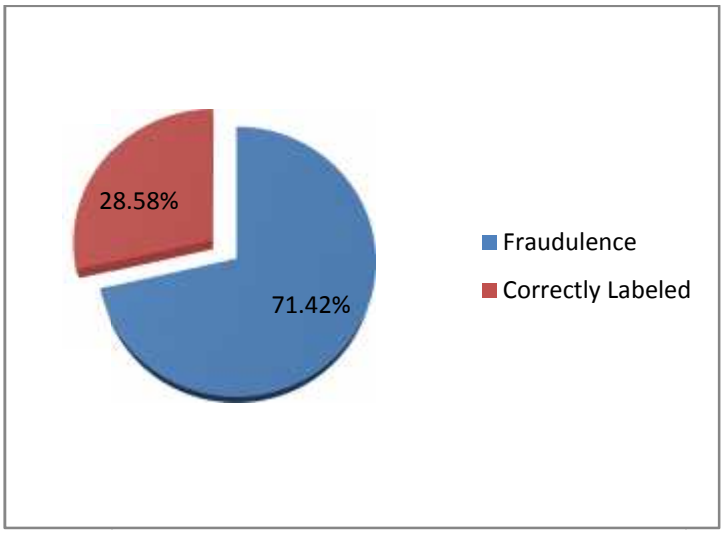


Figure 13. Graphical Representation of the percentage of whole fish fraudulence

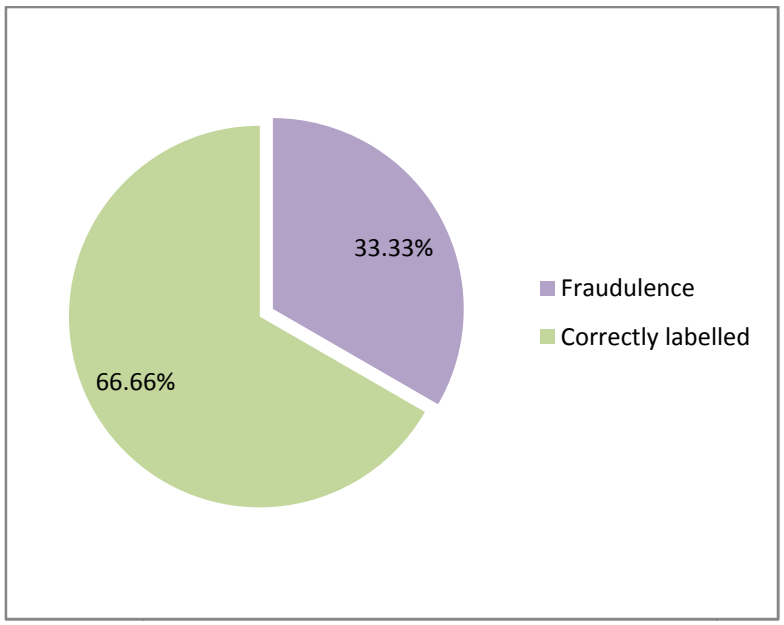


Figure 14. Graphical Representation of the percentage of Fillet fish fraudulence

In relation to international studies, this value corresponds with the seafood mislabelling rates determined for retail outlets in the US (18 %) and Brazil (20 %) (Warner *et al*.

2013), falls below the rate of 32% found for Italian retailers (Filonzi *et al.* 2010), but is considerably higher than those rates reported for retailers in the UK (6%) and Tasmania (0%) (Helyar *et al*. 2014).

In certain circumstances, the determination of whether a species is mislabeled or not depends largely on the geographic area in which it is sold. In this study, a sample sold by Nandan Super shop as Sardin *(Sardinella longiceps*)showed 97 % sequence similarity with *Megalaspis cordyla* . According to the 'seafood list' published in the US (FDA, 2014), 'Sardin' is the legally designated market name only for *Sardinella longiceps*, a highly valued but overexploited species from the Atlantic coast and Gulf of Mexico that has been a target for substitution in North America (Hanner *et al*. 2011).

Nonetheless, it is notable that the common name for Sardin is commonly called as *Mathi* or *Chaala* in Kerala. In north Kerala or Malabar it is called Maththi(മത്തി) and in south Malabar its called Chaala(ചാള). The sardin is called as *Pedvey, Padwa, Washi* in Hindi and Marathi respectively. It is also called as *Tarla* in Hindi and Marathi respectively. It is called as *Tharlae* in Konkani and *Boothai* in Tulu and thus comprehensible that the term 'Sardin' might be preferably selected to appeal to local consumers.

For instance, a sample sold by Meena Bajer Super shop retailer as 'Macerel but identified as *Megalaspis*, while one sold as 'Desi Maguer was rather found to be Shing *(Heteropneustes fossilis).* The fish sample collected from local market Gopibag sold as Chapila (*Gudusia chapra*) but found as small size Hilisha (*Tenualosa ilisha)*.

Among of the 10 retail samples sequenced, 7 sample was whole fish . Among them 5 (71.42%) whole fish sample and the rest of the three were fillet, among them one (33.33%) sample identified as different species , And the overall fraudulence in fish market (both local and super shop) is 60%. On the other hand, mislabeling was not detected in Bagha Ayer (*Bagarius bagarius*) and Bacha fish *(Eutropichthys vacha*) fillet samples since all tilapia fillets were genetically identified as (*Bagarius bagarius*) and Bacha fish respectively . Genetics therefore helped to ascertain the species in 100 % cases of fish fillets.

3.6 Comprehensive labeling requirement

The ambiguities associated with colloquial names in the global marketplace signal a clear need for the utilisation of scientific names in seafood product labeling, as also increasingly called for in the monitoring of international wildlife trade (Gerson *et al.* 2008).

The inclusion of scientific names would not only promote uniformity in seafood trade, but would assist law enforcers to detect fraud or the commercialization of illegal species. For the consumer, however, the need for accurate information on labels exists not only for the species, but should extend to the origin and production method if they are to be fully capable sof making sustainable seafood choices. Regulators in the EU have recognized these factors by effecting legislation requiring that the commercial designation, scientific name, geographical origin and production method (wild or farmed) be declared on fish product labels (EC 2001).

3.7 DNA Barcode analyses and optimizes threshold

Samples were all identifiable by COI sequencing because our results indicated that all the species examined showed a unique sequence clearly distinguishable from the others. After comparison with reference sequences from databases, a high level of mislabeling was detected in the frozen fish fillets analyzed.

PLATE

Species on the Label	Identification Similarity with another species	Real Species
A1 	99%	 A2
Sardenella Longiceps		*Megalaspis Cordyla*
B1 	99%	B2 
Rastreliger Kanagurta		*Megalaspis Cordyla*
C1 	86%	 C2
		Scartelaos gigas
 D1	99%	 D2
Pampus chinensis		*Piaractus Mesopotamicus*

E1 	99%	E2 
Clarius batrachus		*Heteropneustes fossilis*
F1 	99%	F2 
Gudusia Chapra		*Tenualosa ilisha*
G1 	100%	G2 
Bagarius bagarius		*Bagarius bagarius*
H1 	99%	H2 
Clupisoma prateri		*Clupisoma prateri*
I1 	99%	I2 
Tor Putitora		*Mylopharyngodon piceus*

Chapter 4
Discussion

The ambiguities associated with colloquial names in the global marketplace signal a clear need for the utilisation of scientific names in seafood product labelling, as also increasingly called for in the monitoring of international wildlife trade (Gerson *et al*. 2008). DNA barcoding was 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 sequences in forensics has already been investigated for reproducibility, heteroplasmy, mixed DNA samples, chemical treatments, environmental conditions and other factors showing consistent results in which a great range of reference data exist (Dawnay *et al*. 2007).

In this study, we identified commercial samples labeled as sardine through the comparisons of COI mtDNA sequences using the BLAST engine to search Gene Bank. The high rate of substitution of this sea fish species could be due to the fact that the vernacular name of Mackerel and Sardin is well known in the Super market. Therefore, we have strong evidence that intentional mislabeling of cheaper fish is more frequent phenomenon mainly within processed fish. Therefore, by using this mislabel, traders might be able to sell their product in a better price. This is clear as we compare the market price of the fishes sold as Mackerel and Sardin identified in this study. For instance, species labeled bagres, a Bangladesh vernacular name for a less known Surma fish group, are sold at a 63% lower price than fishes under the Mackerel and Surma label.

Specimen (sn1), sold as ‘Sardin’ (*Sardinella longiceps*) but identified in Gene Bank as (*Megalaspis cordyla)* which accession no KM522836.1 according to Gene Bank reference and accession no ABFJ246-07 according to BOLD data base reference and the sold spp is 99% similar with *Megalaspis cordyla* . Samples were all identifiable by COI sequencing because our results indicated that all the species examined showed a unique

sequence clearly distinguishable from the others. After comparison with reference sequences from databases, a high level of mislabeling was detected in the frozen fish fillets analyzed (Table7 and Plate1). The identified species did not match with the one declaredon the labels in 99 % Sardin and Mackerel 99 % panga samples analyzed, given a total 71.42% mislabeling in the analyzed in the case of whole fish and 33.33% mislabeling in the case of fillet fish.(Table 7). Sample of Fish (mm2) which is sold as *Rastreliger kanagurta* but found as *Megalaspis cordyla* and 99%Similiar with it . Accession no of *Megalaspis cordyla* is KM522836 according to Gene Bank database and accession no is ABFJ246-07 according to BOLD database.

Sample cn3 sold as chewa (*Pseudapocryptes elongates*) but identified as (*Scartelaos gigas*) and similar with 86% of *Scartelaos* spp and 86%of *Taeniodes nigrimarginatus* which accession no is GBGCA11774-15 according to BOLD reference, on the other hand no sequence in BOLD database of *Scaretelaos gigas* , the accession no of *Scaretelaos gigas* and *Taeniodes nigrimarginatus* according to Gene Bank is KT277705.1 and KJ865407.1 respectively. (Fig. C1 and C2) and Table 7.

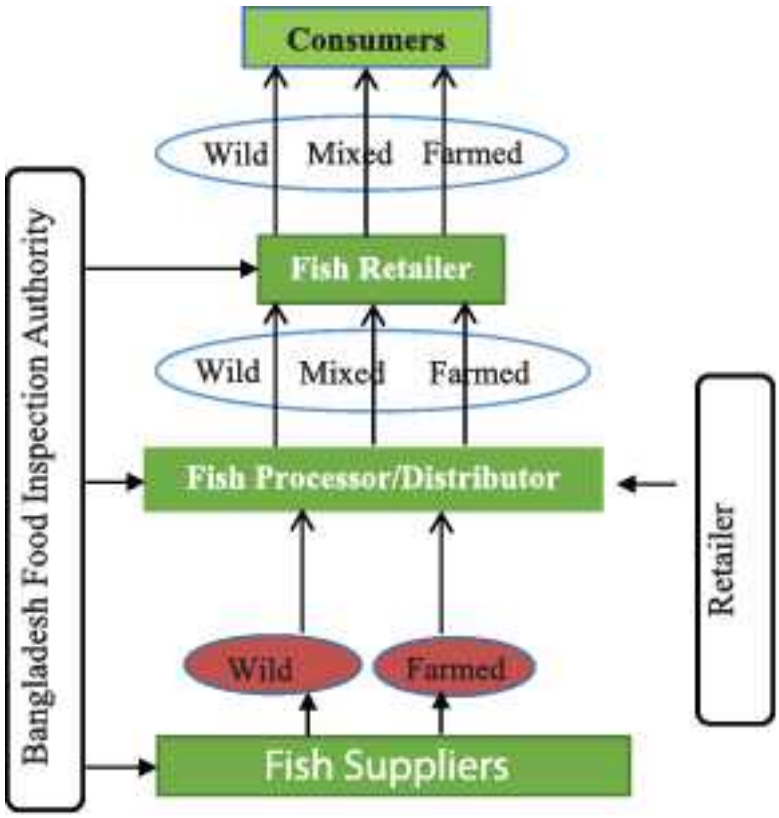
Fish sample lg 4 sold as lal chanda *(Pampus chinencis*) in local market but identified in GeneBank (*Piaractus mesopotamicus*) which is a non-native species in our country it is endemic to South American region and which is 99%similiar with species of *Piaractus mesopotamicus* which accession no is HQ420833.1 according to Gene Bank and BOLD reference is ANGBF6848-12. Sample ma5 *(Clarius batrachus*) sold as desi magur (fillet) but identified as (*Heteropneustes fossilis*) and which is 99% similar with Fish *Heteropneustes fossilis.* Fish sample (bp6) sold as bacha *(Clupisoma prateri*) fish in super shop actually it is bacha fish according to GeneBank accession no is JX983274.1 and also informatie that no sequence present in BOLD database. Fish sample cg7 and Fish c8j which collected from locl market (Gopibag and Jatrabari) but found as (*Tenualosa ilisha*) and their accession no APO11611.1 according to Gene Bank and accession no CYTC3774-12 according to BOLD Database, actually Jhatk which is forbidden to catch and sell .

Fish sample bj9 sold as Bagha Ayer fillet and actually found to Bagha Ayer according to Gene Bank reference and accession no EU417762.1 and ANGBF5989-12 according to Gene Bank and BOLD Database respectively .

Sample of Fish (ta10) Tor *putitora* sold as Mohashol fillet in Agora Super shop but identified as *Mylopharyngodon piceus* spp which accession no according to Gene Bank HQ236003.1 and GBGC6787-09. Accession no of *Tor putitora* ANGBF325-12 according to BOLD database.

Therefore, We have strong evidence that intentional mislabeling of cheaper fish products is a more frequent phenomenon mainly within processed and packaged fish. The establishment of conservation strategies and the normalization of vernacular names for native commercially important Brazilian fishes, together with the molecular inspection of fish products, have the potential to form an important tool for the preservation of our fish fauna and protect consumers from mislabeled products (Ugochukwu *et al*.2015).

Flow Chart 4: Stylized supply chain featuring technology adoption for supply chain monitoring



We strongly recommend the establishment of a valid list of commercial and Latin names for the fishes commercialized in Bangladesh. Such a reference list would make possible for the authority to be able to regulate and detect fraud, substitution and the commercialization of threatened species.

In addition, custom services will have the ability to regulate and inspect imported/exported items, for the purpose of taxation and to protect the consumer from misguidance. Such a list is currently in use together with Barcoding analysis to detect market substitution in North American seafood (Wong *et al*. 2008).

This work represents study using DNA barcoding for fish species identification in Fish markets, focused on the main species that are sold in Super Shop as fillets (Bagha Ayer, Desi fillet Magur, Mohashol, and Bacha fish). Our results revealed a high incidence of mislabeling. In essence, 33.33% of the highly valued Mohashol *(Tor putitora*), was substituted by the fish *Mylopharyngodon piceus* Similarly, 98% Desi Magur (*Clarius batracus)* was also replaced by hybrid low quality (fade colour) shing fish. The reason for this mislabeling could be economic in both cases. Desi magur more expensive than hybrid shing fish needs long growing time and higher level of feed consumption (Thanh 2003). *Sardenella longiceps* and *Rastreliger kanagurta* have 99 % similarity with *Megalaspis cordyla*. For *Pseudocryptes elongate*s 86 % similarity was observed with *Scatelaos gigas* and its sequence is not available in BOLD database. *Pampus chiensis* has 99% similarity with *Piaractus mesopotamicus* and *Clarius batrachus* 99 %similarity with hybrid low quality *Heteropneustes fossilis*. *Gudusia chapra* is 99 %similarity with small size *Tenualosa ilisha*. But no substitution found for the species of *Clupisoma prateri* and *Bagarius bagarius.* The overall substitution in fish market is 80 %because 8 out of 10 samples found mislabeled or misnaming. For specific fraudulence in the case of whole fish was 71.42 % and 33.33 % fraudulence found for fillet samples.

Moreover, such substitution not only impinges on economy but may have also implications for health. Some fish species is replaced by species from distinct genera, morphologically unalike. Interestingly, some of the species found were identified as saltwater species of lower commercial value Food safety and public health may be at risk if this practice of seafood substitution becomes generalized. When considering the International Union for Conservation of Nature (IUCN) Red List Status, two endangered species were being sold in marketplaces such as *Bagarius bagarius,*and another *is Tor putitora* . Both species are considered endangered by the IUCN Red List). The marketing of fish labeled with the Fish markets as a high fidelity tool for species identification.

The governmental program of seafood inspection resulted in law enforcement and financial penalties to the restaurant, supermarket and fishmonger owners for all the

mislabeling cases reported here. Lastly, we also suggest that our Government agencies and Non-Government organization should monitor regularly fish products to impose more strict controls for product labeling.

Several research reports have shown a wide range of replacements of fish worldwide, an economic and potential health problem for consumers. Thus, considering the high mislabeling rate reported here, we suggest that the implementation of a systematic regulatory program conducted by governmental agencies, which are able to apply penalties to those responsible for illegal mislabeling, is necessary to discourage market substitutions by consistently penalizing cases of replacements.

Chapter 5
Conclusion and Recommendations

5.1 Conclusion

The DNA barcoding technology and reference library is an improvement over the existing DNA-based techniques in detecting food fraud, particularly in fish markets. This work has provided an initial overview of the economic incentives for species substitution and mislabeling in fish markets, modeled the effect of private (third party) adoption of DNA barcoding technology for supply chain monitoring, and examined the feasibility of the technology for a typical retail store.

DNA barcoding offers a new level of precision in the application of species names, which is increasingly important in the expanding international market. The ease of generating DNA barcodes and a focus on high quality data records instill increasing confidence in the technique.

Fraudulence is identified worldwide as a significant problem associated with issues of food standards, traceability and security. The results presented here show that seafood fish as sold by retail outlets is not mislabeled and highlight the utility of DNA barcoding for testing regulation. However, comprehensive and absolute naming standards, wherever fish products may be consumed, are required if ambiguous names are to be eliminated from the industry. Further, such a standard might be complemented by the implementation of an appropriate monitoring regime to ensure standard names are properly applied.

5.2 Recommendations

- Study was conducted only in very short period and for that reason few number of replica were used so further study is needed.
- Fraudulence is increasing day by day specially in fish market by substituting higher commercial value added fish by lower commercial value added fish and consumer deprives from real fish species.
- As DNA barcoding technology is very important tool in detecting fraud so it is very essential to apply it to get rid from this type of problem.

LITERATURE CITED

ANDERSON, J. L. 2003. The international seafood trade. Woodhead Publishing, Cambridge.

CARVALHO, D. C., PALHARES, R. M., DRUMMOND, M. G., and FRIGO, T. B. 2015. DNA barcoding identification of commercialized seafood in South Brazil: A governmental regulatory forensic program. Food Control 50: 784–788.

CARVALHOP, D. C., NATO, D.A., BRASIL, B. and OLIVEIRA, D. A. 2011. DNA barcoding unveils a high rate of mislabeling in a commercial fresh water catfish from Brazil. Mitochondrial DNA 22: 97-105.

CAWTHORN, D., STEINMAN, H. A., and HOFFMAN, L. C. 2013. A high incidence of species substitution and mislabeling detected in meat products sold in South Africa. Food Control 32(2): 440–449.

COSTA, F. O. and CARVALHO, G. R. 2007. The barcode of life initiative: synopsis and prospective societal impacts of DNA barcoding of fish. Genomics, Society and Policy 3: 29–40.

DASMAHAPATRA, K. K., and MALLET, J. 2006. DNA barcodes: Recent successes and future prospects. Heredity 97: 254–255.

DAWNAY, N., OGDEN, R., MCEWIN, R., CARVALHO. G. R. and THORPE, R. S. 2007. Validation of the barcoding gene COI for use in forensic genetic species identification. Forensic Science international 173(1): 1-6.

EC (EUROPEAN COMMISSION) 2001. Commission Regulation (EC) No 2065/2001 of 22 October 2001 laying down detailed rules for the application of Council Regulation (EC) No 104/2000 as regards informing consumers about fishery and aquaculture products. Official Journal of the European Communities L278: 6–8.

FILONZI, L., CHIESA, S., VAGHI, M. and NONNIS MARZANO, F. 2010. Molecular barcoding reveals mislabeling of commercial fish products in Italy. Food Research International 43: 1383–1388.

FLOYD, R., LIMA, J., DEWAARD, J., HUMBLE, L. and HANNER, R. 2010.Common goals: policy implications of DNA barcoding as a protocol for identification of arthropod pests. Biol. Invasions **12**(9): 2947–2954.

GERSON, H., CUDMORE, B., MANDRAK, N. E., COOTE, L. D., FARR, K. and BAILLERGEON, G. 2008. Monitoring international wildlife trade with coded species data. Conservation Biology 22: 4–7.

HANNER, R., BECKER, S., IVANOVA, NV. and STEINKE, D. 2011. FISH-BOL and sea food identification: geographically dispersed case studies revel systemic market substitution across Canada. Mitochondrial DNA **22**: 106-122.

HAUCK, M. and KROESE, M. 2006. Fisheries compliance in South Africa: A decade of challenges and reform, 1994–2004. Marine Policy 30: 74–83.

HEBERT, P. D. N., PENTON, E. H., BURNS, H. M., JANZEN. D. H. and HALLWACHS, W. 2004. A. DNA barcoding reveals cryptic species in the neotropical skipper butterfly. Astraptes fulgarator. Proc. Natl Acad. Sci. USA 101: 14 812–14 817.

HOBBS, J. E. 2004. Information asymmetry and the role of traceability systems. Agribusiness. An International Journal 20(4): 397–415.

HEBERT, P. D, N., CYWINSKA, A., BALL, S, L. and WAARD, J. R. 2003. Biological identifications through DNA barcodes. Proc. R. Soc. B 270: 313–322.

HELYAR, S. J., LLOYD, H. A. D., DE BRUYN, M., LEAKE, J., BENNETT, N. and CARVALHO, G. R. 2014. Fish product mislabeling: Failings of traceability in the production chain and implications for Illegal, Unreported and Unregulated (IUU) fishing PLoS ONE 9(6): e98691.

HOFMANN, K. (1987). Fundamental problems in identifying the animal species of muscle meat using electrophoretic methods. FleischwirtschafT 67: 820–826.

HOGG, I. D. and HEBERT, P. D. N. 2004. Biological identification of springtails (Collembola: Hexapoda) from the Canadian Arctic, using mitochondrial DNA barcodes. Canadian Journal of Zoology 82: 749–754.

HUBERT, N., HANNER, R., HOLM, E., MANDRAK, N. E., TAYLOR, E. and BURRIDGE, M. 2008. Identifying Canadian freshwater fishes through DNA barcodes. PLoS One 3(6): 2490.

HUXLEY, I.E., SHAWJ, J.LA., FLETCHER,C., PARNELL,J. and WATTS,PC. 2012. Use of DNA Barcoding to reveal specias compassion of convenience seafood. Conservation biology 26 (2): 367-371.

JHA, B. R. and RAYAMAJHI, A. 2010. "Tor putitora". IUCN Red List of Threatened Species. Version 2011. International Union for Conservation of Nature.

KHALLAF, A. G., ARDURA, A.,GEBA,K.M., BORREL ,Y.J.AND VAZQUEZ ,E.G. 2014. DNA barcoding reveals a high level of mislabeling in Egyptian fish fillets. Food control 46: 441-445.

KOH, T.L., KHOO, G., FAN, L.Q. and PHANG, V.P.E. 1999. Genetic diversity among wild forms and cultivated varieties of discus (*Symphysodon* spp) as revealed by random amplified polymorphic DNA fingerprinting. Aquaculture 173: 485-497.

KUMAR, S., TAMURA, K., and NEI, M. 1994. MEGA: molecular evolutionary genetics analysis software for microcomputers. Comput Appl Biosci. 10: 189–191

LAMENDIN, R., KAREN, M .and ROBERT, D. WARD. 2014. Labelling accuracy in Tasmanian seafood: An investigation using DNA barcoding. Food control 47: 436-443.

LESCHIN-HOAR, C. 2011. Specious species: Fight against seafood fraud enlists DNA testing. Affordable genetic sequencing and new data base of “DNA barcodes” help combat the practice of selling one fish as another. The Scientific American, November 2011.

LOCKLEY, A., and BARDSEY, R. 2000. DNA-based methods for food authentication. Trends in Food Science and Technology 11: 67–77.

MARKO, P.B., LEE, S. C., RICE, A.M., GRAMLING, J. M., FITJHHENRY,T.M. and MCALISTER, J. S.2004. Mislabeling of a depleted reef fish. Nature 430: 309–310.

NELSON, L. A., WALLMAN, J. F and DOWTON, M. 2007. Using COI barcodes to identify forensically and medically important blowflies. Med Vet Entomol 21: 44–52.

NAMAREE, D.C., JOHN, DUNCAN., CHRIS, KASTERN., JUNAID., FRANCIS and LOUWREMS,C.2015. Fish species substitution and misnaming in South Africa.

PEPE, T., TROTTA, M., Di, MARCO, I., ANASTASIO, A., BAUTISTA, J.M. 2007. Fish species identification Surimi based products. J. Agric. Food Chem. 55(9): 3681–3685.

RATNASINGHAM, S., PAUL D. N. and HEBERT 2007. "BOLD: The Barcode of Life Data System (http://www.barcodinglife.org)". Molecular Ecology Notes 7 (3): 355–364.

REHBEIN, H. 1990. Electrophoretic techniques for species identification of fishery products. Zeitschrift für Lebensmittel-Untersuchung und -Forschung 191: 1–10.

ROCK, J., COSTA, F. O., WALKER, D. I., NORTH, A. W., HUTCHNSON, W. F. and CARVALHO,G.R.2008. DNA barcodes of fish of the Antarctic Scotia Sea indicate priority groups for taxonomic and systematic focus. Antarctic science 20: 253-62.

STOECKL, M. November–December 2013. "DNA Barcoding Ready for Breakout". GeneWatch 26 (5).

SUJEEVAN, RATNASINGHAM and PAUL, D. N. 2007. HEBERT.Validation of the barcoding gene COI for use in forensic genetic species identification. Forensic Sci International 173: 1–6.

THANH, N. X. 2003. Case study: Catfish fight, Vietnam's tra and basa fish exports to the US.

TVETERÅS, S., ASCHE, F., BELLEMARE, M.F., SMITH, M.D., GUTTORMSEN, A.G., and LEM, A. 2012. Fish is food — the FAO's Fish Price Index. PLoS ONE **7**(5): 36731.

UGOCHUKWU, A.I., HOBBS, J.E., PHILLIPS, P.W.B. and GRAY, R. 2015. An economic analyses privet incentives to adopt DNA barcoding technology for fish species authentication in Canada 58: 1-9.

VINAS, J. and TUDELA, S. 2009. A validated methodology for jenetic identification of tuna species. PLOS one 4, e7606.

WARD, R. D. 2009. DNA barcode divergence among species and genera of birds and fishes. Mol Ecol Resour **9:** 1077–1085.

WARD, R.D., ZEMAKAK, T. S., INNES, B. H., LAST, P. R and HEBERT, P. D. N. 2005. DNA barcoding Australia's fish species .Phil. Trans. R. Soc. B 360: 1847–1857.

WONG, E.S.K. and HANNER, RH. 2008. DNA barcoding detects market substation in north American sea food. Food Research International 41: 828-837

WORM, B., BARBIER, E. B., BEAUMONT, N., DUFFYJ. E., FOLKE, C. and HALPERN, B.S. 2006. Impact of biodiversity loss on ocan ecosystem services. Science 314:787-790.

YANCY, H. F., ZEMLAK, T. S., MASON, J. A., WASHINGTON, J. D., TENGE, B.J. and NGUYEN, N.-L. T. 2007. Potential use of DNA barcodes in regulatory science: Applications of the regulatory fish encyclopedia. Journal of Food protection 71 (1): 210–217.

APPENDIX-1

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

12.14 g of Trizma base (MW=121.14) was dissolved in 75 ml of distilled water. The pH of this solution was adjusted to 8.0 by adding about 5 ml of concentrated HCl. The volume of the solution was adjusted to a total of 100 ml with de-ionized distilled water. Then it was sterilized by autoclaving and stored at 4ºC.

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

18.61 g of EDTA (EDTA.$2H_2O$, MW = 372.24) was added to 75 ml of distilled water and stirred vigorously with a magnetic stirrer. Approximately 2 g of NaOH pellets was added to adjust the final pH to 8.0. The final volume of the solution was adjusted to 100 ml by adding sterile de-ionized distilled water. The solution was sterilized by autoclaving and stored at 4ºC.

5 M Stock Solution of NaCl (100 ml)

29.22 g of sodium chloride (NaCl, MW = 58.44) was dissolved in 75 ml of distilled water. The total volume of the solution was adjusted to 100 ml with distilled water. The solution was heated in an oven for 15 seconds and stirred vigorously on a magnetic stirrer to dissolve NaCl. It was then sterilized by autoclaving and stored at 4ºC. NaCl was added in small amount at once to be dissolved in solution.

70% Ethanol (100 ml)

30 ml double distilled water was added in 70 ml absolute ethanol to prepare 100 ml 70% ethanol.

Stock Solution of TE (Tris-HCl EDTA) Buffer pH 8.0 (100 ml)

1 ml of 1 M Tris-HCl was added with 0.2 ml (200 µl) of 0.5 M EDTA. The final volume was adjusted to 100 ml with sterile de-ionized distilled water. The solution was sterilized by autoclaving and stored at 4^0C.

3 M Sodium acetate pH 5.2 (100 ml)

40.824 g of sodium acetate with 70 ml of ddH_2O and adjusted the final volume to 100 ml with ddH_2O and p^H was adjusted to 5.2. Then it was sterilized by autoclaving.

50 × TAE Buffer (pH 8.3) (1 liter)

242 g Trizma base (MW=121.14) was dissolved into 900 ml of sterile de-ionized distilled water. Then 57 ml glacial acetic acid was added to the solution. Finally, 100 ml 0.5 EDTA (pH 8.0) was added in it. They were mixed well. The pH of the solution was adjusted by mixing concentrated HCl at pH 8.3. The final volume of the solution was adjusted to 1000 ml.

6× Loading Dye

This is required to load samples in gel electrophoresis for further visualization.

Prepararation of stock solutions

i. 10 ml of a 2% bromophenol blue stock solution.
ii. 10 ml of a 2% xylene cyanol stock solution.
iii. 50% glycerol solution.

The stock solutions were diluted to prepare 10 ml of the final 6X Loading dye with the following component concentrations:

i. 30% glycerol
ii. 0.3% bromophenol blue
iii. 0.3% xylene cyanol

The 6×loading dye solution can be stored indefinitely in the refrigerator. The bromophenol blue, xylene cyanol and glycerol stock solutions can be stored indefinitely at room temperature. The 6× DNA Loading Dye is added to DNA samples to achieve a final dye concentration of 1×.

Ethidium Bromide Solution

For 1ml solution, Ethidium Bromide 10 mg was added to 1 ml of sterile de-ionized distilled water. It was then mixed by hand shaking. The solution was then transferred to a dark bottle and stored at room temperature. Stock solution of 10 mg/ml can also be purchased directly from companies.

APPENDIX-2

Volume of solutions to prepare extraction buffer

Chemical Names	**Molecular Weight**	**Stock Conc.**	**Reference Con./working conc.**	**Working Volume (100 ml)**
NaCl	58.44	5 M	100mM	2 ml
EDTA (pH 8)	372.24	0.5 M	100 mM	20 ml
Tris-Base (pH 8)	121.1	01 M	50 mM	5 ml
Sarcosyl		10%	10%	20ml

Component of PCR Reaction Mixture (for 17 reactions)

Sl. No.	Reagents	Amount per sample	Total
1	Sterile de-ionized distilled water	18.7 µl	317.9 µl
2	*Taq* Buffer A 10X (Tris with 15 mM $MgCl_2$)	2.5 µl	42.5 µl
3	Primer	1.0 µl	17 µl
4	dNTPs 10 mM	0.25 µl	4.25µl
5	*Taq* DNA Polymerase 5U/µl	0.05 µl	.85 µl
6	Template DNA	2.5 µl	42.5µl
Total		**25.0 µl**	**425 µl**

Random primers used in the present study for screening

Primer code	Sequence (5′—3′)	G+C content (%)
Fish F1	TCAACCAACCACAAAGACATTGGCAC	46.154
Fish F2	TCGACTAATCATAAAGATATCGGCAC	38.462
Fish R1	TAGACTTCTGGGTGGCCAAAGAATCA	46.154
FishR2	ACTTCAGGGTGACCGAAGAATCAGAA	46.154

Stock solution for DNA concentration measurement

	Standard Assay Tubes	user sample Assay Tubes
Volume of working solution (From Step 1) to add	190	180-190
Volume of standard (from kit) to add	10	
Volume of user Sample to add		1-20
Total Volume in each Assay tube	200	200

APPENDIX-3

Identification of Collected Samples using the Gene Bank and Bold Search Engines

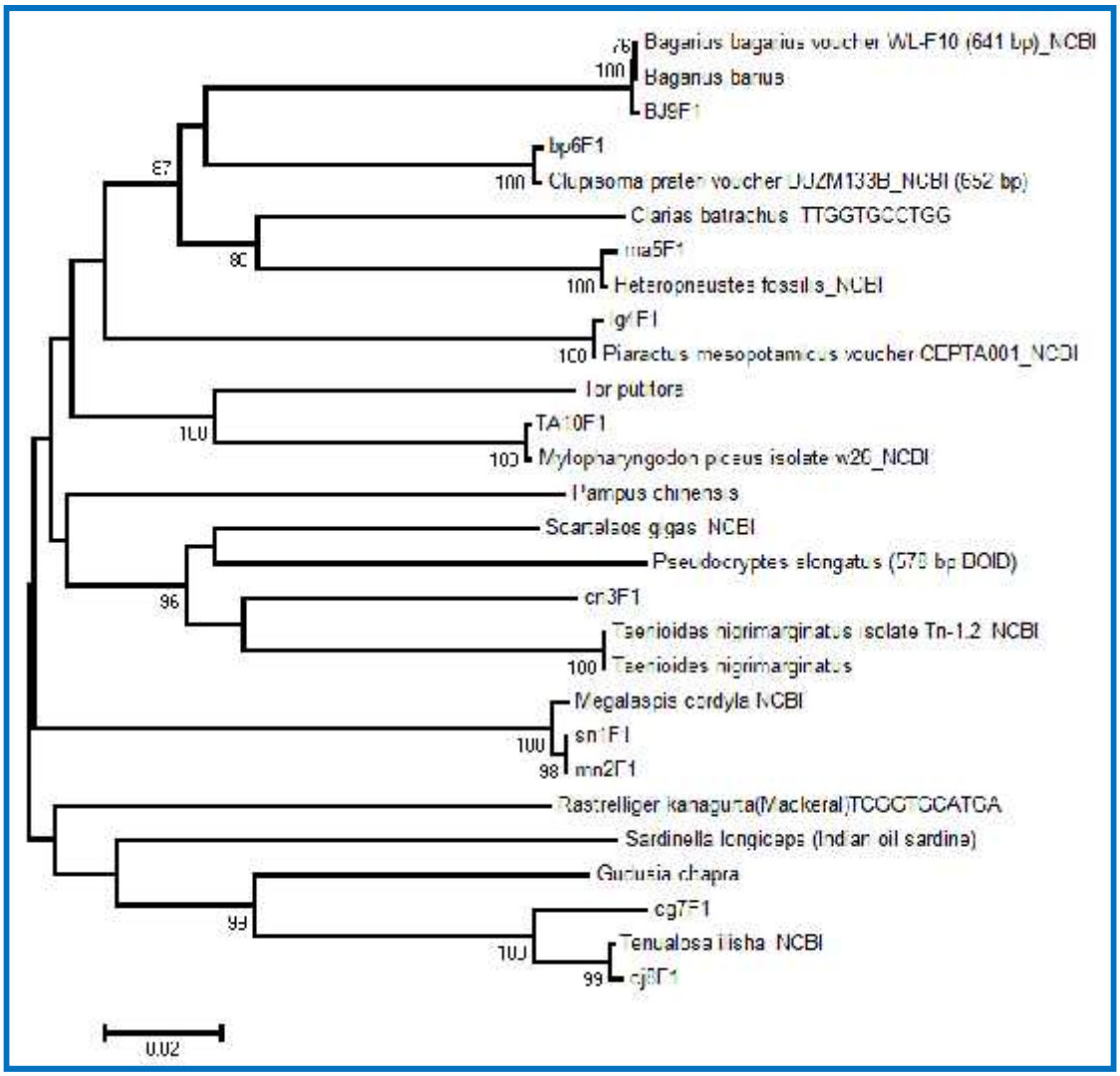
serial	sample ID	Sold as	Description	Max score	Total score	Query cover	E value	Identification	Accession	Bold reference database	Mislabeled
1	sn1	*Sardenella longiceps*	*Megalaspis cordyla*	1184	1184	100%	0	99%	KM522836.1	ABFJ246-07	**Yes**
2	mm2	*Rastreliger kanagurta*	*Megalaspis cordyla*	1184	1184	100%	0	99%	KM522836.1	ABFJ246-07	**Yes**
3	cn3	*Pseudapocryptes elongatus*	*Scartelaos gigas*	686	686	100%	0	86%	KT277705.1	No sequence	**Yes**
			Taenioides nigrimarginatus	686	686	100%	0	86%	KJ865407.1	GBGCA11774-15	
4	lg4	*Pampus chinensis*	*Piaractus mesopotamicus*	1206	1206	100%	0	99%	HQ420833.1	ANGBF6848-12	**Yes**
5	ma5	*Clarius batrachus*	*Heteropneustes fossilis*	1184	1184	100%	0	99%	AP012013.1	ANGBF10687-12	**Yes**
6	bp6	*Clupisoma prateri*	*Clupisoma prateri*	1173	1173	99%	0	99%	JX983274.1	No sequence	No
7	cg7	*Gudusia chapra*	*Tenualosa ilisha*	1201	1201	100%	0	99%	AP011611.1	CYTC3774-12	**Yes**
8	cj8	*Gudusia chapra*	*Tenualosa ilisha*	1201	1201	100%	0	99%	AP011611.1	CYTC3774-12	**Yes**
9	bj9	*Bagarius bagarius*	*Bagarius bagarius*	1179	1179	97%	0	99%	EU417762.1	ANGBF5989-12	No
10	ta10	*tor putitora*	*Mylopharyngodon piceus*	1201	1201	100%	0	99%	HQ236003.1	GBGC6787-09	**Yes**

Note: 8 (out of 10) samples

APPENDIX-4

Neighbor joining phylogentic tree of cn3 (sample collected from a supershop, sold as chewa *(Pseudapocryptes elongotus)* with other sequences downloaded from Gene Bank and BOLD systems based on CO1 gene.

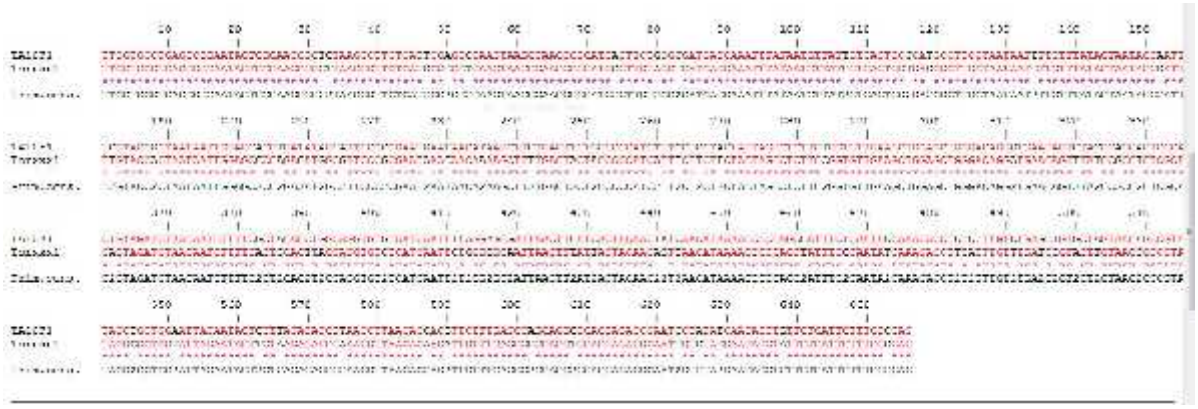
APPENDIX-5



Plylogenetic tree of all sequences of collectedy samples, species sequence which told by seller and downloaded sequences.

APPENDIX-6

Alingement of Sequence of ta10 and *Tor putitora*



Sequence alignment of two spp and confirmation of actual species.

Identification of polymorphic sites between of ta10(collected) and *Tor putitora*

Position	Ta10	original, *Tor putiora*
50	A	G
80	G	A
230	T	C
260	T	A
390	C	T
440	T	A
500	G	A
590	G	A

APPENDIX-7

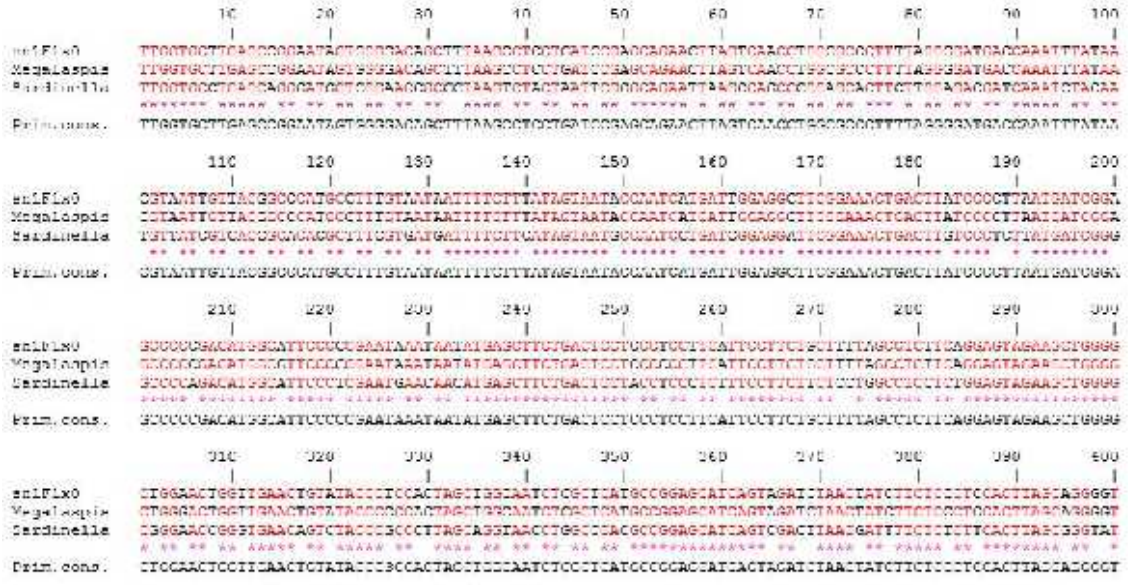


Figure 11. Multiple Sequence alignment among th sample sequence , sequence of sold species and indentified species after submission the sequence at BOLD database

Visualization at a glance the polymorphic sites.

Position	Sn1	*Megalaspis cordyla*	*Sardinella lonceps*
20	A	A	G
50	A	A	G
80	A	A	T
110	T	T	C
140	T	T	C
200	A	A	G
230	T	T	C
260	A	A	T
320	A	A	C
350	T	T	C

APPENDIX-8

All Sequences of Collected Sample, Suspected Species Sequences & Real Species Sequences

>sn1F1

TTGGTGCTTGAGCCGGAATAGTGGGGACAGCTTTAAGCCTCCTGATCCGAGCAGAACTTAGTCAACCTGGCGCCCTTTT
AGGGGATGACCAAATTTATAACGTAATTGTTACGGCCCATGCCTTTGTAATAATTTTCTTTATAGTAATACCAATCATG
ATTGGAGGCTTCGGAAACTGACTTATCCCCTTAATGATCGGAGCCCCCGACATGGCATTCCCCCGAATAAATAATATG
AGCTTCTGACTCCTCCCTCCTTCATTCCTTCTGCTTTTAGCCTCTTCAGGAGTAGAAGCTGGGGCTGGAACTGGTTGAAC
TGTATACCCTCCACTAGCTGGCAATCTCGCTCATGCCGGAGCATCAGTAGATCTAACTATCTTCTCCCTCCACTTAGCA
GGGGTCTCATCAATCCTTGGAGCTATTAATTTCATTACTACGATTATTAATATAAAACCGCCTGCAGTTTCAATATACC
AAATTCCATTATTTGTCTGAGCCGTGCTGATTACAGCCGTCCTCCTCCTCCTCTCTCTTCCAGTCTTAGCTGCTGGGATC
ACGATACTTCTCACAGACCGAAACCTAAACACTGCCTTCTTTGATCCGGCAGGAGGTGGAGATCCAATTCTTTATCAAC
ACCTATTCTGATTCTTTGGCCAC

>**mn2F1**

TTGGTGCTTGAGCCGGaATAGTGGGGACAGCTTTAAGCCTCCTGATCCGAGCAGAACTTAGTAACCTGGCGCCCTTTTA
GGGGATGACCAAATTTATAACGTAATTGTTACGGCCCATGCCTTTGTAATAATTTTCTTTATAGTAATACCAATCATGA
TTGGAGGCTTCGGAAACTGACTTATCCCCTTAATGATCGGAGCCCCCGACATGGCATTCCCCCGAATAAATAATATGA
GCTTCTGACTCCTCCCTCCTTCATTCCTTCTGCTTTTAGCCTCTTCAGGAGTAGAAGCTGGGGCTGGAACTGGTTGAACT
GTATACCCTCCACTAGCTGGCAATCTCGCTCATGCCGGAGCATCAGTAGATCTAACTATCTTCTCCCTCCACTTAGCAG
GGGTCTCATCAATCCTTGGAGCTATTAATTTCATTACTACGATTATTAATATAAAACCGCCTGCAGTTTCAATATACCA
AATTCCATTATTTGTCTGAGCCGTGCTGATTACAGCCGTCCTCCTCCTCCTCTCTCTTCCAGTCTTAGCTGCTGGGATCA
CGATACTTCTCACAGACCGAAACCTAAACACTGCCTTCTTTGATCCGGCAGGAGGTGGAGATCCAATTCTTTATCAACA
CCTATTCTGATTCTTTGGCCAG

>cn3F1

TTGGTGCTTGGGCCGGAaTAGTAGGCACAGCCCTAAGCCTTCTAATTCGTGCTGAACTGAGCCAACCAGGAGCCCTTCT
TGGAGACGATCAGATCTATAATGTAATTGTAACAGCTCATGCCTTTGTAATAATTTTCTTCATAGTTATACCTGTAATG
ATTGGAGGCTTTGGAAACTGACTTGTACCCCTCATGATTGGAGCACCAGACATGGCCTTCCCTCGAATGAACAACATA
AGCTTCTGACTCCTTCCCCCCTCATTTCTCCTTCTCCTTGCATCTTCAGGGGTAGAAGCTGGAGCTGGAACAGGATGAA
CAGTTTATCCCCCACTTGCAGGCAATCTTGCCCATGCAGGAGCTTCTGTTGACTTAACCATTTTTTCACTTCACCTAGCC
GGTATTTCTTCAATTCTAGGGGCTATTAATTTCATTACTACAATTTTAAATATGAAACCCCCTGCCATTTCACAATATCA
AACACCCCTTTTCGTGTGAGCTGTACTAATTACAGCTGTGTTGCTCTTATTATCCCTTCCCGTCCTAGCTGCTGGCATTA
CAATACTTCTAACAGACCGAAACCTAAATACAACCTTCTTTGACCCAGCTGGAGGGGGAGATCCAATTCTTTA
TCAACATCTATTCTGATTCTTTGGCCAC

>**lg4F1**

TTGGTGCCTGAGCCGGAaTAGTTGGAACGGCCCTTAGCCTCTTAATTCGAGCGGAGCTAAGCCAACCCGGATCCCTCTT
AGGTGATGACCAGATCTATAATGTTATCGTTACTGCGCACGCCTTCGTAATAATTTTCTTTATAGTAATACCAATTATG
ATTGGAGGCTTCGGGAATTGATTGGTTCCCCTAATGATTGGTGCACCCGACATAGCATTCCCACGAATAAATAATATAA
GCTTCTGACTCCTACCCCCATCCTTCCTTCTTCTGCTAGCATCCTCAGGAATCGAAGCCGGAGCAGGGACAG
GCTGAACTGTATATCCCCCTCTTGCCGGTAACCTCGCACACGCGGGCGCCTCTGTTGACCTAACCATCTTTTCACTTCAT
CTTGCTGGGGTTTCCTCCATCCTTGGGGCTATTAACTTCATTACAACTATTATTAACATGAAGCCTCCAGCCATTTCACA
ATATCAAACACCCCTATTTGTATGAGCAGTCCTAATCACTGCCGTTCTTCTTCTTCTCTCGCTACCAGTTCTGGCTGCCG
GAATTACTATACTTCTGACAGATCGAAACCTTAACACCACATTCTTTGACCCCGCGGGGGGAGGAGACCCAATTCTCTA
CCAACATTTATTCTGATTCTTTGGCCA

>ma5F1

TTGGTGCCTGAGCCGGAaTAGTCGGCACAGCCCTTAGCTTACTTATCCGGGCAGAATTAGCACAACCTGGTGCTCTACT
GGGTGATGACCAAATTTATAACGTTATTGTTACTGCTCACGCATTCGTAATAATTTTCTTTATAGTAATACCAATTATGA
TTGGAGGCTTCGGAAACTGACTAGTGCCCCTAATGATTGGAGCCCCTGATATAGCATTTCCACGTATGAATAATATAAG
CTTCTGACTACTTCCACCATCTTTCCTACTACTGCTTGCATCTTCTGGAGTTGAAGCGGGGGCAGGAACAGGATGAACA
GTGTATCCACCTCTTGCTGGGAATCTTGCACATGCTGGAGCCTCAGTAGATTTAACCATTTTCTCCCTACACTTAGCAG
GTGTCTCATCTATTCTAGCATCTATTAATTTTATTACTACTATTATTAACATGAAACCCCCAGCCATCTCACAATATCAA
ACACCACTATTTGTTTGATCAGTGTTAATTACAGCCGTACTACTACTACTCTCCCTACCTGTACTAGCCGCTGGAATTAC
CATACTACTAACTGACCGAAATCTAAACACTACATTCTTTGACCCCGCAGGAGGTGGAGACCCCATTCTCTA
CCAGCATCTCTTCTGATTCTTTGGCCAC

>bp6F1

TTGGTGCCTGAGCCGGAaTAGTTGGCACAGCCCTTAGCCTACTAATTCGGGCAGAACTAGCCCAACCTGGTACTCTACT
GGGCGATGACCAGATTTATAATGTTATTGTTACTGCCCATGCCTTCATCATAATTTTCTTTATAGTAATACCAATCATAA
TTGGAGGATTTGGAAATTGACTCGTTCCCCTAATGATTGGGGCACCAGACATGGCATTCCCTCGAATAAATAACATAA
GCTTCTGATTACTACCCCCATCTTTCCTGCTACTTCTTGCCTCATCTGGAGTTGAAGCAGGAGCAGGAACAGGGTGAAC
TGTATACCCCCCTCTCGCTGGCAACCTGGCACATGCAGGAGCTTCTGTAGATTTAACTATCTTCTCCCTCCACCTTGCTG
GGGTTTCATCAATTTTAGGAGCAATTAATTTTATTACAACTATTATTAATATGAAACCCCCAGCTATTTCACAGTATCA
AACACCTCTATTTGTATGAGCCGTATTAATTACAGCCGTACTACTTCTGCTGTCTCTACCAGTATTAGCCGCTGGGATTA
CAATACTACTAACAGATCGAAACCTAAATACCACATTCTTCGACCCGGCAGGGGGAGGAGATCCAATTCTTTATCAAC
ACCTTTTCTGATTCTTTGGCCAC

>cg7F1

TTGGTGCTTGGGCAGGGATGGTTGGCACTGCTTTGAGTCTTCTGATTCAAGCTAAGCTAATCCAACCCGGGTCTCTCCA
TGGAAATGACcAAATCTATAATGTTATCGTTACTGCGCATGCCTTTGTGATGATTTTCTTTATAGTGATGCCCATTCTGA
TTGGGGGGTTTGGAAACTGACTGGTTCCCCTGATAATTGGAGCACCCGACATAGCATTCCCCCGAATAAACAACATAA
GCTTCTGACTCCTCCCTCCCTCCTTTCTCCTCTTGCTCTCTTCCTCTGGAATGGAAGCTGGAGCGGGCACAGGATGAACT
GTATACCCCCCCTTGGCCGGTAACTTGGCCCACGCAGGAGCATCTGTTGATTTAACCATTTTCTCTCTACACCTGGCAG
GGATTTCCTCTATTCTTGGGGCAATCAACTTTATTACCACAATTATTAATATGAAGCCCCCTGCCATTTCACAATATCAA
ACACCACTATTTGTCTGAGCTGTCCTTGTAACTGCCGTTCTCCTTCTGCTCTCGCTTCCTGTCTTGGCTGCCGGAATTACT
ATGCTGCTGACAGATCGAAACCTAAACATCACATTCTTTGACCCTGCAGGGGGAGGAGACCCAATTCTCTACCAACAC
TTATTTTGATTCTTTGGCCAC

>cj8F1

TTGGTGCTTGGGCAGGGATGGTTGGCACTGCTTTGAGTCTTCTGATCCGGGCAGAGCTAAGCCAACCCGGGTCTCTCCT
TGGAAATGACCAAATCTATAATGTTATCGTTACAGCGCATGCCTTTGTGATGATTTTCTTCATAGTGATGCCCATTCTGA
TTGGGGGGTTTGGAAACTGACTGGTTCCCCTGATAATTGGAGCACCCGACATAGCATTCCCCCGAATAAACAACATAA
GCTTCTGGCTCCTCCCTCCCTCGTTTCTCCTCTTGCTCTCTTCCTCTGGAGTGGAAGCTGGGGCGGGCACAGGATGAACG
GTATACCCCCCCTTGTCCGGCAACTTGGCCCACGCAGGAGCATCTGTTGATTTAACCATTTTCTCTCTACACCTGGCAG
GTATTTCGTCTATTCTTGGGGCAATCAACTTTATTACCACAATTATTAATATGAAGCCCCCTGCTATCTCACAATATCAA
ACACCACTATTTGTCTGAGCTGTCCTTGTACTGCCGTTCTCCTTCTGCTCTCCCTTCCTGTCTTGGCCGCCGGAATTACC
ATGCTGCTAACAGATCGAAACCTAAATACCACATTCTTTGACCCTGCAGGGGGCGGAGACCCAATTCTCTACCAACAC
CTCTTTTGGTTCTTCGGGCAC

>bj9F1

TCGGTGCTTGAGCTGGGATAGTTGGCACAGCTCTTAGCCTCCTAATTCGGGCAGAGCTAGCCCAACCTGGCGCCCTTCT
AGGCGATGACCAAATTTATAATGTCATTGTTACTGCTCACGCCTTTGTTATAATTTTCTTTATAGTAATACCAATCATGA
TTGGTGGGTTCGGCAACTGACTAGTGCCACTAATGATTGGAGCTCCCGACATGGCATTCCCTCGAATAAATAACATAA

GCTTCTGACTACTGCCCCCATCCTTTCTACTACTGCTTGCCTCTTCTGGTGTTGAAGCCGGGGCAGGAACAGGATGAAC
CGTATACCCCCCACTTGCAGGAAACCTCGCACATGCAGGAGCTTCCGTGGATTTAACTATTTTTTCACTGCATCTTGCA
GGAATTTCATCAATTCTAGGAGCCATCAACTTTATCACAACTATCATTAATATAAAACCTCCAGCGATCTCCCAGTACC
AAACACCATTATTCGTGTGGGCCGTCCTCATCACAGCAGTACTTCTCCTGCTCTCTCTGCCAGTACTTGCCGCGGGCAT
CACAATGTTATTAACAGACCGAAACCTAAACACCACCTTCTTTGACCCAGCAGGAGGAGGTGATCCAATCCTATATCA
ACATCTTTTCTGATTCTTTGGCCAC

>ta10F1

TTGGTGCCTGAGCCGGAaTAGTGGGAACCGCTCTAAGCCTTCTCATTCGAGCCGAACTAAGCCAACCCGGATCACTTCT
GGGCGATGACCAAATTTATAATGTTATTGTCACTGCCCATGCCTTCGTAATAATTTTCTTTATAGTAATACCAATTCTTA
TTGGAGGATTCGGAAACTGACTCGTACCGCTAATAATTGGAGCACCTGATATAGCATTCCCCCGAATGAATAACATAA
GCTTCTGACTTCTGCCCCCATCTTTCCTCCTACTACTAGCCTCTTCTGGTGTTGAAGCTGGAGCTGGGACAGGGTGAAC
AGTCTACCCACCACTCGCAGGCAATCTTGCACACGCAGGAGCATCTGTAGATCTAACAATCTTTTCGCTACACCTGGCA
GGTGTGTCATCAATTTTAGGAGCGATTAACTTCATCACTACAACTATCAACATAAAACCCCCAGCCATTTCTCAATACC
AAACACCTCTCTTTGTCTGAGCTGTGCTAGTAACAGCCGTACTCCTTCTCCTATCCCTACCAGTCCTAGCTGCTGGAATT
ACAATACTCCTTACAGACCGTAACCTTAACACCACGTTCTTTGACCCAGCAGGCGGAGGAGACCCAATCCTATATCAA
CACCTGTTCTGATTCTTTGGcCAC