J. Bangladesh Acad. Sci., Vol. 41, No. 1, 17-27, 2017

FRAUDULENCE DETECTION IN FISH MARKETING OF BANGLADESH USING DNA BARCODING

SAGORIKA SMRITI, TAHMINA ISLAM[1], MST KHADIZA BEGUM, NUSRAT JAHAN PUNOM, SHAWON MITRA[1], ANWAR HOSSAIN, RHAKA HARI SARKER[1] AND MOHAMMAD SHAMSUR RAHMAN[*]

Department of Fisheries, Faculty of Biological Sciences, University of Dhaka, Dhaka 1000, Bangladesh

ABSTRACT

DNA barcoding, based on the sequencing of a short, standardized region of the cytochrome c oxidase subunit 1 (COI) gene has the potential to be a practical method for fish species identification in raw or processed fish products. The aim of this study was to identify the level of fraudulence in fish marketing particularly in super shops of Bangladesh. Different species of fish as whole or in fillet were collected from different supermarkets and local markets of Dhaka metropolis for examination. COI gene sequence was amplified using random primers (Fish F1, Fish F2, Fish R1 and Fish R2) and then compared with reference sequences from GenBank and BOLD (Barcode of Life Data Systems). The database was able to provide species matches of >85% sequence similarity for ten samples tested. The overall fraudulence was detected 80% but specifically for whole fish it was 85.71% and fraudulence for fish fillet was 66.67%. These naming discrepancies and ambiguities demonstrate that DNA barcoding can be a reliable tool for the detection of fish products mislabeling in Bangladesh. Therefore, it can be a useful tool for fraudulence control, law enforcement and guiding consumers to get rid of fish product fraudulence.

Key words: Fraudulence, Fish Marketing, Fish fillet, DNA Barcoding, COI gene

INTRODUCTION

In Bangladesh, fishes including indigenous, exotic, freshwater and marine 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 is a need to produce more fish to fulfill the demand of growing population. This growing demand leads fisheries marketing to several fraudulent activities in various forms such as species substitution, mislabeling, capture of endangered species etc. 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

[*]Corresponding author: <shamsur@du.ac.bd>.
[1] Plant Breeding and Biotechnology Laboratory, Department of Botany, University of Dhaka, Dhaka 1000, Bangladesh.

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 fish fillet rather than whole fresh fish. 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.

Due to species substitution and mislabeling, 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 (Ugochukwu *et al.* 2015). 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 may also create potential health hazards for consumers with allergies to specific types of fish (Wong and Hanner 2008).

Therefore, to detect species substitution, to implement laws to prevent product substitution, there is an urgent 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. Molecular tools are advantageous for fish and fish products identification because of large number of fish species from distinct life 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 (Carvalho *et al.* 2011).

DNA barcoding has the potential to be a practical method for fish species identification. It aims to provide an efficient method using an array of species specific molecular tags derived from the 5′ region of the mitochondrial cytochrome c oxidase subunit 1 (COI) gene (Cawthorn *et al*. 2015). The utility of the method for fish species identifications is grounded on the premise that the COI sequence shows considerably greater interspecies variation than intraspecies, allowing for the differentiation of 97% of fish species (Ward 2009) and often being more discriminatory than alternative DNA markers used for this purpose (Cawthorn *et al*. 2013). 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). COI barcoding have recently been used to reveal disturbing rates of fish mislabeling (15–56%) in South Africa (Cawthorn *et al*. 2012).

The aim of this study was to identify the level of fraudulence in fish marketing in Bangladesh. The study was conducted to identify the species of fish in fresh or in fillet using DNA barcoding collected from commercial markets and super shops of Bangladesh and to compare the sequenced data with the databases of GenBank and BOLD systems.

MATERIALS AND METHODS

In the present study, 10 different samples of fish were examined. The samples were collected from different supermarkets and local markets of Dhaka metropolitan city. Fish samples were collected in sterilized plastic bag from different markets early in the morning during the periods of April 2015 to August 2015. Samples were then transported to the laboratory using icebox. The fish samples were processed within 2 hours of collection following aseptic techniques. Samples from muscle tissue were examined for each specimen.

DNA was isolated using the modified Sodium Dodecyl Sulfate (SDS) method of Koh *et al*. (1999). COI gene sequence (653 bp long) was amplified using random primers (Fish F1, Fish F2, Fish R1 and Fish R2) for all the experimental specimens (Ivanova *et al.* 2007; Yancy *et al.* 2008).

Table 1. Random primers used in the present study for amplification of COI gene fragments.

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

A 25 µl PCR reaction mixture was prepared containing 17.7 µl sterile de-ionized distilled water, 2.5 µl *Taq* Buffer 10X (Tris with 15 mM $MgCl_2$), 1.0 µl of each primer, 0.25 µl dNTPs 10 mM, 0.05 µl *Taq* DNA polymerase 5U/µl, 2.5 µl template DNA. PCR amplification was done in an oil-free thermal cycler (Biometra, UNO II). The program initially consisted of the following steps: 94°C for 4 min for initial denaturation, then 32 cycles at 94°C for 45s, 54°C for 1 min and 72°C for 1 min, followed by an extension step at 72°C for 5 min. Samples were then maintained at 4°C. After the reaction, the amplicons were separated by horizontal gel electrophoresis on a 1% agarose gel and visualized under UV-transilluminator.

For purification, PCR product was transferred to 1.5 ml eppendrof and 5 times FADF buffer was added and centrifuged at 11000 rpm for 1 min. Then the flow through was discarded and 750 ml buffer was added and centrifuged at 11000 rpm and liquid

discarded and again centrifuged at 14000 rpm for 10 minutes. Then 40 µl of elution buffer was added to the membrane centre of FADF column and waited for 3 minutes for full absorption and centrifuge at 14000 rpm for 2 minutes to elute the DNA. Then the liquid was further placed in column tube and centrifuged at 14000 rpm for 2 minutes.

Bi-directional sequencing were done using BigDye Terminator v 3.1 Cycle Sequencing Kit (Applied Biosystems, USA) according to manufacturer's instruction and capillary electrophoresis was done using ABI Genetic Analyzer (Applied Biosystems, USA). The generated sequences were then compared with reference sequences from different databases (GenBank and BOLD). BLAST was used for comparing primary sequence information. CLUSTALW was used for pairwise sequence alignment. This helped to find out the similarity, dissimilarity or identity between different sequences. MEGA v7.0 was used for the comparative analysis of molecular sequence data for reconstructing the phylogenetic tree.

RESULTS AND DISCUSSION

Full-length DNA barcodes (653 bp) were recovered using the Fish primers Fish F1, Fish F2, Fish R1 and Fish R2 with no detectable insertions, deletions or stop codons. DNA of 10 fish samples obtained from the super shops and local markets were amplified using the primer set and 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 GenBank and BOLD databases. For the 10 samples yielding interpretable COI barcodes, maximum sequence similarity values of 99% were achieved in BOLD and/or GenBank for the top species identifications.

Specimen sn1 was sold as 'Sardine' (*Sardinella longiceps*) but identified as *Megalaspis cordyla* with accession no. KM522836.1 in GenBank and accession no. AAB5271 as per BOLD database and the sold species is 99% similar with *Megalaspis cordyla*. Sardine is usually found in Indian Ocean; marketed fresh, salted, dried salted and also sold as smoked and canned (www.fishbase.org/summary/1511; accessed on 10 December 2016). Sardine fishery is highly commercial. On the other hand, *Megalaspis cordyla* is a pelagic species and generally occurs in inshore areas of the continental shelf; commercial value is lower than the *Sardinella longiceps*. Thus it is comprehensible that the name 'Sardine' might be preferably selected to appeal the local consumers.

Sample mm2 was sold as mackerel (*Rastreliger kanagurta*) but barcoding identified as *Megalaspis cordyla*. Accession no. of *Megalaspis cordyla* is KM522836.1 according to GenBank and AAB5271 according to BOLD systems. *Rastreliger kanagurta* was substituted by the lower value fish *Megalaspis cordyla*.

Table 2. Identification of collected fish samples using the GenBank and BOLD database.

Sample ID	Type of fish	Sold as	Barcode identification (percentages of GenBank similarities)	GenBank Accession	BOLD Accession	Mislabeled
sn1	Whole	Sardine *Sardinella longiceps*	*Megalaspis cordyla* (99%)	KM522836.1	AAB5271	**Yes**
mn2	Whole	Mackerel *Rastrelliger kanagurta*	*Megalaspis cordyla* (99%)	KM522836.1	AAB5271	**Yes**
cn3	Whole	Chewa *Pseudapocryptes elongatus*	*Scartelaos gigas* (86%) *Taenioides nigrimarginatus* (86%)	KT277705.1 KJ865407.1	No match	**Yes**
lg4	Whole	Lal chanda *Pampus chinensis*	*Piaractus mesopotamicus* (99%)	HQ420833.1	AAD6423	**Yes**
ma5	Fillet	Desi magur *Clarias batrachus*	*Heteropneustes fossilis* (99%)	KT364787.1	ACR4875	**Yes**
bp6	Whole	Bacha *Clupisoma prateri*	*Clupisoma prateri* (100%)	KT762369.1	ABA9729	**No**
cg7	Whole	Chapila *Gudusia chapra*	*Tenualosa ilisha* (95%)	AP011610.1	No match	**Yes**
cj8	Whole	Chapila *Gudusia chapra*	*Tenualosa ilisha* (99%)	AP011611.1	AAI0795	**Yes**
bj9	Fillet	Bagha ayer *Bagarius bagarius*	*Bagarius bagarius* (99%)	KX455910.1	AAE2095	**No**
ta10	Fillet	Mohashol *Tor putitora*	*Mylopharyngodon piceus* (99%)	AP011216.1	AAD9723	**Yes**

Note: 8 out of 10 samples were found as mislabeled; 85.71% mislabeling were found for whole fish and 66.67% for fish fillet.

Sample cn3 was sold as chewa (*Pseudapocryptes elongatus*) but identified as *Scartelaos gigas* which shows 86% similarity with GenBank accession no. KT277705.1, on the other hand no sequence is available yet in BOLD systems. The following phylogenetic tree shows the taxonomic position among sample cn3, *Pseudapocryptes elongatus, Scartelaos gigas* and *Taenioides nigrimarginatus* (Fig.1).

Fish sample lg4 was sold as 'lal chanda' (*Pampus chinensis*) in local markets but identified as *Piaractus mesopotamicus* which is non-native to Bangladesh but endemic to South American region (www.fishbase.org/summary/55383; accessed on 10 December 2016). *Pampus chinensis* is mainly distributed in Indo-Pacific region; feeds on centophores, salps, medusae and other zooplankton; harmless for human. To get the higher price, *Piaractus mesopotamicus* was sold as 'lal rupchanda'.

Sample ma5 was sold as fillet of 'desi magur' (*Clarias batrachus*) in super shops but identified as hybrid Shing (*Heteropneustes fossilis*) which is low value fish and the production cost is lower than to collect 'desi magur' from rural areas.

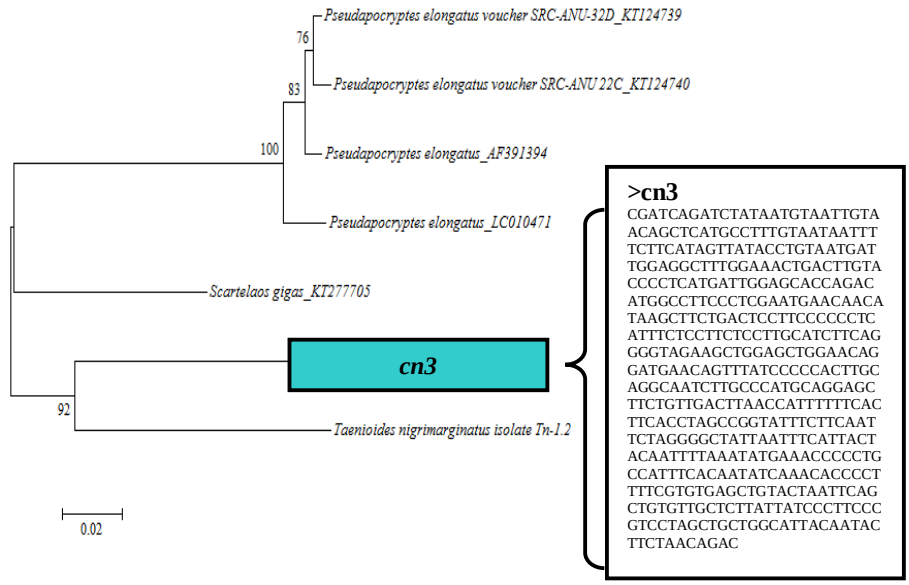

Fig. 1. Neighbor joining phylogentic tree of cn3 (sample collected from a supershop, sold as chewa *(Pseudapocryptes elongotus)* with other sequences downloaded from GenBank and BOLD systems based on COI gene.

Sample bp6 was sold as 'bacha' (*Clupisoma prateri*) in super shops and actually it is 'bacha' fish according to GenBank accession no. KT762369.1 and BOLD accession no. ABA9729.

Fish samples cg7 and cj8 were sold as 'chapila' which were collected from local markets (Gopibag and Jatrabari) but identified as *Tenualosa ilisha*. Capture of small size (<23 cm) *T. ilisha* which is known as 'Jhatka' is forbidden to catch but some dishonest businessman catch and sell them as 'chapila'.

Fish sample bj9 was sold as 'Bagha ayer' fillet and actually identified as 'Bagha ayer' with accession no. KX455910.1 and AAE2095 according to GenBank and BOLD Database, respectively.

Fish sample ta10 was sold as 'Mohashol' (*Tor putitora*) fillet in super shop but identified as Black carp (*Mylopharyngodon piceus*) according to GenBank and BOLD with accession no. AP011216.1 and AAD9723, respectively.

A total of 10 sequenced samples could be readily discriminated at the species level using DNA barcoding (either matching the species under which they were sold or being assigned to an alternative species). Out of 10, eight 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.

Comparison of COI gene sequence of sn1 with the available sequence from GenBank database has been found that the fish was not sardine. After comparing the sequence with that of *Sardinella longiceps*, 143 out of 653 nucleotide bases were found polymorphic (Fig. 2).

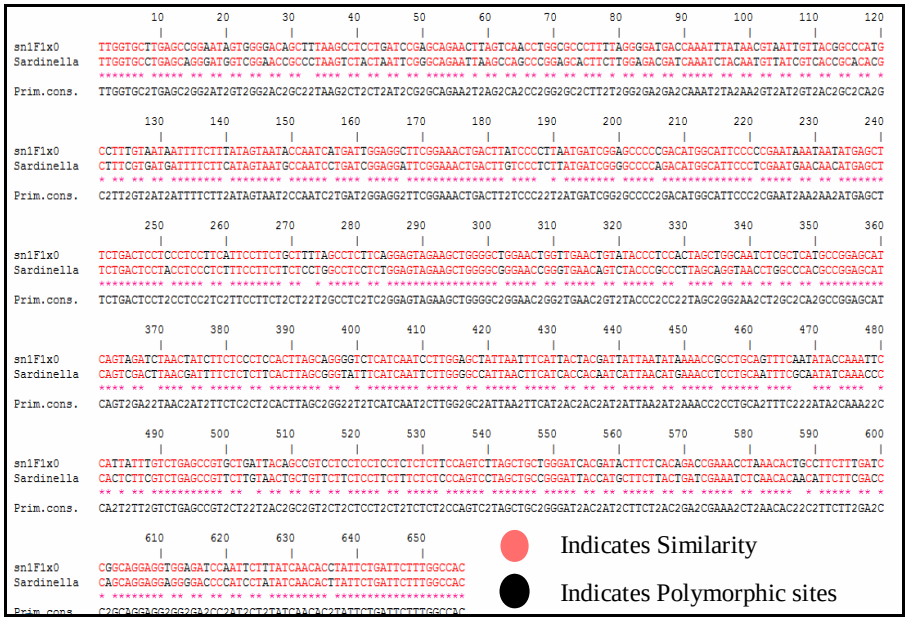

Fig. 2. Sequence alignment of COI gene fragment of 2 substituted fish species. Some representative polymorphic sites are indicated by nucleotide positions.

Table 3. Pairwise global alignment of COI gene fragment between 2 different substitute mislabeled fish species

Serial no.	Mislabeled species	Sequence length (bp)	Similarity (%)	Dissimilarity (%)
1.	*Sardinella longiceps* sn1	653	78.10 (510)	21.90 (143)
2.	*Rastrelliger kanagurta* mn2	636	80.66 (513)	19.34 (123)
3.	*Pseudapocryptes elongates* cn3	486	81.69 (397)	18.31 (89)
4.	*Pampus chinensis* lg4	630	44.13 (278)	55.87 (352)
5.	*Clarias batrachus* ma5	618	84.47 (522)	15.53 (96)
6.	*Gudusia chapra* cg7	639	84.04 (537)	15.96 (102)
7.	*Gudusia chapra* cj8	639	84.98 (543)	15.02 (96)
8.	*Tor putitora* ta10	639	86.23 (551)	13.77 (88)

Alignment length was 653 bp and identity was 78.10% similar with *Sardinella* and the difference was 21.90%. Table 3 also shows the similarity and dissimilarity of other sample sequences with their respective sold name sequences.

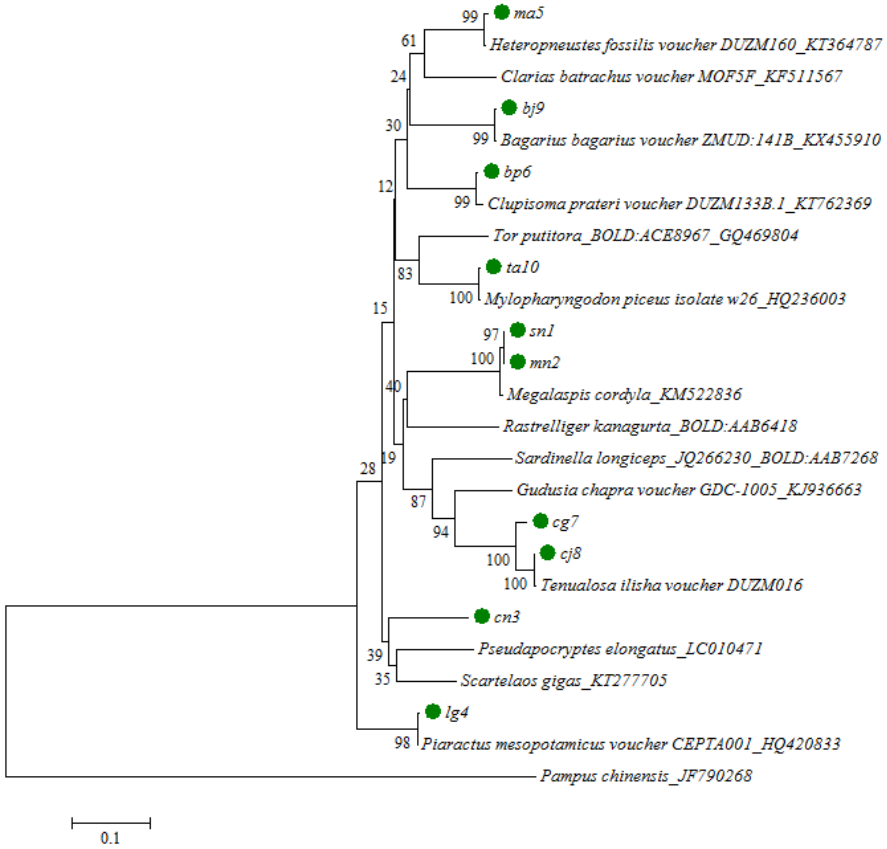

Fig. 3. Evolutionary relationships of taxa (10 sample sequences, sold name sequences and real species sequences)

The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 2.78828690 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the Phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 25 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 415 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

All amplified sequences that exceeded 650 nucleotides in length with no insertions, deletions, or stop codons were observed. Successful matches varied from 86 to 100% pairwise sequence identity (Table 2). Two sample sequences (cn3 and cg7) could not found any match 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 samples (fillets and whole fishes).

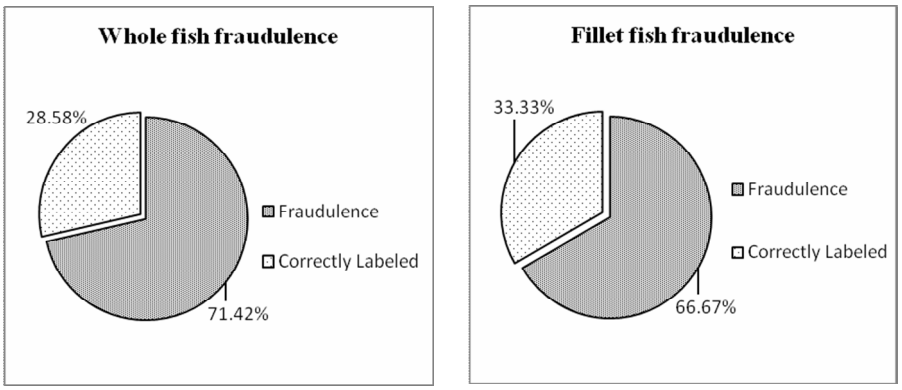

Fig. 4. Graphical Representation of the percentage of whole and fillet fish fraudulence.

Among the 10 samples sequenced, 7 samples were whole fish and the rest of the three were fillet. Six (71.42%) whole fish sample and two (66.67%) fillet sample identified as different species, and the overall fraudulence in fish market (both local and super shop) is 80% (Fig. 4). Mislabeling was not detected in 'Bagha Ayer' (*Bagarius bagarius*) and 'Bacha' fish *(Clupisoma prateri*) fillet samples. Barcoding therefore helped to ascertain the species in 100% cases of fish fillets.

In relation to international studies, this value corresponds with the seafood mislabeling 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 super shop as sardine (*Sardinella longiceps*) showed 99% sequence similarity with *Megalaspis cordyla*. According to the 'seafood list' published in the US (FDA, 2016; http://www.accessdata.fda.gov/scripts/fdcc/?set=seafoodlist&sort=SLSN&order=ASC&startrow=1&type=basic&search=Sardinella%20longiceps; accessed on 10 December 2016), 'Sardine' 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).

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 as well as with whole fish. Intentional mislabeling of cheaper fish products with high value fish is a more frequent phenomenon in both whole and fillet fish. Moreover, such substitution not only related to 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. When considering the International Union for Conservation of Nature (IUCN) Red List Status (IUCN Bangladesh 2000), two endangered species were being sold in marketplaces such as *Bagarius bagarius*, and another *is Tor putitora*.

The DNA barcoding technology is an improvement over the existing morphology-based techniques in detecting food fraud, particularly in fish markets. This work has provided an initial overview of species substitution and mislabeling in fish markets of Bangladesh. DNA barcoding offers a new level of precision in the application of species names, which is increasingly important in the expanding number of super markets in Bangladesh. Further, implementation of an appropriate monitoring regime is necessary to ensure standard names are properly applied.

ACKNOWLEDGEMENTS

This work was funded by the Ministry of Science and Technology of the Government of the People's Republic of Bangladesh under special allocation research fund for the financial year 2014-15.

REFERENCES

Carvalho, D. C., D. A. P. Neto, B. S. A. F. Brasil and D. A. A. Oliveira. 2011. DNA barcoding unveils a high rate of mislabeling in a commercial freshwater catfish from Brazil. *Mitochondrial DNA*. **22**: 97-105

Cawthorn, D., M., H. A. Steinman and R. C. Witthuhn. 2012. DNA barcoding reveals a high incidence of fish species misrepresentation and substitution on the South African market. *Food Res. Int.* **46**(1): 30-40.

Cawthorn, D. M., H. A. Steinman and L. C. Hoffman. 2013. A high incidence of species substitution and mislabelling detected in meat products sold in South Africa. *Food Cont.* **32**(2): 440-449.

Cawthorn, D. M., J. Duncan, C. Kastern, J. Francis and L. C. Hoffman. 2015. Fish species substitution and misnaming in South Africa: An economic, safety and sustainability conundrum revisited. *Food Chem.* **185**: 165-181.

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

Filonzi, L., S. Chiesa, M. Vaghi and F. N. Marzano. 2010. Molecular barcoding reveals mislabeling of commercial fish products in Italy. *Food Res. Int.* **43**: 1383–1388.

Hanner, R., S. Becker, N. V. Ivanova and D. Steinke. 2011. FISH-BOL and seafood identification: geographically dispersed case studies revel systemic market substitution across Canada. *Mitochondrial DNA* **22**: 106-122.

Helyar, S. J., H. D. Lloyd, M. de Bruyn, J. Leake, N. Bennett and G. R. Carvalho. 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.

IUCN Bangladesh. 2000. *Red book of threatened fishes of Bangladesh*. IUCN- The World Conservation Union (Bangladesh). Xii + 116 pp.

Ivanova, N. V., T. S. Zemlak, B. H. Hanner, and P. D. N. Hebert. 2007. Universal primer cocktails for fish DNA barcoding. *Mol. Ecol. Notes* **7**: 54-548

Koh, T. L., G. Khoo, L. Q. Fan and V. P. E. Phang. 1999. Genetic diversity among wild forms and cultivated varieties of Discus (*Symphysodon* spp.) as revealed by Random Amplified Polymorphic DNA (RAPD) fingerprinting. *Aquaculture* **173**: 485-497.

Leschin-Hoar, C. 2011. Specious species: Fight against seafood fraud enlists DNA testing. Affordable genetic sequencing and new database of "DNA bar codes" help combat the practice of selling one fish as another. Scientific American, November 2011 (www.scientificamerican.com/article/dna-testing-for-seafood-fraud/; accessed on 10 December 2016).

Ugochukwu, A. I., J. E. Hobbs, P. W. B. Phillips and R. Gray. 2015. An economic analysis of private incentives to adopt DNA barcoding technology for fish species authentication in Canada. *Genome* **58**: 559-567.

Ward, R. D. 2009. DNA barcode divergence among species and genera of birds and fishes. *Mol. Ecol. Resour.* **9**(4): 1077-1085.

Warner, K., W. Timme, B. Lowell and M. Hirschfield. 2013. Oceana Study Reveals Seafood Fraud Nationwide: Oceana (http://oceana.org/reports/oceana-study-reveals-seafood-fraud-nationwide; accessed on 10 December 2016).

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

Yancy, H. F., T. S. Zemlak, J. A. Mason, J. D. Washington, B. J. Tenge, N. L. Nguyen, J. D. Barnett, W. E. Savary, W. E. Hill, M. M. Moore, F. S. Fry, S. C. Randolph, P. L. Rogers and P. D. Hebert. 2008. Potential use of DNA barcodes in regulatory science: Applications of the Regulatory Fish Encyclopedia. *J. Food Prot.* **71**(1): 210-217.

(*Received revised manuscript on 20 April, 2017*)