Elsevier Editorial System(tm) for Food Control
Manuscript Draft

Manuscript Number: FOODCONT-D-15-00092R1

Title: DNA barcoding reveals commercial and sanitary issues in ethnic seafood sold on the Italian market

Article Type: Research Article

Keywords: DNA Barcoding, Mini DNA Barcoding, COI gene, seafood, mislabeling, traceability

Corresponding Author: Dr. Andrea Armani,

Corresponding Author's Institution:

First Author: Andrea Armani

Order of Authors: Andrea Armani; Lisa Guardone; Riccardo La Castellana; Daniela Gianfaldoni; Alessandra Guidi; Lorenzo Castigliego

Abstract: The number of seafood species sold on Western markets is constantly growing and many unconventional species are sold in ethnic food retailers. In this work, 68 ethnic seafood products variously processed were collected and a molecular analysis was performed by sequencing a full cytochrome c oxidase (COI) DNA barcode (FDB, ~655bp) or a mini COI DNA barcode (MDB, ~139bp) using universal primers. Barcodes were then compared with BOLD and GenBank. In addition, the label information was assessed according to the European legislation. By using the IDs analysis on BOLD a maximum species identity ≥98% was retrieved for 84% of the sequences. Of these, 67% were unambiguously identified at species level (51.3% of the FDB and 74% of the MDB). Using NCBI BLAST, 74% of the sequences scored a maximum species identity ≥98%, of which 73% were identified at species level (52% of the MFDB and 61% of the MDB). Both databases performed better in mollusk identification. Overall, 45 products (66%) were not correctly labelled according to the European requirements. Finally, the comparison between the molecular and the label analysis highlighted that 48.5% of the products presented discrepancies between labeling and molecular identification. In particular, sanitary implications were highlighted for 2 samples labeled as squid but identified as Lagocephalus spp., a poisonous puffer fish species banned from the EU market. The present results confirm DNA barcoding as a reliable tool for protecting health and economic interests of the consumers.

Dear Editor,

we would like to submit the following manuscript for possible publication:

"DNA barcoding reveals commercial and sanitary issues in ethnic seafood sold on the Italian market"

The removal of morphological characteristics induced by preparation and the entrance of new exotic species on the international market represent the main challenges in seafood identification. Moreover, the rapid growth of immigrants' settlements in Western countries has led to the appearance of unconventional seafood products sold in ethnic retails. Among the others, Chinese and Bangladeshi communities are well established on the Italian territory. Ethnic activities, despite a very good business organization, are often characterized by deficiencies in traceability systems. Seafood mislabeling is very frequent in the fishery supply chain and, other than putting consumers at risk of purchasing products not corresponding to their choice, it can represent a sanitary concern when toxic species are marketed and conservation issue if protected species are commercialized. DNA barcoding has been successfully used to enforce traceability regulations in the seafood chain because it is able to overcome the problems related to morphological identification. While Full DNA Barcoding (FDB-655 bp) performed very well when applied to fresh products, Mini DNA Barcoding (MDB-139bp) represents a valid alternative approach in case of processed products.

In this work 68 ethnic seafood products unprocessed or variously processed (dried, salted, roasted, smoked and canned) were purchased and a molecular analysis, based on DNA Barcoding (full or mini), was performed. In addition, the label information was assessed according to the European legislation. Of the total 204 sequences obtained, 158 FDB and 46 MDB were obtained. MDB were mainly obtained from roasted, smoked and canned products. By using the IDs analysis 67% of the sequence were unambiguously identified at species level (in particular, 51.3% of the FDB and 74% of the MDB). Using NCBI BLAST 73% of the sequences were identified at species level (52% of the MDB and 61% of the MDB). Both databases performed better in mollusks identification. Overall, the label analysis highlighted that 43 products (66%) were mislabeled according to the European requirements. The comparison between the molecular and the label analysis highlighted that 48.5% of the products presented discrepancies between labeling and molecular outcomes. In particular, sanitary implications were highlighted for 2 samples labeled as squid but identified as *Lagocephalus* spp., a poisonous puffer fish species banned from the EU market.

The present results thus confirm DNA barcoding as a reliable tool for protecting the health and economic interests of the consumers.

Best regards

Andrea Armani

***Highlights (for review)**

- Unconventional seafood species are sold in ethnic food retailers in Western countries
- DNA barcoding is a useful tool for seafood species identification
- Full and mini-DNA barcodes have been used for ethnic seafood identification
- Full and mini-DNA barcodes show high discriminatory ability
- Molecular and labeling analysis highlighted widespread mislabeling

Dear Editor,
we revised the manuscript as suggested by the Reviewers

Best Regards

Andrea Armani

Reviewers' comments:

Reviewer #1: This paper reports an in-depth investigation into the labelling of ethnic (Chinese and Bangladeshi) seafood products in Italy. The identities of the products were checked for accuracy using DNA barcoding. Full-length DNA barcoding was used where possible, but mini-length DNA barcoding was used for recalcitrant products. Sequences were checked against both BOLD and NCBI databases. Methods used were appropriate and the conclusions reached are valid. The authors are well aware of the various issues that can arise from DNA barcoding, including likely misidentification of a few specimens in the reference databases and the incompleteness of these databases, and took these limitations into account in reaching their conclusions.

The study reveals that about one half of the products investigated showed discrepancies between the label and the molecular analysis. The extent of this mislabelling, including pufferfishbeing mislabelled as squid, is very concerning.

Presentation of the study is generally good - apart from the English style that, although clearly understandable, needs some improvement. I outline a few improvements below, but more are needed. I did like the full presentation of the analysis of the individual items in extensive tables of Supplementary Material. These allow the interested reader to see exactly the results reached for these items, while the general reader will be satisfied with the summarised tables given in the body of the text. This approach keeps the published text to a reasonable length.

Specific comments

1. I wasn't enthusiastic about use of the term 'sanitary' in the title of this paper and many times in the text. This generally refers to hygiene facilities and clean drinking water, which is not what is meant here. I suggest it be replaced with the term 'health', which I think more accurately conveys the intended message in English.
The term ‘sanitary’ has been replaced with ‘health’.

2. I see that the full length DNA barcode (FDB) performed more poorly at unambiguous species delineation (51.3%) than the mini DNA barcode (MDB) (74%). See, for example, lines 302-306. This seems counter-intuitive to me, as I would have expected the FDB, with more information in it, to have been better able to discriminate sibling species than the MDB. The latter is surely more likely to be lacking any diagnostic characters that might separate closely-related species. The authors don't explain this apparent conundrum - I would like to see some brief interpretation/explanation of this given.
Dear reviewer, we were also surprised by these results. A possible explanation would be the higher variability of the fragment of the MDB respect to the whole region of the FDB. This explanation has been added in the text (L 309-301).

3. Table 1. I checked Lamendin et al.'s paper as I was surprised to see their sequencing success

was only 25.5%. In fact this was the failure rate - the success rate was 74.5%. So this table needs to be corrected.

The success rate in Table 1 has been modified.

4. This is a comment - not a suggestion for change: I wondered if the apparent better performance of the reference mollusc database was because many fewer people had contributed to it, and it isn't as well populated as the fish database. Maybe as more people contribute to it, and as more molluscan species are barcoded, perhaps more conflicts and uncertainties will arise….

We agree with your idea and we decided to report this "possible explanation" in the text (line 280-283).

Some English improvements:

1. Line 58 France
2. L63 (IFTN) (Ercsey-Ravasz,
3. L66 immigrant settlements
4. L68 countries
5. L74 most of the foreign food
6. L75other
7. L76 on Italian
8. L105 and are particularly
9. L107 (655bp), is the most common approach in the case
10. L108 (Table 1). This DNA
11. L110 In addition, mini

All the above modifications have been made

12. L117 study helped us to better

Modified according to the suggestion of reviewer 2

13. L124 food stores owned by (or, food stores run by)
14. L136 products comprising the same
15. L138 specimen of each
16. L184 (also see L187) 13 mollusk products (10 cephalopods, 2 bivalves….
17. L186 to find some
18. L187 We identified 57 fish
19. L194 eat". However, a consideration of the
20. L213 step allowed us to
21. L227 In fact, while universal…..DNA regions across…..assure DNA amplification of organisms …..
22. L231 primer pairs
23. L241 mollusk and crustacean DNA (no plurals in adjectives - also elsewhere)
24. L247 In total, 204
25. L277 were not so identified due to the inability of the system to discriminate very closely related species
26. L279 more effective in
27. L297/L299 species
28. L300 label, despite the latter being included in the (I haven't made suggestions for much of the rest of the paper. Perhaps the authors and/or editor should do this)
29. L389 area 61 (NW Pacific), the
30. L458 as pufferfish, Lagocephalus spp.
31. L461 attributed to a lack

All the above modifications have been made

32. Table 3SM why are some products shaded. They seem to be inconsistencies, but please clarify in legend.

An explanation has been added in the caption of the table.

33. Table 5SM why are some 'Total non conformities" boxes shaded? Please clarify in legend.

The grey background has been removed

Reviewer #2: REVIWER COMMENTS: Manuscript Number: FOODCONT-D-15-00092 Title: DNA barcoding reveals commercial and sanitary issues in ethnic seafood sold on the Italian market

General commentary:

This study uses a DNA barcoding approach to assess the extent of seafood mislabelling on the Italian market. As such, the work is of value for the following reasons: 1. It adds to the growing body of literature on the global problem of seafood mislabelling; 2. It specifically addresses ethnic food products, which are often overlooked in other comparative studies, and highlights a number of labelling and traceability concerns in this regard; 3. It compares the efficiency of DNA barcoding for finfish and molluscs, as well as the efficiency of sequence databases for making identifications; 4. It highlights specific health concerns as a consequence of seafood mislabelling. The work is generally thoroughly conducted and written up, however, there are a quite a number of recommendations that could improve the content for publication.

Widely speaking the following needs to be addressed:

1. The references need attention as many are missing or inconsistent

The references have been checked thoroughly.

2. In many cases in the text, especially near the end of the manuscript, the authors make extensive use of internet references, where the URL is placed in line in the text. This severely hampers the readability and these should rather be placed in the end reference list and be referenced in short in text.

When possible the URL been deleted (section 2.2) or replaced by reference (Line 384; line 366). In the other cases, it was impossible to find a name for the document.

3. In general, some parts of the discussion are long-winded, especially as pertains to the discussion on the success of sequencing and the ability of the two databases to identify the species. This is especially relevant since all the data is presented in Table form and the reader can easily judge most of this for themselves. It is suggested that these sections are shortened and only the salient points discussed in the text.

Dear reviewer, the discussion was already written trying to reduce it as much as possible. In fact, many details have been only reported in the tables (in particular in the supplementary materials). Thus, also considering the comment of the reviewer n. 1, we have not further shortened this section.

Specific comments:

Line 1 (Title): it is suggested that the word 'sanitary issues' in the title and throughout be replaced by something to the effect of 'potential health issues'. The word 'sanitary', in the reviewer's opinion, gives the impression of hygiene issues, which was not the case for the puffer fish.

Also considering the suggestion of the other reviewer, the word sanitary has been replaced with health.

Line 29: change to 'sold by ethnic food retailers' or 'sold in ethnic food outlets'.

Done

Line 30: it is suggested that the location of collection is mentioned in the abstract, e.g. "In this work, 68 ethnic seafood products variously processed were collected from the Italian market and a molecular analysis was performed…"
The location of collection has been added.
Line 32: Barcodes were then compared with sequences available in BOLD and GenBank….
Line 33: Should it not be IDS?
Line 37: remove the M in MFDB
Line 64: countries not Countries
Line 75: among the other ethnic, not others
Line 84: check spelling of labeling here vs. labelling in line 100
All the above modifications have been made.
Line 89: Carvalho, Palhares, Drummond &Frigo, 2015 is missing from reference list
The reference has been added.
Line 92: Armani, Castigliego & Guidi, 2012; and should this not be designated as Armani, Castigliego & Guidi, 2012a, since there is an Armani et al. 2012b in line 395.
Done

Line 102: should it not be Council Reg. (EC) as per the reference list: Council Regulation (EC) No 104/2000. It is also suggested that the various regulations throughout be better referenced to aid finding these in the reference list. E.g. (EC, 2000). Please check throughout.
The modification requested has been made. Furthermore we revised all the laws references paying attention to report exactly the same form in the text and in the reference list. In our opinion the laws must be reported in full.

Line 105-106: useful for seafood identification not 'the seafood identification'
Done

Line 110: should it be Hebert, Ratnasingham& de Waard, 2003a as per the reference list?
Done

Line 115: collected from Chinese and Bangladeshi retailers
Done

Line 116: Moreover, this study provided a better understanding of the preferences of these ethnic communities in terms of seafood and delineated their internal market network.
Done

Line 138: one specimen of each species
Done

Line 146: should standard marker not be two words?
Done

Line 185: check spelling of bivalves throughout
Done

Line 190: why is there … after filleted?
… have been removed

Line 213: allowed us to speed up and reduce
Done

Line 216-217: Hebert, Cywinska, & Ball, 2003b?
Done

Line 233: there are a number of places in the text where the authors refer to Table 1 to support a specific statement. It is recommended that the individual authors that had similar findings be referenced as this is too general, unless they all found it in all instances.
It has now been specified whev the statement was referred to specific studies reported in Table 1 (line 261).

Line 259: I do not think that 'sequencability' is a recognised word. Perhaps rather the 'ability to sequence'.
The sentence has been modified (Line 257)

Line 266: Lamendin et al., 2014 not in reference list
The right citation was Lamendin et al., 2015 as in the rest of the text and in the reference list

Line 279: in the case of mollusks. Please also check spelling of mollusc vs. mollusc - both are used in the text.
Done

Line 301: despite inclusion in the database
Sentence has been modified on the basis of the suggestion of reviewer 1

Line 325: allowed an unambiguous identification
Done

Line 388: despite a previous survey finding that …
Done

Line 393: In fact, even though after the accident the EU has imposed…
Done

Line 424: ….manage an ethnic retail shop can put on the market these kind of products.
Done

Line 453: Hu (2014) cannot be found in the reference list
This reference has been added to the reference list

Line 477: Coehen et al - appears misspelled compared to Cohen et al. 2009
Done

DNA barcoding reveals commercial and health issues in ethnic seafood sold on the Italian market

Armani A.*, Guardone L., La Castellana R., Gianfaldoni D., Guidi A., Castigliego L.

FishLab, Department of Veterinary Sciences, University of Pisa, Viale delle Piagge 2, 56124, Pisa (Italy).

***corresponding author**:

Postal address: FishLab, Department of Veterinary Sciences, University of Pisa, Viale dellePiagge 2, 56124, Pisa (Italy)

Tel: +390502210207

Fax: +390502210213

Email: andrea.armani@unipi.it

Abstract

The number of seafood species sold on Western markets is constantly growing and many unconventional species are sold in ethnic food outlets. In this work, 68 ethnic seafood products variously processed were collected from the Italian market and a molecular analysis was performed by sequencing a full cytochrome c oxidase (COI) DNA barcode (FDB, ~655bp) or a mini COI DNA barcode (MDB, ~139bp) using universal primers. Barcodes were then compared with sequences available in BOLD and GenBank. In addition, the label information was assessed according to the European legislation. By using the IDs analysis on BOLD a maximum species identity ≥98% was retrieved for 84% of the sequences. Of these, 67% were unambiguously identified at species level (51.3% of the FDB and 74% of the MDB). Using NCBI BLAST, 74% of the sequences scored a maximum species identity ≥98%, of which 73% were identified at species level (52% of the FDB and 61% of the MDB). Both databases performed better in mollusk identification. Overall, 45 products (66%) were not correctly labelled according to the European requirements. Finally, the comparison between the molecular and the label analysis highlighted that 48.5% of the products presented discrepancies between labeling and molecular identification. In particular, health implications were highlighted for 2 samples labeled as squid but identified as *Lagocephalus* spp., a poisonous puffer fish species banned from the EU market. The present results confirm DNA barcoding as a reliable tool for protecting health and economic interests of the consumers.

Keywords: DNA Barcoding, Mini DNA Barcoding, COI gene, seafood, mislabeling, traceability

1. Introduction

Innovations in processing, transportation and technology have facilitated the trade of a great variety of seafood species worldwide. Nowadays, about 1200 species are commercialized in the European Union (EU). In fact, the EU is the largest market for fishery products and, in 2012, it has absorbed 36% of world importations, which are often made of prepared products. Italy, together with Spain and France, has covered almost 60% of the EU expenditure for seafood importation (EUMOFA, 2014; FAO, 2014a).

The removal of morphological characteristics induced by preparation (gutting, beheading and filleting), together with the continuous entrance of new exotic species, represent the main challenges in seafood identification by consumers and official controllers, while the intricate web of interactions, the International Agro-Food Trade Network (IFTN), (Ercsey-Ravasz, Toroczkai, Lakner & Baranyi, 2012) formed by food fluxes between countries, makes the tracking of goods increasingly complicated.

The rapid growth of immigrant settlements in Western countries has led to a further increase in ethnic food exchanges. In fact, the populations that have moved to USA and Europe have brought with them their own food cultures, increasing the diversity of food available in the host countries. Consequently, ethnic food has become increasingly popular and readily available in many supermarkets, restaurants, and shops (Lee, Hwang & Mustapha, 2014). Chinese communities, with many kinds of food business activities (restaurants, rotisseries and retail markets) are well established in Italy, where Chinese residents increased from 20000 to 210000 in the last twenty years (ISTAT, 2011). In particular, the Chinese communities of Prato and Milan are the largest in Italy and in Europe (Todarello, 2013). Interestingly, in these cities, most of foreign food business operators are Chinese (Fondazione Leone Moressa, 2010). Among the other ethnic groups that are present on Italian territory, the number of people coming from Bangladesh has risen from 5000 in 1992 to 74000 in 2009 (Tarquini, 2010). Currently Italy hosts more than 110000 Bangladeshi

people, representing the second European destination of Bangladeshi emigration. Noteworthy is the fact that around a fourth of the population is involved in activities linked to trade and sale (M.L.P.S., 2013).

Ethnic activities, despite a very good business organization, are often characterized by deficiencies in traceability systems. Our previous surveys put into light their difficulties to conform to the European rules, in particular regarding the enforcement of seafood labeling legislation (Armani *et al*., 2013; D'Amico *et al*., 2014).

Mislabeling, together with other unfair practices (intentional substitution, tampering, or misrepresentation of food) made for economic gain, belongs to the so-called Economically Motivated Adulterations (EMA) or simply food frauds (Spink & Moyer, 2011). Mislabeling is very frequent in the fishery supply chain, where it occurs in different forms and at any stage (Cawthorn, Steinman & Witthuhn, 2012; Carvalho, Palhares, Drummond & Frigo, 2015, Armani *et al*., 2015). A part of mislabeling is probably unintentional due to the fact that different species can be referred to by the same names in different regions and due to the lack of specific denominations, especially for new exotic species (Armani, Castigliego, Guidi, 2012; Lamendin, Miller & Ward, 2015). However, some traders could deliberately use mislabeling to trade illegally-caught fishery products into rightful fish markets. In fact, financial incentives represent the strongest motivation for fish mislabeling. Unfortunately, the consequences of mislabeling go beyond the economy: this practice undermines consumers' confidence in seafood products and distorts their perception on fish stock status, nullifying market-driven conservation efforts (Stiles *et al.*, 2011). Moreover, other than putting consumers at risk of purchasing products not corresponding to their ethical or taste criteria, mislabeled fish can represent a healthconcern when toxic species are marketed (Civera, 2003).

In the EU, considering that honest and accurate food labeling is essential to assure consumers' safety and aware choice, seafood must be labeled with the commercial and the scientific name, the production method, the catch area and the category of fishing gear (Council Regulation (EC) No

104/2000 and, starting from the 13th December the Regulation (EU) No 1379/2013). However, documental traceability is not always sufficient to fulfil these purposes.

DNA-based methodologies represent a valuable tool for food traceability and are particularly useful for seafood identification. DNA barcoding, based on the analysis of the first part of the cytochrome c-oxidase I (*COI*) gene sequence (655 bp) is the most common approach in the case of unprocessed products (Table 1). This DNA region usually shows a greater interspecific than intraspecific variation, allowing efficient discrimination among species (Hebert, Ratnasingham & de Waard, 2003). In addition, mini DNA barcoding (139bp) has been successfully used as an alternative approach for species identification in case of processed products (Armani *et al*., 2015). In fact, the amplification of a shorter region could represents the only chance to get molecular information from products containing degraded DNA.

In this study, full (FDB) and mini (MDB) DNA barcoding was applied for the identification of 68 ethnic processed and unprocessed seafood products collected in Chinese and Bangladeshi retailers in Prato and Pisa (Italy), with the aim to verify the traceability information. Moreover, this study provided a better understanding of the preferences of these ethnic communities in terms of seafood and delineated their internal market network

2. Materials and methods

2.1 Sample collection

Sixty eight ethnic fishery products (fish, mollusks and crustaceans), whole or prepared in various forms (filleted, pieces, threads), unprocessed (simply frozen) or processed (dried, salted, canned, roasted and smoked), variously packaged (in plastic bags, canned, under vacuum), were purchased in retail food markets within the Chinese communities of Prato and in food stores run by Bangladeshi people in Pisa in 2014 (Table 2 and Table 3). Samples were divided in 2 groups (unprocessed and processed) according to the definition provided by Reg. (EC) 852/2004 for unprocessed products: "*foodstuffs that have not undergone processing, and includes products that*

have been divided, parted, severed, sliced, boned, minced, skinned, ground, cut, cleaned, trimmed, husked, milled, chilled, frozen, deep-frozen or thawed".

Each product was brought to our laboratory where a visual inspection was performed according to a simple morphological analysis. Each product was registered by an internal code, photographed and stored (at room temperature or -20°C, depending on the kind of processing) until further analysis.

2.2 Molecular analysis

2.2.1 Tissue sampling, DNA extraction and evaluation of DNA fragmentation by gel electrophoresis. In case of bulk products comprising the same species, at least three samples were taken. In case of a single package (ETN54), composed of 8 different species (Table 1SM), the collection was performed from at least 1 specimen of each species.

Total DNA extraction was performed starting from 100 mg of tissue as described by Armani *et al.,* (2014). Dried and salted samples were washed and re-hydrated in running tap water overnight. The DNA quality and quantity was determined with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, US).

One thousand nanograms of the total DNA extracted from the samples was electrophoresed on 1% agarose gel GellyPhorLE (Euroclone,Wetherby, UK), stained with GelRed™ Nucleid Acid Gel Stain (Biotium,Hayward, CA, USA), and visualized via ultraviolet transillumination. DNA fragment size was estimated by comparison with the standard marker SharpMass™50-DNA ladder and SharpMass™1-DNA ladder (EurocloneS.p.A-Life Sciences Division, Pavia, Italy).

2.2.2 Amplification and sequencing of the full-COI barcode (FDB). A 655-658bp fragment of the *COI* gene was firstly amplified from the DNA extracted from all specimens, using two universal primer pairs (PP1 and PP2) for the *COI* region (Table 4). The following PCR protocol was applied: 20 µl reaction volume containing 2 µl of a 10x buffer (5Prime, Gaithersburg, USA), 100 µM of each dNTP (Euroclone, Pavia, Italy), 250 nM of forward primer, 250 nM of reverse primer, 25ng/µL of BSA (New England BIOLABS® Inc. Ipswich, MA, USA), 2.5 U PerfectTaq DNA

Polymerase (5Prime, USA), 100 ng of DNA and DNase free water (5Prime, USA) with the following cycling program: denaturation at 94 °C for 3 min; 45 cycles at 94°C for 30s, 47-53°C (depending on the primer pair, see Table 4) for 30s, 72°C for 35s; final extension at 72°C for 10 min. Five µL of PCR products were checked by electrophoresis on a 1.8% agarose gel and the presence of expected amplicons was assessed by a comparison with the standard marker SharpMass™50-DNA ladder. Amplicons were purified and sequenced by High-Throughput Genomics Center (Washington, USA).

2.2.3 Amplification and sequencing of the mini-COI barcode (MDB). The DNA of the samples that failed amplification of the FDB region was submitted to the amplification of a ~190bp MDB region (139bp without primers) with the primer pair FISHCOILBC_ts/REVshort1 (Table 4). The PCR was performed following Armani *et al.,* 2015. All the PCR products were purified and sequenced as reported in section 2.2.2.

2.2.4 DNA barcode sequence analysis and comparison with databases. The obtained sequences were analyzed using Clustal W in Bio Edit version 7.0.9. (Hall, 1999). Fine adjustments were manually made after visual inspection. All the sequences were used to run a BLAST analysis on GenBank and analyzed using the Identification System (IDs) on BOLD (Species Level Barcode Records) to assess the concordance between the label information and the molecular analysis. A top match with a sequence similarity of at least 98% was used to designate potential species identification (Barbuto *et al*., 2010). Since the *COI* sequences obtained in this study were not derived from voucher samples or expertly-identified fish specimens, these sequences were submitted neither to GenBank nor to BOLD.

2.3 Labeling analysis

The analysis has been performed on the product labels as described in D'Amico *et al*., (2014). In particular, the information reported on the label were assessed in the light of the requirements of the Council Regulation (EC) No 104/2000.

3. Results and Discussion

3.1 Sample collection

Even though, according to the label, the collected products were 54 fish products, 13 mollusks products (of which 10 cephalopods, 2 bivalvs and 1 gastropod), and 1 crustacean product, the visual inspection, subsequently confirmed by the molecular analysis, allowed us to find some inconsistencies. We identified 57 fish products, 10 mollusks (of which 7 cephalopods, 2 bivalvs and 1 gastropods), and 1 crustacean (Table 3).

Of the 68 products collected, 31 were whole and 37 variously prepared (beheaded, gutted, filleted…); 19 (28%) were only frozen (unprocessed), while the remaining 49 (72%) were variously processed: in particular, 2 were only dried and 6 were only canned, while all the others had undergone more than one type of processing. Interestingly, all the 11 Bangladeshi products were unprocessed, while 41 (72%) of the 57 Chinese ones were produced using more than one method (Table 2). In fact, 24 Chinese products (42.1%) were "ready to eat". However, a consideration of the most recent studies using DNA barcoding for seafood authentication highlighted that mainly unprocessed products had been sampled and analyzed (Table 1). Moreover, in some of these studies frozen samples have been considered as processed, despite the definition provided by Regulation (EC) 852/2004 (see Section 2.1).

3.2 Molecular analysis

3.2.1 DNA extraction and evaluation of DNA fragmentation by gel electrophoresis. All the DNA samples extracted showed good values of quality and quantity after spectrophotometric analysis.

The electrophoretic analysis of total DNA showed that DNA samples extracted from processed products were more degraded than those obtained from unprocessed products. In fact, heat exposure, low pH, and drying can produce depurination and hydrolysis provoking fragmentation of DNA molecules (Teletchea, 2009). These effects have also been widely observed by several other studies (Table 1). It is interesting to note that we also confirmed the high level of degradation already observed in case of fresh/frozen products (Lamendin *et al.*, 2015; Armani *et al.*, 2015).

206 The analysis of total DNA was used as a criterion to divide the samples in 2 different groups,
207 according to the level of degradation (low and high), defined on the basis of the fragment length. In
208 particular, samples that showed DNA fragments shorter than 200bp were considered highly
209 degraded.
210 By providing useful information on the DNA quality, this simple preliminary evaluation step
211 allowed us to speed up and reduce the cost of the analysis by optimizing the amplification
212 procedure (see Section 3.2.2).
213 *3.2.2 Amplification and sequencing.* Even though there is no formula that can predict the length
214 of the sequence that must be analyzed to ensure species-specific diagnosis (Hebert, Cywinska, &
215 Ball, 2003) it is evident that the longer the sequence, the greater the amount of information it
216 contains. This evidence could explain why previous studies have tried to obtain a FDB from DNA
217 samples extracted from both unprocessed and processed products. However, in most of the cases
218 they failed in obtaining long sequences (Table 1).
219 In this study, considering the results of the electrophoretic analysis, only DNA samples that did
220 not show high level of degradation were submitted to the amplification of FDB with PP1. However,
221 using this PP on DNA samples of molluscs and crustaceans, no PCR products could be obtained
222 showing that PP1 was able to amplify only fish DNA. For this reason, an additional PP was
223 introduced (Table 4). In particular, considering that other than fish, most of the products were
224 composed of mollusks, the primers of Mikkelsen, Bieler, Kappner, & Rawlings, (2006) (PP2) were
225 selected. In fact, while universal primers are designed to amplify conserved DNA region across
226 different species groups, they cannot assure DNA amplification of all kind of organisms belonging
227 to different taxa (Carrera *et al*., 2000).
228 Finally, in case of degraded DNA samples and when it was impossible to obtain FDB with the
229 aforesaid primer pairs, the PP3 was used for the amplification of a MDB. In fact, previous studies
230 have highlighted, on one hand, the impossibility to obtain a FDB (~ 655bp) in the case of processed
231 products (Table 1) and, on the other, the potentialities of MDB in species discrimination (Armani *et*

232 *al*., 2015). Therefore, our study has put into practice an approach often advocated by previous
233 authors (Holmes, Steinke, & Ward, 2009; Haye, Segovia, Vera, Gallardo, & Gallardo-Escárate,
234 2012; Cawthorn *et al*., 2012).

235 At least one PCR product (FDB or MDB) was amplified with one of the PP for all the sampled
236 products, giving an overall amplification success of 100%.

237 Using PP1, we succeeded in the amplification of 42 out of 57 fish DNA samples (74%) (Table
238 3). One fish sample (identified as *Carcharhinus brachyurus*, Table 1SM) was only amplifiable with
239 PP2. With the PP2 we amplified 8 of 11 (73%) mollusk and crustacean DNA samples. Finally, the
240 18 DNA samples (from both fish and mollusks) that failed the previous amplification were all
241 amplified using the PP3, demonstrating the potentiality of the REVshort1 primer, initially designed
242 for Porgies species, to amplify organism belonging to different and distant taxa.

243 At least one sequence was obtained for 63 products. Therefore, the overall rate of sequencing
244 success was 93%. The number and percentage of sequencing success for each product category are
245 reported in Table 3. In total, 204 sequences were obtained from the 68 products collected. Of these
246 sequences, 173 were obtained from the 57 fish products, 27 from the 10 mollusks products and 4
247 from the crustacean product. Concerning fish products, for 42 of them, at least one FDB was
248 obtained, producing 133 long sequences (average length 632.6 bp), while for other 11 products only
249 40 MDB were produced (average length 139 bp). In the case of mollusks products, 21 FDB
250 (average length 653.9) were obtained from 7 products, and 6 MDB (average length 139bp) from the
251 remaining 2 products. Finally, for the crustacean samples 4 FDB (average length 658bp) were
252 produced. No insertions, deletions or stop codons were observed in the *COI* sequences. In
253 particular, in the case of MDB, no nuclear DNA sequences originating from mtDNA (NUMTs),
254 described by Zhang & Hewitt (1996), were sequenced.

255 As already observed during PCR amplification, sequencing was also affected when DNA
256 extracted from thermally treated products was used. While drying and salting, even combined with
257 freezing, affected less the recovery of sequences, thermal processing (roasting and canning) and

smoking negatively affected the recovery of FDB. In particular, in the case of the 6 canned products it was possible to obtain only MDB (Table 2). Low percentage of FDB amplification and limitation in the length of recovered FDB has been widely observed in processed products analyzed in previous studies (Cawthorn *et al*., 2012; Haye *et al*., 2012). However, also in the case of frozen products, which in this study were considered as unprocessed, the sequencing success was lower than 100% (Table 2). This failure could be due to a certain DNA degradation that it is known to occur not only in processed products, but also in fresh and frozen products, even though to a less extent (Armani *et al*., 2015; Lamendin *et al*., 2014). In fact, during transportation, handling and storage, fish could have suffered freezing/thawing processes that can influence DNA quality and affect the success of barcoding analysis. This is even more plausible if considering that products sold in small ethnic retail markets often suffers from the lack of good manipulation or storage procedures.

3.2.3 Comparison with the databases. During the comparison with the databases the raw data have been analyzed, to solve ambiguous results that could create misidentification, according to the revision process used by Armani *et al*., (2015). In fact, a low degree of reliability for some sequences in the databases has already been observed and discussed (Landi *et al*., 2014).

By using the IDs analysis on BOLD a maximum species identity in the range of 98–100% was obtained for 172 sequences (84%) (Table 3). Of these, 115 (67%) were unambiguously identified at species level while the remaining 57 sequences were not so identified due to the inability of the system to discriminate very closely related species (Table 1SM). We found that the system was more effective in identifying mollusks species than fish species. In fact, in case of mollusks, all the 24 sequences with a match higher than 98% were unambiguously identified, while in case of fish only 63% (Table 1SM). This apparent better performance of the reference mollusk database could be due to the fact that fewer studies had contributed to it and, currently, it is not as well populated as the fish database. Maybe as more molluscan species are barcoded, more conflicts and uncertainties will arise as in the case of fish species.

The IDs analysis did not identify among fish species belonging to Engraulidae, Tetraodontidae, Gadidae, and Carangidae. In the Engraulidae family, the system was not able to discriminate among different species of the genus *Engraulis*, giving overlapping values of identity ranging between 98-100%. The same ambiguous results using the *COI* gene had been obtained in the study of Ardura, Planes, & Garcia-Vazquez, (2013). Also in the case of the Tetraodontidae family, the system was not able to correctly identify the sequence to the species level. In fact, it scored a value higher than 98% with 4 different species. This occurrence was already highlighted by Cohen *et al*., (2009) (see Section 3.5). The same outcome was obtained for the species *Gadus morhua* and *Gadus chalcogrammus* (valid name according to Fishbase *Theragra chalcogramma*) and for the species of the genus *Decapterus* (Carangidae) (Table 1SM).

When the top match did not reach a value equal or higher than 98%, the database resulted in "no match", with the exception of the sequences obtained from the sample labeled as *Ompok bimaculatus*, which scored an identity value of 97.75% with *Ompok pabda* (Table 1SM). In some cases (ETN7, ETN10 and ETN53), a top match below 98% was due to the absence of deposited reference sequences of the species declared in the label, namely *Corica soborna*, *Neotropius acutirostris*, and *Otolithoides pama*, preventing specific and also genus level identification. However, in 4 other cases (ETN22, ETN33, ETN34 and ETN54.10), even though the reference sequences for the declared species were not present on the database an unambiguous match was obtained with another species, enabling specific identification. On the contrary, for other products, a high scored match was obtained with species different from those declared on the label, despite the latter being included in the database. All these samples were thus considered mislabeled.

Considering separately the results obtained with the IDs analysis for the FDB and the MDB, we observed that 129 FDB out of 158 (82%) and 43 out of 46 MDB sequences (94%) obtained a maximum species identity in the range of 98–100%. Among these, 81 FDB (51.3%) and 34 MDB (74%) allowed an unambiguous identification at the species level on the BOLD system. These

309 unexpected results confirm the potential discriminatory power enclosed in the MDB (Armani *et al*.,
310 2015) and could be related to an higher variability of this region.

311 When analyzed by BLAST a maximum species identity in the range of 98–100% was obtained
312 for 151 sequences (74%). Of these 110 (73%) were unambiguously identified at species level, while
313 the remaining 41 sequences were not identified due to the close phylogenetic relationship between
314 species. The occurrence of the high matching with more than one species, already observed during
315 IDs analysis for Engraulidae, Tetraodontidae, Gadidae, and Carangidae, also recurred for NCBI
316 database (Table 1SM). As already mentioned for BOLD, a higher performance was obtained in
317 identifying mollusks species rather than fish species (Table 1SM). When the top match did not
318 reach a value equal or higher than 98%, an identity value between 83 and 96% was observed (Table
319 1SM). As reported for BOLD, samples ETN7, ETN10 and ETN53 were not identified, due to the
320 absence of reference sequences. Also on this database, even though the COI sequences of the
321 species declared on the label of ETN22, ETN33, ETN34, and ETN54.10 products were not
322 available, the BLAST analysis allowed unambiguous identification with other species. Finally,
323 matching with species other than those declared on the label was obtained for the same products to
324 which this occurrence was observed by performing analysis using BOLD database.

325 Considering separately the results obtained by the BLAST analysis for the FDB and the MDB,
326 we observed that 105 FDB (66%) reached a maximum species identity equal or higher than 98%
327 and, among these, 82 (52%) could be unequivocally attributed to a definite species. Concerning
328 MDB, all 46 MDB sequences (100%) obtained a maximum species identity in the range of 98–
329 100%. Among these, 28 MDB (61%) allowed an unambiguous identification at the species level on
330 the NCBI system.

331 On both databases, an intraspecific variability higher than 2%, a threshold considered effective in
332 distinguish different species (Hebert *et al*., 2003a), was found (highlighted in gray in Table 1SM
333 and described in details in Table 3SM). On BOLD, this issue was observed for 9 species from 12
334 products. In particular, while for 2 species the divergence was only slightly higher (2.56-2.88%),

the *COI* genetic divergence within some species was much wider than those previously described (Ward & Holmes, 2007), reaching values of 14.43-15.67%. In the case of freshwater species, such as *Anabas testidineus*, this divergence may be explained by the effect of highly restricted gene flow attributable to the fragmented nature of freshwater ecosystems (April, Mayden, Hanner, & Bernatchez, 2011). In other cases, high values of intraspecific distance can be attributed to cryptic or undescribed species (Ward, Holmes, & Yearsley, 2008; April *et al*., 2011). However, in the remaining cases, the observed distance could represent an indication of misidentification, as already reported for other species (Landi *et al*., 2014; Ardura *et al*., 2013). The BLAST analysis also gave some ambiguous results, although a lower number of species were affected by this problem with respect to NCBI database (3 instead of 8) (Table 2SM).

The comparison among the identification results obtained on BOLD and GenBank shows that the discrimination power of these two databases was different according to the kind of barcode analyzed. While on both databases the FDB was able to identify ~52% of the samples, a difference was observed for the MDB. In particular, MDB performed better on BOLD than on GenBank (74% vs 61%). In this study, the higher resolution of BOLD with respect to GenBank in correctly identifying FDB, already found by other authors (Wong & Hanner, 2008; Cawthorn *et al*., 2012), was also confirmed for the MDB.

3.3 Label analysis

The presence and the correctness of the mandatory information, such as the commercial (trade) name, the scientific name, the production method (aquaculture or fishery product) and the catch area were verified on the label. Information on traceability was examined according to Council Regulation (EC) No 104/2000, which is applicable to fresh, chilled, frozen, dried, salted, brined, and smoked products (Circular of the Italian Minister of Agriculture, Food and Forestry (MIPAAF) n. 21329 of 27th May 2002). In the case of canned products, only the correctness of label information was assessed.

First of all, the presence of a label in Italian, in English or in other languages was verified. In fact, in order to facilitate the understanding of the information the language used for label details must be intelligible by consumers (Regulation (EU) No 1169/2011). It was observed that an Italian label was present in 87% of the cases (59/68), while 4% of the products presented only an English label and the remaining 9% only a label in other languages (mainly Chinese) (Table 5).

Then, the correspondence between the trade and the scientific name was assessed, consulting both the official list of the Italian Ministry of Agricultural, Food and Forestry Policies (MIPAF, 2008) and the list of provisional denominations proposed by the Veneto Region (http://www.izsvenezie.it/documenti/temi/identificazione-specie-ittiche/catalogo-specie-ittiche/denominazione-prodotti-pesca-veneto.pdf). Considering that one of the products contained a mix of 8 species, the total number of scientific and trade names (75) is higher than the total number of collected products (68). On the labels, the trade and the scientific name corresponded in 40% of the cases (30/75); did not corresponded in 29% of the cases (22/75); and one or both denominations were not present in the remaining 31% (23/75) (Table 5). Overall, we found that 6 trade names were still not present in the official list (of these, 4 were not even included among the provisional denominations). Regarding the scientific name, 12 species were absent from the official list (of these, 7 were not even included among the provisional denominations). In particular, for three products both the trade and the scientific name (*Corvina macrocefala* - *Collichthys lucidus* and *Gamberetto cinese* - *Acetes chinensis*) were not included in any lists.

The assessment of the correspondence between scientific and commercial denomination on the basis of the Italian list could affect the objectivity of results, providing data not comparable with those obtained in other studies/countries. Therefore, the utilization of an official international accepted list could allow to normalize the mislabeling rate, allowing comparison among different countries. In fact, also the Regulation (EU) 1379/2013 suggests to the Member States to draw up a list of commercial/scientific designations on the basis of the FAO list (FAO, 2014bhttp://www.fao.org/fishery/collection/asfis/en), with the aim to reduce the discrepancies among

commercial denominations used in the EU territory. However, the lack of Italian commercial denomination in the aforesaid list did not allow us to proceed in the comparison.

Of the 68 products collected, 25 (37%) did not report the catch area, not fulfilling European requirements. The other products came from FAO areas 61 (n=32), 04 (8), 57 (1), 71 (1) and 87 (1) (Table 5). Despite a previous survey found that only a few species routinely commercialized on the Italian market originated from FAO area 61 (NW, Pacific), the high rate of ethnic products coming from this area should be noted. Fish products are of particular concern for their capability of bioaccumulation and those coming from FAO area 61 are at risk for the presence of radioactive contaminations, due to the spilling of contaminated waters in the Chinese Sea after the nuclear accident of March 2011 in Fukushima, Japan. In fact, even though after that accident the EU has imposed additional tests for the products imported from Japan, these restrictive measures do not apply to products imported from other countries, such as China, although originating from the same area (Armani *et al.*, 2012b). Overall, 66% of the products presented labeling non-conformities.

Considering the high rate of label non-conformities found (79% in Chinese products and 54% in Bangladeshi products) we also decided to verify the origin of the products. In fact, fishery and aquaculture products can only be imported into the EU if they come from an approved establishment of a third Country included in a positive list of eligible Countries for the relevant product (Regulation (EC) 854/2004). Therefore, we verified the identification numbers of the factories reported on the labels by consulting the available official lists reporting the companies approved for importation: China (https://webgate.ec.europa.eu/sanco/traces/output/CN/FFP_CN_en.pdf; Bangladesh: https://webgate.ec.europa.eu/sanco/traces/output/BD/FFP_BD_en.pdf) Indonesia (https://webgate.ec.europa.eu/sanco/traces/output/ID/FFP_ID_en.pdf;) Myanmar (http://fishexporters.org/index.php?option=com_content&view=article&id=19&Itemid=122;) and Vietnam (https://webgate.ec.europa.eu/sanco/traces/output/VN/FFP_VN_en.pdf). On the basis of the verifiable information, all the products purchased in Chinese shops originated from China

(mainly from Zehjian Province), while the 11 products bought in Bangladeshi shops were variously distributed: 6 were produced in Bangladesh, 3 in Myanmar, 1 in Indonesia and 1 in Vietnam.

Interestingly, while all the Bangladeshi products came from approved processing plants, 23 Chinese products reported a plant not included in the official lists. This evidence suggests the hypothesis of possible "paralegal" imports of Chinese products, alongside legal imports (Armani, Castigliego, Gianfaldoni, & Guidi, 2011; Pramod, Nakamura, Pitcher, & Delagran, 2014) and it is further supported by the fact that both products (ETN 39 and 43) containing toxic puffer fish come from "unapproved" establishments. Pramod *et al*. (2014) found that Chinese seafood exported to US largely belong to illegal and unreported sourced fish and that supply chains for seafood products transiting China are rife with opportunities for obfuscation and the laundering of illegal catches into legitimate trade flows. Moreover, another issue associated to seafood products sold inside the ethnic retails is related to the so-called personal importation. In fact, importation on fish, bivalves (dead) and fish/fishery products (no more than 20kg total weight (fish must be gutted if fresh, or processed, i.e. cooked, cured, dried or smoked), 2kg total weight combined per person) are allowed from non-EU countries (Commission Regulation (EC) No 206/2009) Thus, food business operators that manage an ethnic retail shop can put on the market these kind of products. In fact, 91% of the Chinese products coming from unapproved plants showed the QS (Quality and Safety) label (Figure 1), which is required for selling products within the territory of the People's Republic of China (http://www.asianlii.org/cn/legis/cen/laws/irftsaaotqsotfmapet1398/), instead of the CIQ (China Inspection Quarantine), which should be reported on the sale packaging of Chinese food products certified for export (http://search.mofcom.gov.cn/swb/recordShow.jsp?flag=0&lang=1&base=iflow_4&id=english2007 09050919481&value=%28Announcement%20and%2085%20and%202007%29). The remaining 2 products coming from unapproved plants did not show any label.

3.4 Comparison between molecular and label analysis

437 The correspondence between the information reported on the label (commercial and scientific
438 denomination) and the results of the molecular analysis were assessed. For some products, even
439 though it was not possible to achieve unequivocal identification, it was still possible to detect the
440 presence of frauds, due to the fact that the species declared on the label was not confirmed by the
441 molecular analysis (see Section 3.2.3). Considering these frauds altogether, 33 products out of 68
442 (48.5%) were mislabeled. In particular, these were 5 out of the 11 Bangladeshi products (45%) and
443 28 out of the 57 Chinese products (49%) (Table 1SM; Table 5).
444 Furthermore, 3 products among the mislabeled contained species belonging to the genus
445 *Carcharhinus* sp., considered by the International Union for the Conservation of Nature (IUCN) as
446 near threatened (Table 3SM).
447 Finally, the results of the molecular analysis of the products that were specifically identified
448 were compared with the catch area reported on the label: 3% showed a discrepancy between the
449 catch area and the geographical habitat of the species identified by barcoding. Moreover, an
450 analysis was carried out on the habitat of the species unequivocally identified by molecular analysis
451 (Table 4SM): on the whole, 19 species where marine or marine-brackish and 10 species lived in
452 freshwater or fresh-brackish water. Interestingly, a difference could be observed between products
453 purchased in Chinese shops and in Bangladeshi shops: in fact, while species contained in Chinese
454 products were mostly marine (17 VS 1 from freshwater), species identified in Bangladeshi products
455 were mainly from freshwater habitats (9 VS 2 from marine water). This difference seems to support
456 the survey of Hu (2014), who showed that Chinese consumers have a traditional food culture on
457 freshwater fish, but the consumption patterns are progressively changing according to the life style
458 changes and economic growth. All the observed non conformities were summarized in Table 5SM.
459 ***3.5 Health implications***
460 Health implications were highlighted for 2 samples labeled as squid but identified at the genus
461 level as puffer fish, *Lagocephalus* spp. In fact, the molecular analysis failed in identifying these
462 samples at the species level giving similar identity values for *Lagocephalus spadiceus, L. inermis*,

463 *L. gloveri* and *L. wheeleri* (Tetraodontidae family). This issue was already found by Cohen *et al*.,
464 (2009) and was attributed to a lack of authenticated standard on the database. However, this
465 impossibility to make a specific diagnosis is not a limit for the goal of our study, considering that,
466 according to the current European legislative requirements (Regulation (EC) 853/2004; Regulation
467 (EC) 854/2004), all the fish belonging to the family Tetraodontidae must not be placed on the
468 market. In fact, these species contain the tetrodotoxin (TTX), a heat stable neurotoxin that blocks
469 sodium conductance and neuronal transmission in skeletal muscle, and leads to weakness or
470 paralysis and potentially death if ingested in sufficient quantities (Mosher & Fuhrman, 1984). While
471 the muscle of most commercial Asian puffer species is non or weakly toxic, some species such as
472 *Lagocephalus lunaris*, *L. spadiceus*, and *L. inermis*, are known to contain TTX (Chulanetra *et al*.,
473 2011). Incidents related to poisonous puffer fishes have been frequently reported within Countries
474 of the Indo-West Pacific: Japan, China, Taiwan, Philippines, Thailand and Bangladesh (Hwang &
475 Noguchi, 2007).
476 In US, the legal importation of puffer fish is limited to a single Japanese importer certified by the
477 Japanese Ministry for Health and Welfare
478 (http://www.fda.gov/InternationalPrograms/Agreements/MemorandaofUnderstanding/ucm107601.h
479 tm). However, previous cases of TTX poisoning demonstrate that illegal importation of puffer fish
480 into the United States continues in response to consumer demand (Centers for Disease Control and
481 Prevention, 1996; Coehn *et al*., 2009). In particular, in 2007, 2 people became ill after consuming
482 puffer fish imported from China (Cohen *et al*.,
483 2009http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/2007/ucm108920.htm).
484 In Italy, the first and only case of death due to ingestion of these poisonous species had been
485 recorded in 1997, when 3 people consumed frozen fillets of "monkfish" that had been fraudulently
486 replaced with "fish tail ball" (probably *Lagocephalus lunaris*) imported from Taiwan. Other
487 attempts to commercialize these species have been reported in many Italian regions during the
488 subsequent years (Pucci, 2014). The risk associated to the importation of these toxic species might

489 increase, considering that, in recent years, farmed pufferfish has become more and more popular
490 (Tao, Wang, Gong, & Liu, 2012).

491 Sale of the toxic fish for home or commercial preparation puts consumers, who may not be
492 aware of the illegality of the sale, at risk for tetrodotoxin poisoning. In fact, even though such
493 labeling incongruence is sometime due to mere negligence, it can also be due to an aforethought
494 modification of the reported information with the aim of circumventing the normative related to the
495 importation. These ethnic food stores and restaurants are nowadays frequented by consumers
496 belonging to many different nationalities, included Italian people, due to a change in food tastes and
497 to the lower prices. In fact, also as a consequence of the recent economic recession, the
498 consumption of ethnic products is augmented so much that some of these products are appearing in
499 many district markets. In US, the Mintel Group estimates that, between 2012 and 2017, sales of
500 ethnic foods in grocery stores will grow more than 20 percent
501 (http://reports.mintel.com/display/590141/#). European ethnic food sales were worth around EUR
502 4.12 billions in 2006 with Chinese and Oriental foods accounting for 42% of value sales
503 (http://www.thinkethnic.com/portfolio/european-ethnic-food-growth/). In UK, the Europe largest
504 market ahead of France, Germany and the Netherlands, the overall ethnic foods market recorded a
505 24% increase in value sales between 2007 and 2011 (http://store.mintel.com/ethnic-foods-uk-
506 september-2012).

507 The severity of the illness associated with these species represents a public health and a safety
508 issue that can be prevented by official control. In this study, Local Health Authorities were
509 informed immediately after the output of the molecular analysis with the aim to proceed to a rapid
510 seizure of this dangerous food product from ethnic markets. In fact, other than from the European
511 Regulations (Regulation (EC) No 853/2004; Regulation (CE) No 854/2004) the offering of sale of
512 toxic species represent a healthfraud in the light of the Italian Penal Code (art. 442 and 444)

513 **4. Conclusions**

In this work, the DNA barcoding approach was used to assess the label information of seafood ethnic products collected in Chinese and Bangladeshi shops. Our study confirmed the reliability of both FDB and MDB in fish identification even in the case of processed products submitted to different technological treatments. Overall, this approach highlighted a high rate of incorrect labeling (48.5%), which, in 2 cases, was associated with healthissues due to the presence of toxic fish species belonging to the Tetraodontidae family. Our survey has further demonstrated as the ethnic communities are still characterized by law disregard and tend to maintain "paralegal" commercial circuits that often operate outside the European rules framework on food safety. Finally, this work confirmed that the molecular inspection of seafood should be routinely used as support for the official control for ensuring the regulatory compliance.

Acknowledgements

The authors would like to thank the anonymous referees for their suggestions which helped improving the manuscript.

Funding source

This project has been found by the Tuscany Region Project titled "*Analisi delle problematiche igienico-sanitarie e commerciali legate alla vendita e somministrazione dei prodotti alimentari etnici all'interno della Regione Toscana*".

References

April, J., Mayden, R. L., Hanner, R. H., & Bernatchez, L. (2011). Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Sciences, 108*(26), 10602-10607.

Ardura, A., Planes, S., & Garcia-Vazquez, E. (2013). Applications of DNA barcoding to fish landings: authentication and diversity assessment. *ZooKeys, 365*, 49.

Armani, A., Guardone, L., Castigliego, L., D'Amico, P., Messina, A., Malandra, R., Gianfaldoni, G., & Guidi, A. (2015). DNA and Mini-DNA Barcoding for the identification of Porgies species (family Sparidae) of commercial interest on the international market. *Food Control, 50*, 589-596.

542 Armani, A., Tinacci, L., Xiong, X., Titarenko, E., Guidi, A., & Castigliego, L. (2014). Development
543 of a simple and cost-effective bead-milling method for DNA extraction from fish muscles.
544 *Food Analytical Methods, 7*, 946–955.
545 Armani, A., Tinacci, L., Giusti, A., Castigliego, L., Gianfaldoni, D., & Guidi, A. (2013). What is
546 inside the jar? Forensically informative nucleotide sequencing (FINS) of a short
547 mitochondrial *COI* gene fragment reveals a high percentage of mislabeling in jellyfish food
548 products. *Food Research International, 54*(2), 1383-1393.
549 Armani, A., Castigliego, L., & Guidi, A. (2012). Fish fraud: The DNA challenge. *CAB Animal*
550 *Science Reviews*, 7, 227-239.
551 Armani, A., D'Amico, P., Castigliego, L., Betti, B., Gianfaldoni, D., Guidi, A., & Gandini, G.
552 (2012)b. I prodotti ittici dopo Fukushima: le garanzie di sicurezza per il consumatore
553 comunitario. *Industrie alimentari, 527,* 19-27.
554 Armani, A., Castigliego, L., Tinacci, L., Gianfaldoni, D., & Guidi, A. (2011) Molecular
555 characterization of icefish (Salangidae family), using direct sequencing of mitochondrial
556 cytochrome b gene. *Food Control, 22*, 888–895.
557 Barbuto, M., Galimberti, A., Ferri, E., Labra, M., Malandra, R., Galli, P., & Casiraghi, M. (2010).
558 DNA barcoding reveals fraudulent substitutions in shark seafood products: the Italian case of
559 ''palombo" (*Mustelus* spp.). *Food Research International, 43*, 376–381.
560 Carrera, E., Garcia, T., Cespedes, A., Gonzalez, I., Fernandez, A., Asensio, L.M., Hernandez, P.E., &
561 Martin, R. (2000). Identification of smoked Atlantic salmon (*Salmo salar*) and rainbow trout
562 (*Oncorhynchus mykiss*) using PCR-restriction fragment length polymorphism of the p53 gene.
563 *Journal of AOAC International, 83*(2), 341–346.
564 Carvalho, D. C., Palhares, R. M., Drummond, M. G., & Frigo, T. B. (2015). DNA Barcoding
565 identification of commercialized seafood in South Brazil: A governmental regulatory forensic
566 program. *Food Control*, *50*, 784-788.
567 Cawthorn, D.M., Steinman, H.A., & Witthuhn, R.C. (2012). DNA barcoding reveals a high incidence
568 of fish species misrepresentation and substitution on the South African market. *Food*
569 *Research International, 46*, 30–40.
570 Centers for Disease Control and Prevention (1996). Tetrodotoxin poisoning associated with eating
571 puffer fish transported from Japan—California, 1996. *Morbibity and Mortality Weekly*
572 *Report, 45*, 389–391.
573 Circular of the Italian Minister of Agriculture, Food and Forestry (MIPAAF) n. 21329 of May 27,
574 2002-Subject: Reg. 2065/2001 of the Commission of 22 October 2001 of 22 October 2001
575 laying down detailed rules for the implementation of Regulation (EC) 104/2000, as regards

576 information to consumers in the field of fisheries and aquaculture. Ministerial Decree of 27
577 March 2002
578 Civera, T. (2003). Species identification and safety of fish products. *Veterinary Research*
579 *Communications, 27*, 481–489.
580 Cohen, N.J., Deeds, J.R., Wong, E.S., Hanner, R.H., Yancy, H.F., White, K.D., Thompson,
581 T.M., Wahl, M., Pham, T.D., Guichard, F.M., Huh, I., Austin, C., Dizikes, G., & Gerber, S. I.
582 (2009). Public health response to puffer fish (tetrodotoxin) poisoning from mislabeled
583 product. *Journal of Food Protection, 72*(4), 810-817.
584 Commission Regulation (EC) No 206/2009 of 5 March 2009 on the introduction into the Community
585 of personal consignments of products of animal origin and amending Regulation (EC) No
586 136/2004. *Official Journal of the European Union, L77.*
587 Council Regulation (EC) No 104/2000 of 17 December 1999 on the common organization of the
588 markets in the fishery and aquaculture products. *Official Journal of the European*
589 *Communities, L17.*
590 Chulanetra, M., Sookrung, N., Srimanote, P., Indrawattana, N., Thanongsaksrikul, J., Sakolvaree, Y.,
591 Chongsa-Nguan, M., Kurazono, H., & Chaicumpa, W. (2011). Toxic Marine Puffer Fish in
592 Thailand Seas and Tetrodotoxin They Contained. *Toxins, 3*, 1249-1262.
593 Cutarelli, A., Amoroso, M. G., De Roma, A., Girardi, S., Galiero, G., Guarino, A., & Corrado, F.
594 (2014). Italian market fish species identification and commercial frauds revealing by DNA
595 sequencing. *Food Control, 37*, 46-50.
596 D'Amico, P., Armani, A., Castigliego, L., Sheng, G., Gianfaldoni, D., & Guidi, A. (2014). Seafood
597 traceability issues in Chinese food business activities in the light of the European
598 provisions. *Food Control, 35*(1), 7-13.
599 Ercsey-Ravasz, M., Toroczkai, Z., Lakner, Z., & Baranyi, J. (2012). Complexity of the international
600 agro-food trade network and its impact on food safety. *PloS one, 7*(5), e37810.
601 EUMOFA, European Market Observatory for Fisheries and Aquaculture Products (2014). The EU
602 fish market: 2014 edition. Available at http://ec.europa.eu/fisheries/market-observatory
603 Accessed 21/11/2014.
604 FAO (2014a). The state of world fisheries and aquaculture 2014: opportunities and challenges.
605 Available at http://www.fao.org/3/a-i3720e.pdf Accessed 21/11/2014.
606 FAO (2014b). ASFIS List of Species for Fishery Statistics Purposes. Available at
607 http://www.fao.org/fishery/collection/asfis/en.

608 Fondazione Leone Moressa (2010). L'influsso della cucina etnica sulle abitudini alimentari degli
609 italiani. Available at http://www.fondazioneleonemoressa.org/newsite/wp-
610 content/uploads/2010/10/comunicato_54.pdf Accessed 21/11/2014
611 Galal-Khallaf, A., Ardura, A., Mohammed-Geba, K., Borrell, Y. J., & Garcia-Vazquez, E. (2014).
612 DNA barcoding reveals a high level of mislabeling in Egyptian fish fillets. *Food Control*, *46*,
613 441-445.
614 Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis
615 program for Windows 95/98/NT. In *Nucleic acids symposium series* (Vol. 41, pp. 95-98).
616 Handy, S.M., Deeds, J.R., Ivanova, N.V., Hebert, P.D.N., Hanner, R.H., Ormos, A., Weigt, L.A.,
617 Moore, M.M., &Yancy, H.F. (2011). A single-laboratory validated method for the generation
618 of DNA barcodes for the identification of fish for regulatory compliance. *Journal of*
619 *Association of Official Analytical Chemists International*, *94*, 201-210.
620 Haye, P. A., Segovia, N. I., Vera, R., Gallardo, M. D. L. Á., & Gallardo-Escárate, C. (2012).
621 Authentication of commercialized crab-meat in Chile using DNA barcoding. *Food*
622 *Control*, *25*(1), 239-244.
623 Hebert, P.D.N., Ratnasingham, S., & de Waard, J.R. (2003)a. Barcoding animal life: cytochrome c
624 oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal*
625 *Society of London. Series B*, *270*, 596–599.
626 Hebert, P.D., Cywinska, A., & Ball, S.L. (2003)b. Biological identifications through DNA
627 barcodes. *Proceedings of the Royal Society of London. Series B: Biological*
628 *Sciences*, *270*(1512), 313-321.
629 Holmes, B.H., Steinke, D., & Ward, R.D. (2009). Identification of shark and ray fins using DNA
630 barcoding. *Fisheries Research*, *95*(2), 280-288.
631 Hwang, D.F., & Noguchi, T. (2007). Tetrodotoxin Poisoning. *Advances in Food and Nutrition*
632 *Research*, *52*, 141-236.
633 ISTAT, National Institute of Statistics (2011). La popolazione straniera presente in Italia. Available
634 at http://www.istat.it/it/archivio/39726 Accessed 21/11/2014
635 Lamendin, R., Miller, K., & Ward, R.D. (2015). Labeling accuracy in Tasmanian seafood: An
636 investigation using DNA barcoding. *Food Control*, *47*, 436-443.
637 Landi, M., Dimech, M., Arculeo, M., Biondo, G., Martins, R., Carneiro, M., Carvalho, G.R., Lo
638 Brutto, S., & Costa, F. O. (2014). DNA Barcoding for Species Assignment: The Case of
639 Mediterranean Marine Fishes. *PloS one*, *9*(9), e106135.
640 Lee, J. H., Hwang, J., & Mustapha, A. (2014). Popular Ethnic Foods in the United States:A Historical
641 and Safety Perspective. *Comprehensive Reviews in Food Science and Food Safety*,*13*, 2-17.

642 Mikkelsen, P.M., Bieler, R., Kappner, I., & Rawlings, T.A. (2006). Phylogeny of Veneroidea
643 (Mollusca: Bivalvia) based on morphology and molecules. *Zoological Journal of the Linnean*
644 *Society*,*148*(3):439–521.
645 MIPAF, 2008. Decreto del Ministero delle Politiche Agricole Alimentari e Forestali del 31 gennaio
646 2008 (G.U. n. 45 del 22 febbraio 2008) e successive integrazioni fino al decreto 19 novembre
647 2012 (G.U. n. 27 del 1 febbraio 2013). Available at
648 http://www.izsvenezie.it/documenti/temi/identificazione-specie-ittiche/catalogo-specie-
649 ittiche/denominazione-latina-specie-ittiche-di-interesse-commerciale.pdf.
650 M.L.P.S., Ministero del Lavoro e delle Politiche Sociali (2013) La Comunità Bengalese in Italia.
651 Rapporto annuale sulla presenza degli immigrati – 2013. Available at:
652 http://www.integrazionemigranti.gov.it/Attualita/IlPunto/Documents/2013%20-
653 %20Comunita%20Bengalese%20v_0.pdf Accessed 21/11/2014
654 Mosher, H. S., & Fuhrman, F. A. (1984). Occurrence and origin of tetrodotoxin. In: *Seafood Toxins*
655 (Ragelis EP, ed). Washington, DC: American Chemical Society, 333-344.
656 Pucci, M. (2014). Il Ruolo del Veterinario Ufficiale nel riconoscimento delle frodi nel comparto
657 ittico. Thesis of Master Degree, University of Pisa
658 Pramod, G., Nakamura, K., Pitcher, T.J., & Delagran, L. (2014). Estimates of illegal and unreported
659 fish in seafood imports to the USA. *Marine Policy, 48*, 102-113.
660 Regulation (EC) No 852/2004 of the European Parliament and of the Council of 29 April 2004 on the
661 hygiene of foodstuffs. *Official Journal of the European Union, L139.*
662 Regulation (EC) No 853/2004 of the European Parliament and of the Council of 29 April 2004 laying
663 down specific hygiene rules for on the hygiene of foodstuffs. *Official Journal of the European*
664 *Union, L139.*
665 Regulation (EC) No 854/2004 of the European Parliament and of the Council of 29 April 2004 laying
666 down specific rules for the organisation of official controls on products of animal origin
667 intended for human consumption. *Official Journal of the European Union, L 139.*
668 Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011.
669 (22th November 2011). On the provision of food information to consumers, amending
670 Regulations (EC) No 1924/2006 and (EC) No 1925/2006 of the European Parliament and of
671 the Council, and repealing Commission Directive 87/250/EEC, Council Directive
672 90/496/EEC, Commission Directive 1999/10/EC, Directive 2000/13/EC of the European
673 Parliament and of the Council, Commission Directives 2002/67/EC and 2008/5/EC and
674 Commission Regulation (EC) No 608/2004. *Official Journal of the European Union, L*
675 *304/18.*

676 Regulation (EU) No 1379/2013 of the European Parliament and of the Council of 11 December 2013
677 on the common organisation of the markets in fishery and aquaculture products, amending
678 Council Regulations (EC) No 1184/2006 and (EC) No 1224/2009 and repealing Council
679 Regulation (EC) No 104/2000. *Official Journal of the European Union, L 354.*
680 Spink, J., & Moyer, D. C. (2011). Defining the public health threat of food fraud. *Journal of food*
681 *science*, *76*(9), 157-163.
682 Steffens, D. L., Sutter, S. L., & Roemer, S. C. (1993). An alternate universal forward primer for
683 improved automated DNA sequencing of M13. *Biotechniques*, 15, 580-582.
684 Stiles, M.L., Lahr, H., Lahey, W., Shaftel, E., Bethel, D., Falls, J., & Hirshfield, M. S., (2011).
685 Oceana, Bait and switch: how seafood fraud hurts our Oceans, our wallets and our health.
686 http://oceana.org/sites/default/files/reports/Bait_and_Switch_report_2011.pdf Accessed
687 21/11/2014
688 Tao, N.P., Wang, L.Y., Gong, X., & Liu, Y. (2012). Comparison of nutritional composition of
689 farmed pufferfish muscles among *Fugu obscurus*, *Fugu flavidus* and *Fugu rubripes*. *Journal*
690 *of Food Composition and Analysis*, *28*(1), 40-45.
691 Tarquini, A. (2010). Le rotte migratorie Bangladesh-Italia Available at
692 http://www.repubblica.it/solidarieta/cooperazione/2010/11/16/news/le_rotte_migratorie_bangl
693 adesh-italia_una_guida_per_una_partenza_consapevole-9186035/ Accessed 21/11/2014
694 Teletchea, F. (2009). Molecular identification methods of fish species: reassessment and possible
695 applications. *Reviews in Fish Biology and Fisheries, 19*(3), 265-293.
696 Todarello, M. (2013). Prato e le Chinatown italiane. Lettera 43. Available at
697 http://www.lettera43.it/economia/aziende/prato-e-le-chinatown-italiane_43675114978.htm
698 Accessed 21/11/2014
699 Ward, R.D. & Holmes, B.H. (2007). An analysis of nucleotide and amino acid variability in the
700 barcode region of cytochrome c oxidase I (cox1) in fishes. *Molecular Ecology Notes, 7*, 899–
701 907.
702 Ward, R.D., Holmes, B.H., & Yearsley, G.K. (2008). DNA barcoding reveals a likely second species
703 of Asian sea bass (barramundi)(*Lates calcarifer*). *Journal of Fish Biology*, *72*(2), 458-463.
704 Wong, E.H.K., & Hanner, R.H. (2008). DNA barcoding detects market substitution in North
705 American seafood. *Food Research International*, *41*, 828–837.
706 Zhang, D.X., Hewitt G.M. (1996). Nuclear integrations: challenges for mitochondrial DNAmarkers.
707 *Trends in Ecology & Evolution*, *6*, 247-251.
708
709 Figure caption

Figure 1

QS label The “Quality and safety”, QS label, which is required for products sold within the territory of the People's Republic of China.

Figure

Click here to download high resolution image

Table

References	DNA barcoding	Total samples	Unprocessed		Processed		Sequences length
			N	% sequencing success	N	% sequencing success	
This study	Full barcode	68	19	79	49	67	~655 bp
	Mini barcode			95		92	139 bp
Carvalho *et al*., 2015	Full barcode	30	17	100	13	100	≥520 - 655 bp[a]
Cutarelli *et al*., 2014	Full barcode	58	40	100	18[b]	100	~655 bp
Galal-Khallaf *et al*., 2014	Full barcode	90	90	100	--	--	604-625 bp
Lamendin *et al*., 2014	Full barcode	51	51	74.5	--	--	≥530 - 655 bp
Cawthorn *et al*., 2012	Full barcode	257	248	100	9	0	~655 bp
Haye *et al*., 2012	Full barcode	333	275	39.3	58	10.3	~655 bp
Holmes *et al*., 2009	Full barcode	211	--	--	211	91.5	≥398 - 655 bp
Wong & Hanner, 2008	Full barcode	96	92	97.9	4	50	~655 bp

Table 1 Comparison between the present study and other studies that applied DNA barcoding for the identification of unprocessed and processed seafood products. [a]The average length was 643bp for unprocessed samples and 555bp for processed products; [b]frozen samples have been included in this category by the authors. N = number.

Table

Product type	N of products sampled	Product origin		N of products sequenced	Tot seq obtained	N FDB	Mean lenght FDB	N MDB	Mean lenght MDB
		C	B						
Whole	31	21	10	30 (97%)	106	95 (89.6%)	637.2	11 (10.4%)	139
Prepared	37	36	1	33 (89%)	98	65 (66.3%)	633	33 (33.7%)	139
Product treatment									
Only frozen (unprocessed)	**19**	**8**	**11**	**18 (94.7%)**	**64**	**53 (82.8%)**	**648.5**	**11 (17.2%)**	**139**
Only dried	2	2	0	2	6	6 (100%)	658	0	-
Frozen and dried	16	16	0	16	53	50 (94.3%)	620.1	3 (5.7%)	139
Frozen and salted	3	3	0	2	6	6 (100%)	642	0	-
Frozen, dried and salted	3	3	0	3	10	10 (100%)	656.2	0	-
Dried and roasted	18	18	0	18	54	35 (64.8%)	629.7	19 (35.2%)	139
Dried and smoked	1	1	0	0	0	0	-	0	-
Canned	6	6	0	4	11	0	-	11 (100%)	139
Processed[a]	**49**	**49**	**0**	**45 (93.7%)**	**140**	**107 (76.4%)**	**640.3**	**33 (23.6%)**	**139**

Table 2 Product type and treatment in relation to their origin and sequencing results. N = number; C = products purchased in Chinese shops; B= products purchased in Bangladeshi shops. [a]includes only dried, frozen and dried, frozen and salted, frozen, dried and salted, dried and roasted, dried and smoked and canned products.

e-component

Click here to download e-component: Table 3.doc

Table

Primer pair	Primer name	Sequence code	Position	Amplicon lenght (bp)	Annealing temperature	Reference
PP1	FISHCOILBC_ts	CTCAACYAATCAYAAAGATATYGGCAC	4418-4444	705	55°C	Handy *et al.*, 2011
	FISHCOIHBC_ts	ACTTCYGGGTGRCCRAARAATCA	5100-5123			
PP2	COIF-ALT	ACAAATCAYAARGAYATYGG	4422-4441	698	47°C	Mikkelsen *et al*., 2006
	COIR-ALT	TTCAGGRTGNCCRAARAAYCA	5100-5120			
PP3	REVshort1	GGYATNACTATRAAGAAAATTATTAC	4584-4610	192[a]	51°C	Armani *et al*., 2015
Tails	M13F	CACGACGTTGTAAAACGAC	--	--	--	Steffens *et al*., 1993
	M13R	GGATAACAATTTCACACAGG	--			

Table 4 Universal primers for the amplification of the *COI* gene from fish used in this study. The primers' position has been calculated on the sequence of *Pagrus auriga*, (GenBank Accession Number AB124801). [a]The length refers to the amplicon generated using the forward FISHCOILBC_ts.

Table

Labels	- Italian - English - Other languages	59 (87%) 3 (4%) 6 (9%)
Trade name[a]	- corresponding to scientific name - non corresponding to scientific name - not verifiable (absence of trade and/ or scientific name; trade name and or scientific name not included in Ministerial lists)	30 (40%) 22 (29%) 23 (31%)
Catch area	- not reported - FAO61 - FAO04 - FAO57 - FAO71 - FAO87	25 (37%) 32 (47%) 8 (11.5%) 1 (1.5%) 1 (1.5%) 1 (1.5%)
Label information and molecular identification	- correspondence between label and molecular identification at species level - non correspondence between label and molecular identification - not verifiable i) maximum identity <98% ii) only genus level molecular identification iii) only family level molecular identification iv) absence of scientific name/non readable sequences	15 (22%) 33 (48.5%) 5 (7.3%) 1 (1.5%) 7 (10.3%) 7 (10.3%)

Table 5 Label information and comparison between labels and molecular results. [a]One of the products contained a mix of 8 species, therefore the total number of scientific and trade names (75) is higher than the total number of collected products (68).

Supplementary material for online publication only

Code	Species declared in the label	Seq. length (bp)	BOLD ID System	BLAST NCBI (Max id.)
ETN1.1	*Engraulis japonicus*	655	*Encrasicholina punctifer* 99.85% (1seq) *Engraulis australis* 99.69 (1 seq) - 98% *Engraulis japonicus* 99.23-98.77%	*Engraulis japonicus* 99-98% *Engraulis australis* 98% (2 seq) *Engraulis encrasicolus* 98% *Engraulis eurystole* 98% (1 seq)
ETN1.2		655	*Engraulis japonicus* 99.54-98.77% *Encrasicholina punctifer* 98.92% (1seq) *Engraulis australis* 98.77-98.%	*Engraulis japonicus* 99% *Engraulis australis* 98% (2 seq) *Engraulis encrasicolus* 98% *Engraulis eurystole* 98% (1 seq)
ETN1.3		655	*Engraulis japonicus* 99.54-99.08% *Encrasicholina punctifer* 99.23% (1seq) *Engraulis australis* 99.08-98.31%	*Engraulis japonicus* 99-97% *Engraulis australis* 98% (2 seq) *Engraulis encrasicolus* 98% *Engraulis eurystole* 98% (1 seq)
ETN1.4		590	*Engraulis japonicus* 99.26-98.07% *Encrasicholina punctifer* 98.88% (1seq) *Engraulis australis* 98.7-98.14%	*Engraulis japonicus* 99-98% *Engraulis australis* 98% (2 seq) *Engraulis encrasicolus* 98% *Engraulis eurystole* 98% (1 seq)
ETN2.1 ETN2.2	*Coilia nasus*	587-655	*Stolephorus* cf. *waitei* 4 100-99.69% *Stolephorus waitei* 100-84.49%	*Stolephorus waitei* 85% *Stolephorus chinensis* 85% *Anchoviella* spp. 85%
ETN2.3 ETN2.4		655	*Stolephorus* cf. *waitei* 4 100-99.85% *Stolephorus waitei* 100-84.33%	*Engraulis encrasicolus* 84% *Engraulis japonicus* 84% *Anchoviella* spp. 84%
ETN3.1	*Engraulis japonicus*	650	*Encrasicholina punctifer* 100-97.22% *Stolephorus indicus* 99.85% (1 seq)	*Encrasicholina punctifer* 100%
ETN3.2		655	*Engraulis japonicus* 99.54-98.77% *Encrasicholina punctifer* 98.92% (1seq) *Engraulis australis* 98.77-98%	*Engraulis japonicus* 99-98% *Engraulis encrasicolus* 98% *Engraulis australis* 98% *Engraulis eurystole* 98%
ETN3.3		485	*Engraulis japonicus 100-99.17%* *Encrasicholina punctifer* 98.17% (1seq) *Engraulis australis 99.17-98.13%*	*Engraulis japonicus* 100-98% *Engraulis australis* 99% *Engraulis encrasicolus* 98% *Engraulis eurystole* 98%
ETN3.4		492	*Engraulis japonicus* 99.79-98.97% *Encrasicholina punctifer* 98.18% (1seq) *Engraulis australis* 98.97-98.15%	*Engraulis japonicus* 99-98% *Engraulis australis* 99-98% *Engraulis encrasicolus* 98% *Engraulis eurystole* 98%
ETN4.1	*Engraulis japonicus*	655	*Engraulis japonicus* 99.54-99.08% *Encrasicholina punctifer* 99.23% (1seq) *Engraulis australis* 99.08-98.31%	*Engraulis japonicus* 99% *Engraulis australis* 98% *Engraulis encrasicolus* 98%

				Engraulis eurystole 98% (1seq)
ETN4.2		655	*Engraulis japonicus* 99.08-98.31% *Encrasicholina punctifer* 98.46% (1seq) *Engraulis australis* 98.31-97.54%	*Engraulis japonicus* 99% *Engraulis australis* 98% *Engraulis encrasicolus* 98%
ETN4.3		655	*Engraulis japonicus* 99.39-98.92% *Encrasicholina punctifer* 99.08% (1seq) *Engraulis australis* 98.92-98.15%	*Engraulis japonicus* 99% *Engraulis australis* 98% *Engraulis encrasicolus* 98% *Engraulis eurystole* 98% (1seq)
ETN5.1 ETN5.2 ETN5.3	*Coilia mystus*	655	*Encrasicholina heteroloba* 100-85.27%	*Engraulis encrasicolus* 85% *Lycengraulis grossidens* 85%
ETN6.1 ETN6.2 ETN6.3	*Ompok bimaculatus*	655	*Ompok pabda* 97.75%	*Ompok bimaculatus* 95%
ETN7.1 ETN7.2	***Corica soborna*[a,b]**	655	No match	*Perca fluviatilis* 84%
ETN7.3		655	No match	*Perca fluviatilis* 83% *Scomberomorus sierra* 83% *Moxostomum papillosum* 83%
ETN8.1 ETN8.2 ETN8.3	*Puntius ticto* (v.n. *Pethia ticto*)	655	*Puntius chola* 99.84-99.54% *Puntius conchonius* 99.26% (1 seq)	*Puntius chola* 99% *Puntius conchonius* 99% (1 seq) *Puntius fraseri* 99% (1 seq)
ETN9.1 ETN9.2 ETN9.3	*Miichthys miiuy*	655	No match	*Atrobucca nibe* 92%
ETN10.1 ETN10.2	***Neotropius acutirostris*[a,b]**	655	No match	*Neotropius khavalchor* 90% *Neotropius atherinoides* 90% *Mystus tengara* 90%
ETN10.3		655	No match	*Neotropius atherinoides* 90% *Mystus tengara* 90%
ETN11	No sequences obtained			
ETN12.1 ETN12.2 ETN12.3	*Trichiurus haumela* (v. n. *Trichiurus lepturus*)	655	*Trichiurus japonicus* 99.85-99.52% (v. n. *Trichiurus lepturus*) *Trichiurus lepturus* 99.83-99.5%	*Trichiurus japonicus* 99% (v. n. *Trichiurus lepturus*) *Trichiurus lepturus* 99%
ETN13.1 ETN13.2	*Stromateoides argenteus* (v. n. *Pampus argenteus*)	655	No match	*Chaetodipterus faber* 86%
ETN13.4		655	No match	*Chaetodipterus faber* 85%
ETN14.1 ETN14.2 ETN14.3	*Pseudosciaena polyactis* (v. n. *Larimichthys polyactis*)	655	*Larimichthys polyactis* 100%	*Larimichthys polyactis* 99%

ETN15.1 ETN15.2 ETN15.3	*Carcharocles megalodon*	655	*Carcharhinus macloti* 100-99.54% *Rhizoprionodon taylori* 100% (1 seq) *Carcharhinus* spp. 99.54%(1 seq)	*Carcharhinus macloti* 100-99%
ETN16.1 ETN16.2	*Pseudosciaena crocea* (v.n. *Larimichthys crocea*)	655	*Larimichthys crocea* 100-99.83%	*Larimichthys crocea* 100-99%
ETN17.1 ETN17.2 ETN17.3	*Argyrosomus argentatus* (v.n. *Pennahia argentata*)	655	No match	*Atrobucca nibe* 92%
ETN18.1 ETN18.2 ETN18.3	*Argyrosomus argentatus* (v.n. *Pennahia argentata*)	655	No match	*Atrobucca nibe* 92%
ETN19.1 ETN19.2 ETN19.3	*Dodicus gigas*	658	*Dodicus gigas* 100%	*Dodicus gigas* 100%
ETN20.1 ETN20.2	*Ommastrephes bartramii*	658	*Ommastrephes bartramii* 100%	*Ommastrephes bartramii* 99%
ETN20.3		658	*Ommastrephes bartramii* 98.84-98.31%	*Ommastrephes bartramii* 99%
ETN21.1 ETN21.2 ETN21.3	*Dodicus gigas*	658	*Dodicus gigas* 100%	*Dodicus gigas* 100%
ETN22.1 ETN22.2 ETN22.3	***Lepidotrigla microptera***[a,b]	655	*Upeneus japonicus* 100-98.98% *Upeneus* cf. *asymmetricus* 99.69-99.53% (2seq) *Upeneus* cf. *japonicus* 99.2-98.09%	*Upeneus japonicus* 100%
ETN23.1 ETN23.2 ETN23.3	*Lophius litulon*	139	*Lophius litulon* 100-98.55%	*Lophius litulon* 100-99% *Grammistes sexilineatus* 99% (1seq)
ETN24	No sequences obtained			
ETN25	No sequences obtained			
ETN26	No sequences obtained			
ETN27	No sequences obtained			
ETN28.1 ETN28.2 ETN28.3	*Lophius litulon*	139	*Lophius litulon* 100%	*Grammistes sexilineatus* 100% (1seq) *Lophius litulon* 100-99%
ETN29.1 ETN29.2	*Lophius litulon*	139	*Lophius litulon* 99.28-98.55%	*Lophius litulon* 99%
ETN29.3		488	*Lophius litulon* 100-99.79%	*Lophius litulon* 100-99%
ETN30.1 ETN30.2 ETN30.3	*Lophius litulon*	139	*Lophius litulon* 100%	*Grammistes sexilineatus* 100% (1seq) *Lophius litulon* 100%
ETN31.1	*Lophius litulon*	139	*Lophius litulon* 100-98.55%	*Lophius litulon* 100-99%

ETN31.3				
ETN31.2 ETN31.4		553	*Lophius litulon* 100-99.64%	*Grammistes sexilineatus* 100% (1seq) *Lophius litulon* 100-99%
ETN32.1 ETN32.2	*Argyrosomus argentatus* (v.n. *Pennahia argentata*)	564	*Otolithes ruber* 99.82% *Pennahia macrophthalmus* 99.8-99.33%	*Atrobucca nibe* 86% *Cynoscion jamaicensis* 86% *Roncador steamsii* 86%
ETN33.1 ETN33.2 ETN33.3	***Lepidotrigla microptera***[a,b]	*655*	*Upeneus japonicus* 100-98.98% *Upeneus cf. asymetricus* 99.38% (2 seq) *Upeneus cf. japonicus* 99.2-99.06%	*Upeneus japonicus* 99%
ETN34.1 ETN34.2 ETN34.3	***Lepidotrigla microptera***[a,b]	586-655	*Upeneus japonicus* 100-98.92% *Upeneus cf. asymetricus* 99.69-99.65% (2 seq) *Upeneus cf. japonicus* 99.2-98.95%	*Upeneus japonicus* 100-99%
ETN35.1 ETN35.2 ETN35.3	*Collichthys lucidus*	596	*Pennahia macrocephalus* 98.65-98.45%	*Protonibea diacanthus* 89% *Pennahia argentata* 88% *Pennahia anea* 88%
ETN36.1	*Collichthys lucidus*	627	*Atrobucca nibe* 99.84-93.99%	*Atrobucca nibe* 95% *Argyrosomus inodorus* 88-87% *Umbrina canariensis* 87%
ETN36.2 ETN36.3		541	*Atrobucca nibe* 100-93.83%	*Atrobucca nibe* 95-94% *Argyrosomus regius* 88% *Argyrosomus inodorus* 88-87%
ETN37.1 ETN37.2	"*brandello del calamaro*"	618-655	*Upeneus japonicus* 99.67-98.64% *Upeneus cf. asymetricus* 99.53-99.34% (2 seq) *Upeneus cf. japonicus* 99.06-98.81%	*Upeneus japonicus* 99%
ETN38.1 ETN38.2 ETN38.3	*Lophius litulon*	655	*Upeneus japonicus* 100-99.77% *Upeneus cf. asymetricus* 99.77-99.54% (2 seq) *Upeneus cf. japonicus* 99.08-98.85%	*Upeneus japonicus* 100-99%
ETN39.1 ETN39.2 ETN39.3	"*brandello del calamaro*" (english label: marine fish fillet)	578-655	*Lagocephalus spadiceus* 100-93.51% *Lagocephalus inermis* 100-99.82% (2 seq) *Lagocephalus gloveri* 100% *Lagocephalus wheeleri* 100-99.3% *Lagocephalus cf spadiceus* 99.82% (1 seq)	*Lagocephalus spadiceus* 100-99% *Lagocephalus gloveri* 100% *Lagocephalus wheeleri* 100-99%
ETN40.1 ETN40.2 ETN40.3	"*calamaro*"	658	*Dosidicus gigas* 100-99.24%	*Dosidicus gigas* 100-99%
ETN41.1 ETN41.2 ETN41.3	fish fillet	139	*Gadus morhua* 100% *Gadus chalcogrammus* 100%	*Gadus morhua* 100% *Gadus chalcogrammus* 100% (v.n. *Theragra chalcogramma*)
ETN42.1 ETN42.2 ETN42.3	fish fillet	139	*Gadus morhua* 100% *Gadus chalcogrammus* 100% (v.n. *Theragra chalcogramma*)	*Gadus morhua* 100% *Gadus chalcogrammus* 100% (v.n. *Theragra chalcogramma*)

ETN43.1 ETN43.2 ETN43.3	"*brandello del calamaro*" (english label: marine fish fillet)	598-655	*Lagocephalus spadiceus* 100-93.93% *Lagocephalus inermis* 100-99.82% (2 seq) *Lagocephalus gloveri* 100% *Lagocephalus wheeleri* 100-99.38%* *Lagocephalus cf spadiceus* 99.66% (1 seq)	*Lagocephalus spadiceus* 100-99% *Lagocephalus gloveri* 100% *Lagocephalus wheeleri* 100-99%
ETN44.1 ETN44.2 ETN44.3	*Engraulis japonicus*	487-489	*Engraulis japonicus* 99.79-98.13% *Encrasicholina punctifer* 99.17-98.96% (1seq) *Engraulis australis* 98.96-98.13%	*Engraulis japonicus* 99% *Engraulis australis* 98% *Engraulis encrasicolus* 98% *Engraulis eurystole* 98%
ETN44.4		500	*Engraulis japonicus* 99.39-98.77% *Encrasicholina punctifer* 98.98% (1seq) *Engraulis australis* 98.77-98.36%	*Engraulis japonicus* 99% *Engraulis encrasicolus* 98% *Engraulis australis* 98%
ETN45.1 ETN45.2 ETN45.3	*Acetes chinensis*	658	No match	*Acetes chinensis* 96% *Acetes japonicus* 89% *Acetes serrulatus* 88%
ETN46.1 ETN46.2 ETN46.3	*Prionace glauca*	587-655	*Carcharhinus sorrah* 100-98.82% *Carcharhinus limbatus* 99.53-99.47%	*Carcharhinus sorrah* 100-99%
ETN47.1 ETN47.2 ETN47.3	*Trichiurus lepturus*	139	*Trichiurus lepturus* 99.26-98.52% *Trichiurus japonicus* 99.26-98.52% (v.n. *Trichiurus lepturus*)	*Trichiurus lepturus* 99% *Trichiurus japonicus* 99% (v.n. *Trichiurus lepturus*)
ETN48.1 ETN48.2 ETN48.3	*Miichthys miiuy*	655	No match	*Atrobucca nibe* 92%
ETN49.1 ETN49.2 ETN49.3	*Harpadon nehereus*	541-587	*Harpadon nehereus* 100-99% *Harpadon microchir* 99%	*Harpadon nehereus* 100% *Harpadon microchir* 100%
ETN50.1 ETN50.2 ETN50.3	*Carcharhinus* spp.	555	*Rhizoprionodon taylori* 100-99.64%	*Rhizoprionodon taylori* 99%
ETN51.1 ETN51.2 ETN51.3	*Gudusia chapra*	655	*Sardinella fimbriata* 100%	*Sardinella jussieu* 91%
ETN52.1 ETN52.2 ETN52.3	*Puntius sarana* (v.n. *Systosomus sarana*)	139	*Barbonymus gonionotus* 100%	*Barbonymus gonionotus* 100%
ETN53.1	***Otolithoides pama*** [a,b]	592	No match	*Panna microdon* 87% *Otolithoides biauritus* 87% *Bahaba taipingensis* 87%
ETN53.2 ETN53.3		655	No match	*Argyrosomus thorpei* 87% *Panna microdon* 87%

				Otolithoides biauritus 87%
ETN54.1 ETN54.2	*Nandus nandus*	655	*Nandus nandus* 99.69-98.99%	*Nandus nandus* 100-98%
ETN54.3	*Colisa fasciatus* (v.n. *Trichogaster fasciata*)	139	*Colisa fasciata* 98.64%	*Colisa fasciata* 99%
ETN54.4 ETN54.5		655	*Colisa fasciata* 99.85-99.38%	*Colisa fasciata* 99%
ETN54.6 ETN54.7		655	*Colisa fasciata* 99.69-99.38%	*Colisa fasciata* 99%
ETN54.8 ETN54.9	*Puntius sophore*	658	*Puntius sophore* 100%	*Puntius sophore* 100-97%
ETN54.10	***Macrognathus aculeatus***[b]	587	*Macrognathus pancalus* 100-98.14%	*Macrognathus pancalus* 99%
ETN54.11 ETN54.12	*Mystus cavasius*	655	*Mystus gulio* 99.84-96.59%	*Mystus gulio* 99-97%
ETN54.13 ETN54.14	*Mystus vittatus* *Labeo bata* *Glossogobius giuris*	655	*Chelon parsia* 100% *Paramugil parmatus* 98.96-98.65%	*Chelon parsia* 99% *Paramugil parmatus* 98%
ETN54.15		139	*Parapercis ommatura* 98.25%	*Parapercis ommatura* 98%
ETN54.16 ETN54.17		655	*Chelon parsia* 100-99.64% *Paramugil parmatus* 99.13-98.48%	*Chelon parsia* 99% *Paramugil parmatus* 98%
ETN55.1 ETN55.2 ETN55.3	*Anabas testudineus*	655	*Anabas testudineus* 100-91.4%	*Anabas testudineus* 100-92%
ETN56.1 ETN56.2 ETN56.3	*Prionace glauca*	658	*Carcharhinus brachyurus* 100-99.49% *Carcharhinus brevipinna* 98.61% *Carcharhinus falciformis* 98.61% *Carcharhinus sp.*98.61% (1 seq)	*Carcharhinus brachyurus* 99%
ETN57.1	*Clarias fuscus*	602	*Clarias gariepinus* 100-98.49% *Clarias cf stappersii* 99.33% (1seq) *Barbus altianalis* 98.83% *Barbus cercops* 98.83-98.66% (2 seq) *Clarias ngamensis* 98.66%	*Clarias gariepinus* 100-98.49%
ETN58.1 ETN58.2 ETN58.3	*Argyrosomus argentatus* (v.n. *Pennahia argentata*)	655	*Atrobucca nibe* 100-94.19%	*Atrobucca nibe* 95%
ETN59.1 ETN59.2 ETN59.3	--	139	*Sinonovacula constricta* 100-97.44%	*Sinonovacula constricta* 100-99%
ETN60.1 ETN60.2 ETN60.3	--	139	No match	*Bullacta exarata* 100-98%
ETN61.1 ETN61.2	*Tegillarca granosa*	658	*Tegillarca granosa* 99.83-98.67%	*Tegillarca granosa* 99-93%

ETN61.3				
ETN62.1 ETN62.2 ETN62.3	*Sepia esculenta*	629	*Uroteuthis duvauceli* 100-88.68%	*Uroteuthis duvauceli* 100-89%
ETN63.1 ETN63.2 ETN63.3	*Sepia esculenta*	658	*Uroteuthis chinensis* 99.85-98.92%	*Uroteuthis chinensis* 99-98%
ETN64.1 ETN64.2 ETN64.3	*Engraulis japonicus*	139	*Decapterus akaadsi* 100% *Trachurus japonicus* 100% (1seq) *Decapterus maruadsi* 100-98.89% *Decapterus russelli* 100-98.99% *Decapterus macarellus* 98.89% (1 seq)	*Decapterus maruadsi* 99% *Trachurus japonicus* 99% (1seq) *Decapterus akaadsi* 99% *Decapterus russelli* 99%
ETN65.1 ETN65.2	*Coilia mystus*	655	*Stolephorus* cf. *waitei* 4 100-99.85% *Stolephorus waitei* 100-84.33%	*Engraulis encrasicolus* 84% *Anchoviella* spp. 84% *Atherinella* spp. 84%
ETN65.3		655	*Encrasicholina heteroloba* 99.39-84.96%	*Atherinella* spp. 85% *Anchoa hepsetus* 85% *Engraulis encrasicolus* 85%
ETN66.1 ETN66.2 ETN66.3	*Engraulis japonicus*	139	*Cirrhinus molitorella* 100%	*Cirrhinus molitorella* 100%
ETN67.1 ETN67.2 ETN67.3	*"Pesce"*	139	*Decapterus akaadsi* 100% *Trachurus japonicus* 100% (1 seq) *Decapterus maruadsi* 100% *Decapterus russelli* 100 -99.26% *Decapterus macarellus* 98.89%	*Decapterus maruadsi* 100% *Trachurus japonicus* 100% (1seq) *Decapterus akaadsi* 100%
ETN68.1 ETN68.2	*"Pesce"*	139	*Oreochromis sp.* 100% (1 seq) *Sarotherodon galilaeus* 100% *Oreochromis aureus* 100% *Oreochromis sp. TP* 100% (1seq) *Sarotherodon lohbergeri* 100% (1seq) *Tilapia zilli* 100% *Oreochromis niloticus* 100% *Oreochromis mossambicus* 100%	*Sarotherodon galilaeus* 100% *Oreochromis niloticus* 100% *Oreochromis aureus* 100% *Sarotherodon lohbergeri* 100%

Table 1SM Results of the comparison of the sequences obtained in this work with BOLD and NCBI databases. Frauds are in grey background. Species showing an intraspecific range higher that 2% are highlighted in dark grey. [a]reference sequences absent from GenBank; [b]reference sequences absent from BOLD.

Sample	Species	Intraspecific variability range on BOLD	Intraspecific variability range on GenBank	Possible explanation
ETN2.1-2.2	*Stolephorus waitei*	15.51%	0%	Possible misidentification
ETN2.3-2.4		15.67%	Not retrieved	
ETN65.1-65.2		15.67%	Not retrieved	
ETN3.1	*Encrasicholina punctifer*	2.88%	0%	Possible misidentification
ETN5.1-5.3	*Encrasicholina heteroloba*	14.73%	Not retrieved	Possible misidentification
ETN65.3		14.43%	Not retrieved	
ETN36.1	*Atrobucca nibe*	6.01%	0%	Misidentification of specimens based on old classification (Sasaki, 1995)
ETN36.2-36.3		6.17%	1%	
ETN58.1-58.3		5.81%	0%	
ETN39.1-39.3	*Lagocephalus spadiceus*	6.49%	1%	Misidentification already reported by Cohen *et al*., (2009)
ETN43.1-43.3		6.07	1%	
ETN55.1-55.3	*Anabas testudineus*	8.6%	8%	A high intraspecific divergence (2-5.2%) between different populations of *A. testudineus* was already observed for the mitochondrial DNA control region (Jamsari *et al*., 2010)
ETN59.1-59.3	*Sinonovacula constricta*	2.56%	1%	Possible misidentification
ETN60.1-60.3	*Tegillarca granosa*	1.16%	6%	The high genetic distance (15.3%) found by Zheng *et al*., (2009) between two groups suggested significant genetic differentiation and the existence of different subspecies.
ETN62.1-62.3	*Uroteuthis duvauceli*	11.32%	11%	Previous genetic studies found percentage of *COI* sequence divergence varying from 6.8% (Da *et al.*, 2012) to 13.2% (Munasinghe & Thushari, 2014), suggesting the existence of cryptic species.

Table 2SM Species with an intraspecific variability range higher than 2% retrieved on one or both databases and hypothesized explanations.

References

Cohen, N.J., Deeds, J.R., Wong, E.S., Hanner, R.H., Yancy, H.F., White, K.D., Thompson, T.M., Wahl, M., Pham, T.D., Guichard, F.M., Huh, I., Austin, C., Dizikes, G., & Gerber, S.I. (2009). Public health response to puffer fish (tetrodotoxin) poisoning from mislabeled product. *Journal of Food Protection,*72(4), 810-817.

Da, L., Zheng, X., Kong, L., Li, Q. (2012) DNA barcoding analysis of Coleoidea (Mollusca: Cephalopoda) from Chinese waters. *Molecular Ecology Resources*, *12*, 437–447.

Jamsari, A.F.J., Muchlisin, Z.A., Musri,M., & Siti Azizah, M.N. (2010). Remarkably low genetic variation but high population differentiation in the climbing perch, *Anabas testudineus* (Anabantidae), based on the mtDNA control region. *Genetics and Molecular Research*, *9*(3), 1836-1843.

Munasinghe, D.H.N., & Thushari, G.G.N. (2014). Length-weight relationship and molecular phylogenetic analysis to infer status of *Uroteuthis duvaucelii* (d'Orbigny 1835) population in the southern coastal region of Sri Lanka. *International Journal of Fisheries and Aquatic Studies*, *2*(1): 223-231.

Sasaki, K. (1995). Two new species of *Atrobucca* (Sciaenidae) from the Bay of Bengal. *Japanese Journal of Ichthyology*, *42*, 269-276.

Zheng, W.J., Zhu, S.H., Shen, X.Q., Liu, B.Q., Pan, Z.C., Ye, Y.F. (2009) Genetic differentiation of *Tegillarca granosa* based on mitochondrial COI gene sequences. *Zoological Research*, 30, 17–23.

e-component

Click here to download e-component: Table 3SM revised.doc

Species habitat	Molecular identification					
	Total		Chinese shops		Bangladeshi shops	
	Samples	Species	Samples	Species	Samples	Species
Marine	24	14	23	13	1	1
Marine-brackish	6	5	5	4	1	1
Brackish-freshwater	6	5	0	0	6	5
Freshwater	5	5	1	1	4	4
Marine-brakish-freshwater	1	1	0	0	1	1
Marine + marine-brackish	30	19	28	17	2	2
Freshwater + brackish-freshwater	11	10	1	1	10	9

Table 4SM Habitat of the species identified with BOLD ID System and NCBI BLAST. Samples have been further subdivided according to their purchase in Chinese or Bangladeshi shops. Note that the number of samples is generally higher than the number of species because different products contained the same species.

e-component

Click here to download e-component: Table 5SM revised.doc