original scientific paper

DNA Barcoding as a Reliable Method for the Authentication of Commercial Seafood Products

Silvia Nicolè[1], Enrico Negrisolo[2], Giulia Eccher[1], Roberto Mantovani[3], Tomaso Patarnello[2], David L. Erickson[4], W. John Kress[4] and Gianni Barcaccia[1]*

[1]Laboratory of Genetics and Genomics, Department of Environmental Agronomy and Crop Science
[2]Department of Animal Science, Faculty of Agriculture;
[3]Department of Public Health, Comparative Pathology and Veterinary Hygiene, Faculty of Veterinary Medicine, University of Padova, Campus of Agripolis – Viale dell'Università 16, 35020 Legnaro, Padova, Italy
[4]Department of Botany and Laboratories of Analytical Biology – National Museum of Natural History, Smithsonian Institution, P.O. Box 37012, Washington, DC 20013-7012, USA

Received: January 14, 2011
Accepted: February 7, 2012

UNCORRECTED PROOF

[*]Corresponding author; Phone: +39 049 827 2814; Fax: ++39 049 827 2839; E-mail: gianni.barcaccia@unipd.it

The nucleotide sequences of the mtDNA barcode regions that were obtained from this study have been recorded in the GenBank/EMBL databases under the accession numbers GU324135-GU324234

Summary

Animal DNA barcoding allows researchers to identify different species by analyzing a short nucleotide sequence, typically the mitochondrial gene *cox1*. In this paper, we use DNA barcoding to genetically identify seafood samples that were purchased from various locations throughout Italy. We adopted a multi-locus approach and analyzed the *cob*, *16S-rDNA* and *cox1* genes, and compared our sequences to reference sequences in the BOLD and GenBank online databases. Our method is a rapid and robust technique that can be used to genetically identify crustaceans, mollusks and fishes. This approach could be applied in the future to conservation, particularly for monitoring illegal trade of protected and endangered species. Additionally, this method could be used for authentication in order to detect mislabeling of commercial processed seafood.

Key words: DNA barcoding, genetic traceability, BOLD, seafood, mislabeling

Introduction

Nowadays DNA barcoding is an established technique that involves sequencing gene segments and comparing the results with orthologous reference sequences in public databases (*1*). The goal of DNA barcoding is to identify the sample species by sequencing a single gene that is universally amplified across metazoans using as template genetic materials isolated from a small portion of organisms in any stage of their life history; theoretically, all species are delineated by their unique barcode sequence, or by a tight cluster of very similar sequences (*2*). The core assumption of DNA barcoding is that variation in the nucleotide sequence is lower within a species than between different species (*3*). Typically, genes within the mitochondrial genome are used for DNA barcoding of animals. The mitochondrial genome has a higher rate of mutation compared to the nuclear genome, is maternally inherited, has less hybridization and has a high copy number, which facilitates PCR amplification and sequence recovery from degraded tissue (*4*,*5*). Furthermore, the mitochondrial genome lacks introns, pseudogenes and repetitive sequences, which facilitates sequence alignments of the amplified genes (*6*). Finally, complete

UNCORRECTED PROOF

mtDNA genome sequences are publicly available; primers can therefore be designed to amplify and sequence any species that has a published mtDNA genome (*7-10*).

The *cox1* gene, which encodes for the cytochrome oxidase subunit I, was originally proposed to be a specific mitochondrial marker for animal DNA barcoding. Researchers identified a 648 nucleotide sequence near the 5' end of the gene that could be amplified and sequenced with a pair of universal primers for a wide range of animal species (*1,7,11*). Molecular identification based on *cox1* sequencing has provided reliable results for several animal groups that have been tested (*6,12*), and has also been expanded to include high resolution at the species level for fish (*13-18*). Based on these encouraging results, the barcoding community has established a Fish-BOL (Fish Barcode of Life) initiative that seeks to assemble a comprehensive reference system based on *cox1* sequencing for all of the estimated 20,000 marine and 15,000 fresh-water fish species (http://www.fishbol.org/index.php). The main goal of this project is to help manage fish biodiversity and develop the Catalogue of Life (http://www.catalogueoflife.org/search.php), which is an exhaustive database that contains all of Earth's known species. Although *cox1* is the standard gene for DNA barcoding in animals, other mitochondrial genes have been suggested as barcode markers: *cob*, which encodes for Apocytochrome-b (*6,19*), is typically used as a marker for phylogenetic analyses in several taxa; *cox2* and *cox3*, which encode for the mitochondrial cytochrome oxidase subunits II and III, respectively (*20*); *nad1*, which encodes for the NADH dehydrogenase 1 subunit (*11*) and mitochondrial *16S-rDNA* (*21*).

DNA barcoding can be applied to several fields, including biodiversity monitoring (*e.g.*, taxonomic, ecological and conservation studies) and forensic science. Additionally, DNA barcoding can be used to identify organisms that lack distinctive morphological features (*i.e.*, are in the larval stage (*22*) or because of homoplasy and phenotypic plasticity of a given diagnostic character to environmental factors (*23*). Moreover, DNA barcoding could be used to monitor the illegal trade of wildlife, such as protected or endangered species (*24,25*), or to identify the species origin of commercially processed food (*16,26*). Application of DNA barcoding for food authentication has recently gained attention because of food safety concerns, including incorrect food labeling, food substitutions (*27,28*) or recent food contamination (*29*). Proper identification of species contained in food is important for protecting consumers against potential food adulteration, ingredient mislabeling (*26*), GMOs (*30*) or food poisoning (*31*). According to U.S.

UNCORRECTED PROOF

Food and Drug Administration, fish substitutions in seafood derivatives are becoming increasingly common; thus, analytical methods to verify food labeling are needed (http://www.fda.gov/Food/FoodSafety/Product-SpecificInformation/Seafood/RegulatoryFishEncyclopediaRFE/ucm071528.htm).

Several approaches can be used to identify fish species, including traditional morphological observations and molecular analyses, such as genomic and proteomic techniques (*32-34*). In this paper, we report the use DNA barcoding to genetically identify the marine species present in seafood products. Sequence-specific universal primers were designed to amplify by PCR three distinct mitochondrial genes (*16S-rDNA, cox1* and *cob*) in raw, frozen and processed commercial seafood. Two approaches were employed to assess the ability of DNA barcoding for the authentication of the originating species: a similarity search was conducted with the DNA identification engine at BOLD (Barcode of Life Database), based on the HMM algorithm (*35*) and BLAST algorithm of GenBank (*36*); and NJ trees were built with a distance-based approach to illustrate sequence identity based on tree topology. The main goal of this study was to verify the label information of several seafood products with a multi-locus, DNA barcoding strategy; additionally, we sought to estimate and compare the reliability of the two most common gene repositories that are used for phylogenetic and forensic purposes: the GenBank and BOLD databases.

Materials and Methods

Seafood sample collection

A total of 37 seafood samples, including raw, frozen and processed products from various commercial brands were collected from markets and groceries of North-Eastern Italy. Most of the samples were labeled with the genus and species, as well as the common name and capture location, which is required by law. Additional samples were obtained at the Chioggia marketplace (Venice), and these samples were labeled with only the common name and the area of origin. Specifically, the 37 samples was comprised of 30 fish, 3 crustacean and 4 mollusk samples; some of the samples were sold as fresh or frozen skinned fillets, while others had undergone different treatments, such as heat treatment or canning (Table 1). Three seafood

UNCORRECTED PROOF

products included more than one species, and the scientific names of the organisms were indicated on the label.

DNA extraction, amplification and sequencing

The total genomic DNA was extracted and purified from the 37 samples with a GenElute Mammalian Genomic DNA Miniprep Kit (SIGMA) according to the manufacturer's instructions with a few changes. The specific DNA barcode region of each target gene was amplified in two independent reactions. All of the PCR experiments were performed with a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA), and the amplification of the target DNA sequences was performed according to instructions at the Barcoding Animal Life website (http://www.dnabarcoding.ca/primer/Index.html).

The PCR conditions for *cox1* amplification were initial denaturation at 94 °C for 1 min; five cycles of denaturation at 94 °C for 30 s, annealing at 50-55 °C for 40 s, and extension at 72 °C for 1 min; 30-35 subsequent cycles of 94 °C for 30 s, 55-60 °C for 40 s, and 72 °C for 1 min and a final extension at 72 °C for 10 min; all of the amplified sequences were stored at 4 °C. One pair of universal primers was used for *16S-rDNA* and *cob*, whereas several primer pairs were used for the *cox1* gene. We first tested the *cox1* universal primers, described by Ward and coworkers (*2*), which unexpectedly failed in some specimens, and subsequently adopted additional *cox1* specific primers, which were formerly designed by Former and coworkers (*7*). The primer sequences, as well as the respective annealing temperatures and related references are listed in Table 2. PCR reactions, performed in a volume of 25 μl, were comprised of 1× PCR buffer (100 mM Tris-HCl pH 9.0, 15 mM $MgCl_2$ and 500 mM KCl), 0.2 mM dNTPs, 0.4 μM of each primer, 1 U of *Taq* DNA polymerase and 15 ng of genomic DNA template. The PCR products were enzymatically purified with EXO/SAP (Amersham) and were sequenced bi-directionally with an original Rhodamine terminator cycle sequencing kit (ABI PRISM Applied Biosystems) using an ABI31030xl genetic analyzer (Applied Biosystems). The sequences were assembled into contigs, screened for errors with Mega V 4.1 beta (*37*) and exported in FASTA format for future database searches and tree-based alignments.

UNCORRECTED PROOF

BLAST and phylogenetic analyses

For species identification, both a similarity analysis and a phylogenetic approach were used to compare the amplified sequences to reference sequences in the GenBank and BOLD databases. BOLD is a new sequence repository that was created as a reference for large-scale DNA barcoding projects (http://www.barcodinglife.org/views/login.php).

Similarity searches were conducted by comparing the BLAST algorithm to the GenBank database, as well as by comparing a global alignment with the Hidden Markov Model (HMM) to the BOLD engine. GenBank was used as a reference database for all of the markers, while BOLD was used only for the *cox1* sequences. Two tiers of comparison were used for specimen identification with BOLD. The first tier compared a reference subset of the database, which was comprised of only validated sequences, to at least three voucher samples. If BOLD interrogation reported no match, we used the full database of *cox1* barcode sequences, which includes non-validated records that have no species identification. Phylogenetic analysis of *cox1* was performed with the CLC Sequence Viewer 6.2 software (CLC bio). The genetic distances were calculated with the Kimura model (*38*), and the Neighbor-Joining method was used as the tree-building algorithm (*39*). When possible, four validated sequences were downloaded from the BOLD database for each species, and these four sequences were analyzed with the experimental sequences. When the *cox1* sequence was not available in the BOLD database, we used sequences that corresponded to the sample species that were downloaded from GenBank. All of the sequences downloaded from GenBank or BOLD that were used to construct the distance trees are listed in Supplementary Table 1S. The reliability of the species tree clusters was evaluated with a bootstrap test with 1,000 replications. An additional NJ tree was produced with the Mega v.4.1 software for genera in which it was difficult to separate the individual species, such as *Thunnus*, *Macruronus* and *Gadus*. In these cases, all of the *cox1* sequences were downloaded from BOLD; if BOLD had a small number of entries, then the sequences were downloaded from GenBank. The downloaded *cox1* sequences were used to draw genus-specific trees to clarify the relationships among species within these genera.

UNCORRECTED PROOF

Results and Discussion

DNA extraction and PCR-based amplification

We successfully isolated total genomic DNA from all 37 of the seafood-derived specimens, including raw and frozen processed products, as well as skinned fillets. Moreover, the isolated DNA was successfully amplified via PCR. The PCR conditions, as well as the universal primers that were used (see Table 1) followed protocols listed on the official barcode website (http://www.dnabarcoding.ca). The primers reproducibly generated single amplification products that had an average length of approximately 700 bp for *cox1*, 500 bp for *16S-rDNA* and 850 bp for *cob*. All of the mtDNA sequences were deposited in the NCBI database on December 12, 2009 under the GenBank accession numbers GU324135-GU324234.

Primers that were designed for *16S-rDNA* universally amplified DNA from all of the commercial samples except one. The *16S-rDNA* primers reproducibly generated a single amplification product from the fish, mollusk and crustacean seafood derivatives, whereas the *cox1* and *cob* primers did not amplify DNA from all of the samples. For instance, *cox1* amplification failed in two crustaceans and one mollusk, while the *cob* primers did not amplify DNA from any of the mollusks. Because food sample mislabeling generally involves fresh fillets that are sold in local marketplaces, our goal was to focus the analysis on fresh raw fillets and frozen samples, thereby avoiding problems that are related to isolation of genomic DNA from processed food (see Table 1). However, future studies must be performed to verify that the primer pairs and the PCR conditions used in this study can be used to amplify specimens that have been exposed to highly denaturating treatments, such as high temperature and low pH, which can damage the DNA and hamper amplification of target regions that are longer than 200 bp (*40*,*41*).

Selected marker validation

To verify the species that were listed on the label of each seafood product, we analyzed and compared sequences in the reference barcode regions of *cox1*, *16S-rDNA* and *cob*. These gene sequences were chosen because they represent common regions that are typically used for species identification; additionally, these genes have the widest taxonomic representation in the NCBI nucleotide databases (Supplementary Fig. 1S). We chose to sequence more than one target gene to obtain independent replicates and confirm our results. Furthermore, because the BOLD

UNCORRECTED PROOF

sequence database is incomplete, sequences of two additional genes improved the success rate of species identification, especially when *cox1* sequences were unavailable (*16*). Thus, a reliable and comprehensive database with adequate reference sequences is needed for accurate species identification (*42*).

BLAST and NJ distance-based analyses

A double approach was used to verify the identity of our samples: a similarity search to probe the GenBank and BOLD databases, and a distance-based approach that is routinely used for barcoding analyses. With the similarity search, we compared the experimentally obtained DNA sequences to publicly available sequences in the GenBank and BOLD databases. To compare the data to GenBank sequences we used the BLAST program, which is an algorithm that locates regions with local similarity between sequences; in contrast, the BOLD engine identifies a species by rapidly aligning a query sequence to a global alignment of all reference sequences, followed by a linear search of the reference library. However, BOLD genetic identification can only identify the species if there is less than 1 % divergence from the reference sequence (*43*). Because the DNA barcoding technology is standardized and affordable, the major limitation is the breadth and availability of reference sequences; larger databases that contain more sequences will increase the identification of unknown samples. Because BOLD was developed with voucher samples, the BOLD repository should contain only validated sequences that can be used for identification purposes (*44*). However, only a subset of the BOLD repository is validated and it includes sequence records from species that were identified by three or more individuals and have less than a 2 % sequence divergence. In contrast, GenBank contains a mixture of verified and unverified sequences because there was no quality control during the sequence submission phase (*45*). We used GenBank when *cox1* sequences for the target species were not publicly available in the BOLD database. Utilization of both the BOLD and GenBank databases also allowed us to verify the accuracy of BOLD-based identification. BLAST analysis provides a list of sequences that are most similar to the query sequence, as well as a BIT score, which estimates the percent identity (%), and an E-value. In contrast, BOLD analysis identifies the species by the degree of nucleotide variation, with similar species having a divergence value of less than 1 %; if no match is obtained, the query sequence is assigned to a genus with a similarity divergence of less than 3%. With the mitochondrial DNA barcodes that were utilized in this study, 16 out of 37

UNCORRECTED PROOF

samples were identified as a single species from all three of the marker genes, and these species matched the species that were labeled on the food (Table 3). Additionally, 12 seafood samples were identified as a single species from two of the marker genes, and the identified species matched the label information. The remaining five samples were identified based on only one marker gene, and the identified species did not match the label information (see Table 3). For instance, sample number 28 was labeled as river perch, which belongs to the genus and species *Perca luviatilis*. However, our molecular analyses of sample number 28 show that it is not *Perca luviatilis*, as all three molecular markers suggest that it belongs to the genus *Paralichthys*, which is a flounder that has a lower market value than perch. This product could have been falsely labeled to increase the market value, as is common.

The marker that had the highest number of positive and unambiguous matches was the *cox1* gene; of the 37 samples, 26 matched the reference sequences in the GenBank database and 21 matched the reference sequences in the BOLD database. This actual number of matched samples may be lower than expected because there were problems with identifying the various species of tuna, which could only be placed into the appropriate genus. With the *16S-rDNA* data we were able to identify the species for 28 of the samples, although nine of these matches produced equal identity scores for more than one species. In contrast, the *cob* gene was the least reliable marker and only identified the species of 21 products, five of which had equal scores for several species. Although *cob* was originally thought to be an optimal marker for phylogenetic and forensic studies in animals, new international studies have begun to use the *cox1* gene instead. Five samples could not be identified by their *cob* sequence because the reference sequence was unavailable in GenBank. This problem was much more frequent with *cob* rather than *cox1* or *16S-rDNA* (see Table 3) as only two *16S-rDNA* and four *cox1* sequences were unavailable in GenBank and BOLD, respectively.

A phylogenetic approach based on construction of a Neighbor-Joining tree with validated *cox1* reference sequences was employed to validate the results from the similarity searches (Fig. 1). In the NJ tree, samples that belong to the same species are clustered in the same monophyletic group; the exceptions are the *Thunnus, Macruronus* and *Gadus* species, which had poor resolution of the species clusters. The majority of samples that we collected were grouped with the species listed on the label; however, a few species had a *cox1* sequence that clustered them with a different species, indicating that the food samples were either substituted or mislabeled.

UNCORRECTED PROOF

Classification and distinction among species within the *Thunnus*, *Macruronus* and *Gadus* genera is often questioned because of conflicting data from systematic approaches, as well as morphological and molecular analyses. In this study, species identification based on the nucleotide divergence of the *cox1* gene did not correctly identify samples for these genera. To identify the correct species, a second NJ tree for each problematic genus was constructed with and without (data not shown) the sequences that correspond to the experimental specimens.

The trees obtained for the genus *Thunnus* were well resolved, except for species that belong to the subgenus *Neothunnus* (*i.e., T. albacares, T. atlanticus* and *T. tonggol*), where the *T. albacares* sequences did not group together. Furthermore, *T. alalunga* and *T. orientalis* could not be differentiated because the marker used was unable to differentiate between these two species that have similar genomes (*46*). It was recently shown that *cox1* cannot be reliably exploited to differentiate between tuna species because of its low genetic variability compared to other markers, such as mitochondrial DNA (*47*). Species resolution with an NJ tree is not limited by the number of sequences available in a database, and thus our inability to discriminate between various species with sequences from the BOLD database may have been caused by two things: incorrect identification of the reference sequences, or by the fact that the *Thunnus* genus has a complex phylogenetic history that contains frequent introgression events. While individual species could be identified for the *Thunnus* genus with *16S-rDNA* and *cob*, only genera could be identified with *cox1*. The inability of *cox1* to identify a species may be related to evolution of the eight *Thunnus* species; *cox1* may not be sensitive enough to detect frequent introgressive hybridizations among tuna species (*47*).

The NJ tree for the *Macruronus* species had two major clusters that did not distinctly separate the two common species, *M. novaezelandiae* and *M. magellanicus*. Similar to the *Thunnus* genus, *Macruronus* is another genus that has taxonomic uncertainty. Division into these two species based upon morphometric analysis and geographical distributions (*48*) has been recently disputed (*49*). Morphological similarities for the larval and adult stages, as well as low genetic divergence in the mitochondrial *cob* and *cox1* genes suggest that these two species may be the same species (*50*). *Gadus* taxonomy is also problematic. Some researchers divide the genus into the three species based on morphological traits during the larval phase, *G. morhua, G. ogac* and *G. macrocephalus*. However, the mitochondrial *cob* sequence is similar for the *G. ogac* and *G. macrocephalus* species, which suggests that they are a single species (*51*). Our *cox1*

UNCORRECTED PROOF

marker was not able to differentiate between *G. ogac* and *G. macrocephalus*, again suggesting that the two species are the same organism.

In general, our results from this study support the use of *cox1*-based identification of fish samples with the BOLD engine. Although the GenBank database is the best online tool for forensic studies, the BOLD database had enough sequences to allow us to identify a majority of our collected samples. There are several methods that can be used to identify a species, including morphological observations, molecular proteomic techniques and/or DNA sequencing. Compared to proteins, DNA is a much more stable molecule that can withstand harsh treatment conditions (*52*). Older gel-based sequencing methods, such as PCR-RFLP or PCR-SSCP, could identify a sample sequence of a short target region from highly processed food (*52*,*53*). Unlike a previous procedure termed FINS (Forensically Informative Nucleotide Sequencing) (*54*), which uses different genetic markers for different taxonomic groups, DNA barcoding allows researchers to develop a *cox1*-based library for all of the metazoans on the Earth. DNA barcoding is a rapid and sensitive method that can be used to monitor commercial food for mislabeling or substitution (*55*).

Conclusions

In recent years, the DNA barcoding has been exploited for both phylogenetic (research) and commercial (legal) purposes in many taxa, not only fishes but also crustaceans and mollusks (*56*). In fact, if it is true that DNA barcoding was initially exploited almost exclusively for biodiversity studies in marine organisms, it is also true that more recently DNA barcoding was also proposed for genetic traceability of seafood products (*56-67*). Moreover, barcodes have been used for population and conservation genetics, phylogeographic studies, detection of invasive species, forensics, and seafood safety (*56* and citations herewith). More controversially, barcodes have been adopted to delimit species boundaries, reveal cryptic species, and discover new species.

The results from this study show that conventional DNA barcoding is an efficient tool that can be used to identify food components and thus validate label information contents. This methodology may be offset by the need to have relatively long barcodes for sequence comparison. For the analysis of degraded samples or processed products, other unique markers may be much shorter and hence useful to this aim. In general, mtDNA barcoding is simple, robust and cost-effective, which make it suitable for seafood authentication assays, even with

UNCORRECTED PROOF

degraded samples, processed products or small portions of any life-history stage of a variety of fishes, crustaceans and mollusks. Our data also show that misidentification can be caused by absent or erroneous reference sequence entries, thus highlighting the need to expand and improve the *cox1* sequences within the BOLD reference database. A reliable and comprehensive DNA sequence library is essential for successful application of DNA barcoding. There were several instances of ambiguous species identification because of poor resolution of the constructed NJ trees in samples that had low nucleotide variation, such as the *Macruronus, Thunnus* and *Gadus* genera. These species could eventually be identified by using additional genetic (nuclear and/or mitochondrial) markers that have appropriate mutation rates. If several genetic markers indicate a species that is not included in the food label, this may indicate false labeling, which was detected for five of the samples used in this study.

In terms of innovation, emerging frontiers may be the use of barcodes for rapid single-species identity and automated among-species diversity assessment by high-throughput sequencing of informative DNA barcodes. For species identification by means of DNA-based assays, there are two basic approaches: sequencing of multiple barcodes or detection of unique markers. It is well known that the number of commercial fish species sold on the markets is low compared to the total number of species living in the seas. Nevertheless, genomes of commercial species are increasingly characterized and some of them are fully or partially sequenced. This means that instead of using one or a few target DNA regions for all species, in a near future it will be methodologically simpler and less expensive to use a set of barcodes for a group of marine organisms, taking one or two unique markers for each species.

Acknowledgements

The authors gratefully acknowledge funding for this research from the Italian Ministry of University, Research, Science and Technology (Project: “Development of molecular diagnostic assays for the genetic traceability of agrifood products”). The authors wish to thank Silvio Collani for his help with editing of the nucleotide sequences. Silvia Nicolè learned the DNA barcoding technique that is the foundation of this study during her internship at the Laboratories of Analytical Biology of the Smithsonian Institution, Washington DC (USA).

UNCORRECTED PROOF

References

1. P.D.N. Hebert, A. Cywinska, S.L. Ball, J.R. deWaard, Biological identifications through DNA barcodes, *Proc. Royal Society B. 270* (2003) 313-321.
2. R.D. Ward, T.S. Zemlak, B.H. Innes, P.R. Last, P.D.N. Hebert, DNA barcoding Australia's fish species, *Philos. Trans. Royal Society B. 360* (2005) 1847-1857.
3. C.P. Meyer, G. Paulay, DNA barcoding: error rates based on comprehensive sampling, *PLoS Biology, 3* (2005) e422.
4. C. Saccone, G. De Carla, C. Gissi, G. Pesole, A. Reynes, Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system, *Gene, 238* (1999) 195-210.
5. P.D.N. Hebert, M.Y Stoeckle, T.S. Zemlak, C.M. Francis, Identification of birds through DNA barcodes, *PloS Biology, 2* (2004) 1657-1663.
6. W.F. Lin, C.Y. Shiau, D.F. Hwang, Identification of four *Thunnus* tuna species using mitochondrial cytochrome b gene sequence and PCR-RFLP analysis, *J. Food Drug Analysis, 13* (2005) 382-387.
7. O. Folmer, M. Black, W. Hoeth, R. Lutz, R. Vrijenhoek, DNA primers for amplification of mithocondrial cytochrome c oxydase subunit I from diverse metazoan invertebrates, *Mol. Marine Biol. Biotech. 3* (1994) 294-299.
8. J.F. Montiel-Sosa, E. Ruiz-Pesini, J. Montoya, P. Roncales, M.J. Lopez-Perez, A. Perez-Martos, Direct and highly species-specific detection of pork meat and fat in meat products by PCR amplification of mitochondrial DNA, *J. Agric. Food Chem. 48* (2000) 2829-2832.
9. R.B. Simmon, S.J. Weller, Utility and evolution of cytochrome *b* in insects, *Mol. Phylogen. Evol. 20* (2001) 196-210.
10. W.F. Lin, D.F. Hwang, Application of species-specific PCR for the identification of dried bonito product (Katsuobushi), *Food Chem. 106* (2008): 390-396.
11. J. Rach, R. Desalle, I.N. Sarkar, B. Schierwater, H. Hadrys, Character-based DNA barcoding allows discrimination of genera, species and populations in *Odonata*, *Proc. Royal Society B. 275* (2008) 237-247.
12. P.D.N. Hebert, E.H. Penton, J.M. Burns, D.H. Janzen, W. Hallwachs, Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*, *Proc. Natl. Acad. Sci. USA, 101* (2004) 14812-14817.

UNCORRECTED PROOF

13. I.D. Hogg, P.D.N. Hebert, Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes, *Canad. J. Zoology, 82* (2004) 749-754.

14. M. Hajibabaei, D.H. Janzen, J. Burns, W. Hallwachs, P.D.N. Hebert, DNA barcodes distinguish species of tropical Lepidoptera, *Proc. Natl. Acad. Sci. USA, 103* (2006) 968-971.

15. H.S. Yoo, J.Y. Eah, J.S. Kim, M.S. Min, W.K. Paek, H. Lee, C.B. Kim, DNA barcoding Korean birds, *Molecules Cells, 22* (2006) 323-327.

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

17. N.V. Ivanova, T.S. Zemlak, R.H. Hanner, P.D.N. Hebert, Universal primer cocktails for fish DNA barcoding, *Mol. Ecology Notes, 7* (2007) 544-548.

18. N. Hubert, R. Hanner, E. Holm, N.E. Mandrak, E. Taylor, M. Burridge, D. Watkinson, P. Dumont, A. Curry, P. Bentzen, J. Zhang, J. April, L. Bernatchez, Identifying Canadian freshwater fishes through DNA barcodes. *PLoS ONE, 3* (2008) e2490.

19. T. Pepe, M. Trotta, I. Di Marco, A. Anastasio, J.M. Bautista, M.L. Cortesi, Fish species identification in surimi-based products. *J. Agric. Food Chemistry, 55* (2007) 3681-3685.

20. M.H. Park, C.J. Sim, J. Baek, G.S. Min, Identification of genes suitable for DNA barcoding of morphologically indistinguishable Korean halichondriidae sponges, *Molecules Cells, 23* (2007) 220-227.

21. S. Willows-Munro, T.J. Robinson, C.A. Matthee, Utility of nuclear DNA intron markers at lower taxonomic levels: phylogenetic resolution among the nine *Tragelaphus* spp., *Molec. Phylogen. Evol. 35* (2005) 624-636.

22. G.G. Pegg, B. Sinclair, L. Briskey, W.J. Aspden, mtDNA barcode identification of fish larvae in the southern Great Barrier Reef, Australia, *Scientia Marina, 70* (2006) 7-12.

23. M. Vences, M. Thoma, R.M. Bonett, D.R. Vieites, Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philos. Trans. Royal Society B. 360* (2005) 1859-1868.

24. C.S. Baker, G.M. Lento, F. Cipriano, S.R. Palumbi, Predicted decline of protected whales based on molecular genetic monitoring of Japanese and Korean markets. *Proc. Royal Society B. 267* (2000) 1191-1199.

UNCORRECTED PROOF

25. M. Shivji, S. Clarke, M. Pank, L. Natanson, N. Kohler, M. Stanhope, Genetic identification of pelagic shark body parts for conservation and trade monitoring, *Conservation Biology, 16* (2002) 1036-1047.

26. S. Tanabe, E. Miyauchi, A. Muneshige, K. Mio, C. Sato, M. Sato, PCR method of detecting pork in foods for verifying allergen labelling and for identifying hidden pork ingredients in processed foods, *Biosci. Biotech. Biochem. 71* (2007) 1663-1667.

27. Y.H.P. Hsieh, Species substitution of restaurant fish entrees, *J. Food Quality, 21* (1996) 1-11.

28. P.B. Marko, S.C. Lee, A.M. Rice, J.M. Gramling, T.M. Fitzhenry, J.S. McAlister, G.R. Harper, A.L. Moran, Mislabelling of a depleted reef fish, *Nature, 430* (2004) 309-310.

29. F. Teletchea, C. Maudet, C. Hanni, Food and forensic molecular identification: update and challenges, *Trends Biotech. 23* (2005) 359-366.

30. S.B. Ronning, K. Rudi, K.G. Berdal, A. Holst-Jensen, Differentiation of important and closely related cereal plant species (Poaceae) in food by hybridization to an oligonucleotide array, *J. Agric. Food Chemistry, 53* (2005) 8874-8880.

31. Y.H. Hsieh, Y.C. Shiu, C.A. Cheng, S.K. Chen, D.F. Hwang, Identification of toxin and fish species in cooked fish liver implicated in food poisoning. *J. Food Sci. 67* (2002) 948-952.

32. H. Rehbein, M. Etienne, M. Jerome, T. Hattula, L.B. Knudsen, F. Jessen, J.B. Luten, W. Bouquet, I. Mackie, A.H. Ritchie, R. Martin, R. Mendes, Influence of variation in methodology on the reliability of the isoelectric focusing method of fish species identification, *Food Chemistry, 52* (1995) 193-197.

33. I. Martinez, A.K. Danielsdottir, Identification of marine mammal species in food products, *J. Sci. Food Agric. 80* (2000) 527-533.

34. M. Trotta, S. Schonhuth, T. Pepe, M.L. Cortesi, A. Puyet, J.M. Bautista, Multiplex PCR methods for use in real-time PCR for identification of fish fillets from grouper (*Epinephelus* and *Mycteroperca* species) and common substitute species, *J. Agric. Food Chemistry, 53* (2005) 2039-2045.

35. S.R. Eddy, Profile hidden Markov models. *Bioinformatics, 14* (1998) 755-763.

36. S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, *J. Mol. Biol. 215* (1998) 403-410.

UNCORRECTED PROOF

37. S. Kumar, J. Dudley, M. Nei, K. Tamura, MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences, *Brief. Bioinformatics, 9* (2008) 299-306.

38. M. Kimura, A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol. 16* (1980) 111-120.

39. N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Mol. Biol. Evol. 4* (1987) 406-425.

40. M.J. Chapela, C.G. Sotelo, R.I. Perez-Martin, M.A. Pardo, B. Perez-Villareal, P. Gilardi, J. Riese, Comparison of DNA extraction methods from muscle of canned tuna for species identification, *Food Control, 18* (2007) 1211-1215.

41. R.S. Rasmussen, M.T. Morrissey, DNA-based methods for the identification of commercial fish and seafood species, *Comprehensive Reviews Food Sci. Food Safety, 7* (2008) 280-295.

42. T. Ekrem, E. Willassen, E. Stur, A comprehensive DNA sequence library is essential for identification with DNA barcodes, *Mol. Phylogen. Evol. 43* (2007) 530-542.

43. S. Ratnasingham, P.D.N. Hebert, BOLD: The Barcode of Life Data system, *Mol. Ecology Notes, 7* (2007) 355-364.

44. E. H.-K.Wong, R.H. Hanner, DNA barcoding detects market substitution in North-American seafood, *Food Res. Intl. 41* (2008) 828-837.

45. P. Forster, P. To err is human, *Annals Human Genetics, 67* (2003) 2-4.

46. J.R. Alvarado Bremer, I. Naseri, B. Ely, Orthodox and unorthodox phylogenetic relationships among tunas revealed by the nucleotide sequence analysis of the mitochondrial control region, *J. Fish Biology, 50* (1997) 540-554.

47. J. Vinas, S. Tudela, A validated methodology for genetic identification of tuna species (Genus *Thunnus*). *PLoS ONE, 4* (2009) e7606.

48. T. Inada, Family Merluccidae. In: Cohen D. M., Inada T., Iwamoto T., Scialabba N., *An annotated and illustrated catalogue of cods, hakes, grenadiers and other gadiform fishes known to date* (Ed. Gadiform fishes of the world (Order Gadiformes)) Rome, Italy: FAO Fisheries Synopsis 125, pp. 319-346.

49. F. Balbontin, F. Uribe, R. Bernal, M. Braun, Descriptions of larvae of *Merluccius australis*, *Macruronus magellanicus*, and observations on a larva of *Micromesistius australis* from

UNCORRECTED PROOF

Southern Chile (Pisces: Gadiformes), *New Zealand J. Marine Freshwater Res. 38* (2004) 609-619.

50. C. Olavarria, F. Balbontin, R. Bernal, C.S. Baker, Lack of divergence in the mitochondrial cytochrome b gene between *Macruronus* species (Pisces: Merlucciidae) in the Southern Hemisphere, *New Zealand J. Marine Freshwater Res. 40* (2006) 299-304.

51. S.M. Carr, D.S. Kivlichan, P. Pepin, D.C. Crutcher, Molecular systematics of gadid fishes: implications for the biogeographic origins of Pacific species, *Canad. J. Zoology, 77* (1999) 1-12.

52. M. Unseld, B. Beyermann, P. Brandt, R. Hiesel, Identification of the species origin of highly processed meat products by mitochondrial DNA sequences, *PCR Methods Applications, 4* (1995) 241-243.

53. M. Hajibabaei, M.A. Smith, D.H. Janzen, J.J. Rodriguez, J.B. Whitfield, P.D.N. Hebert, A minimalist barcode can identify a specimens whose DNA is degraded, *Mol. Ecology Notes, 6* (2006) 959-964.

54. S.E. Bartlett, W.S. Davidson, FINS (Forensically Informative Nucleotide Sequencing): a procedure for identifying the animal origin of biological specimens, *Biotechniques, 12* (1992) 408-411.

55. C.A. Logan, S.E. Alter, A.J. Haupt, K. Tomalty, S.R. Palumbi, An impediment to consumer choice: overfished species are sold as Pacific red snapper, *Biological Conservation, 141* (2008) 1591-1599.

56. S. Kim, H.S. Eo, H. Koo, J.K. Choi, W. Kim, DNA barcode-based molecular identification system for fish species. *Mol. Cells*, 30 (2010) 507-512.

57. S. Arami, M. Sato, S. Futo, Applicability of DNA barcode for identification of fish species [in Japanese]. *Shokuhin Eiseigaku Zasshi*, 52 (2011) 205-210.

58. L.M. Aquino, J.M. Tango, R.J. Canoy, I.K. Fontanilla, Z.U. Basiao, P.S. Ong, J.P. Quilang, DNA barcoding of fishes of Laguna de Bay, Philippines. *Mitochondrial DNA*, 22 (2011) 143-153.

59. W.S. Lakra, M.S. Verma, M. Goswami, K.K. Lal, V. Mohindra, P. Punia, A. Gopalakrishnan, K.V. Singh, R.D. Ward, P. Hebert, DNA barcoding Indian marine fishes. *Mol. Ecol. Resour*., 11 (2011) 60-71.

UNCORRECTED PROOF

60. S.M. Handy, J.R. Deeds, N.V. Ivanova, P.D. Hebert, R.H. Hanner, A. Ormos, L.A. Weigt, M.M. Moore, H.F. Yancy, A single-laboratory validated method for the generation of DNA barcodes for the identification of fish for regulatory compliance. *J. AOAC Int*., 94 (2011) 201-210.

61. M. Moftah, S.H. Abdel Aziz, S. Elramah, A. Favereaux, Classification of sharks in the Egyptian mediterranean waters using morphological and DNA barcoding approaches. *PLoS One*, 6 (2011) e27001.

62. M. Kochzius, C. Seidel, A. Antoniou A., S.K. Botla, D. Campo, A. Cariani, E.G. Vazquez, J. Hauschild, C. Hervet, S. Hjörleifsdottir, G. Hreggvidsson, K. Kappel, M. Landi, A. Magoulas, V. Marteinsson, M. Nölte, S. Planes, F. Tinti, C. Turan, M.N. Venugopal, H. Weber, D. Blohm, Identifying fishes through DNA barcodes and Microarrays. *PLoS One*, 5 (2010) e12620.

63. H. Zhang, Y. Zhang, T. Gao, P. Li, H. Xu, Genetic identification of two species of *Pleuronichthys* by DNA barcoding. *Chinese J. Oceanology Limnology*, 29 (2011) 967-972.

64. D.M. Cawthorn, H.A. Steinman, R. Corli Witthuhn, Establishment of a mitochondrial DNA sequence database for the identification of fish species commercially available in South Africa. *Mol. Ecol. Resour*., 11 (2011) 979-991.

65. J. Zhang, Species identification of marine fishes in China with DNA barcoding. *Evid. Based Complement Alternat.* Med. 978253. [Epub 2011 May 11].

66. E. Naro-Maciel, B. Reid, K.E. Holmes, D.R. Brumbaugh, M. Martin, R. DeSalle, Mitochondrial DNA sequence variation in spiny lobsters: population expansion, panmixia, and divergence. *Marine Biology*, 158 (2011) 2027-2041.

67. Liu J., Li Q., Kong L., Yu H., Zheng X. Identifying the true oysters (Bivalvia: Ostreidae) with mitochondrial phylogeny and distance-based DNA barcoding. *Mol. Ecology Res*., 11 (2011) 820-830.

UNCORRECTED PROOF

Table 1. Commercial samples analyzed by the multi-locus DNA barcoding approach developed in this study.

No.	Product description	Origin	Species declared in the label	Organism	Family	Processing treatments
16	Blue shark	Pacific Ocean, FAO 71	*Prionace glauca*	F	Carcharhinidae	Frozen fillet
15	Atlantic herring	n.a.	*Clupea harengus*	F	Clupeidae	Smoked, vacuum packaged
33	European anchovy	n.a.	*Engraulis encrasicolus*	F	Engraulidae	Brine, canned in vegetal oil
34	Atlantic cod	n.a.	*Gadus morhua**	F	Gadidae	Raw fillet
24	Pacific cod	n.a.	*Gadus macrocephalus*	F	Gadidae	Dried salted (baccalà)
53	Mako shark	n.a.	*Isurus oxyrhincus*	F	Lamnidae	Frozen fillet
9	Nile perch	n.a.	*Lates niloticus*	F	Latidae	Frozen fillet
27	Nile perch	Victoria lake, Africa	*Lates niloticus**	F	Latidae	Raw fillet
21	Angler	n.a.	*Lophius piscatorius*	F	Lophiidae	Raw fillet
3	South Pacific hake	South-West Pacific Ocean and Atlantic Ocean	*Merluccius gayi/productus*	F	Merlucciidae	Frozen, pre-cooked
5	Atlantic hake	South-East Atlantic Ocean	*Merluccius hubbsi*	F	Merlucciidae	Frozen fillet
8	Scarlet snapper	SouthAfrica Ocean and Indian Ocean	*Merluccius capensis/paradoxus*	F	Merlucciidae	Frozen, pre-cooked
6	Patagonian grenadier	Pacific Ocean	*Macruronus magellanicus*	F	Merlucciidae	Frozen fillet
29	Striped catfish	n.a.	*Pangasius hypophthalmus**	F	Pangasidae	Raw fillet
50	Striped catfish	n.a.	*Pangasius hypophthalmus*	F	Pangasidae	Raw fillet
13	Turbot	South-East Atlantic Ocean	*Paralichthys isosceles*	F	Paralichthydae	Frozen fillet
28	European perch	n.a.	*Perca fluviatilis**	F	Percidae	Raw fillet
4	European plaice	North-East Atlantic Ocean	*Pleuronectes platessa*	F	Pleuronectidae	Frozen fillet
51	European plaice	n.a.	*Pleuronectes platessa*	F	Pleuronectidae	Raw fillet
12	Rainbow trout	Farmed in Italy	*Oncorhynchus mykiss*	F	Salmonidae	Smoked, vacuum packaged
19	Atlantic salmon	n.a.	*Salmo salar*	F	Salmonidae	Smoked, vacuum packaged
30	Yellow-fin tuna	n.a.	*Thunnus albacares**	F	Scombridae	Raw fillet
36	Tuna chunks sashimi	n.a.	*Thunnus albacares*	F	Scombridae	Raw fillet
35	Yellow-fin tuna fillets	n.a.	*Thunnus albacares*	F	Scombridae	Raw fillet
31	Tuna	n.a.	*Thunnus albacares*	F	Scombridae	Carpaccio
23	Malabar grouper	n.a.	*Epinephelus malabaricus*	F	Serranidae	Raw fillet
22	Common sole	n.a.	*Solea solea*	F	Soleidae	Raw fillet
17	Smoked swordfish	n.a.	*Xiphias gladius*	F	Xiphiidae	Smoked, vacuum packaged
32	Swordfish carpaccio	n.a.	*Xiphias gladius*	F	Xiphiidae	Carpaccio
37	Swordfish fillets	n.a.	*Xiphias gladius**	F	Xiphiidae	Raw fillet
2	Greenshell mussel	Pacific Ocean	*Perna canaliculus*	M	Mytilidae	Frozen
25	Common octopus	n.a.	*Octopus vulgaris*	M	Octopodidae	Raw
52	Jumbo squid	n.a.	*Dosidicus gigas*	M	Ommastrephidae	Raw
18	Great Atlantic scallop	North-East Atlantic Ocean	*Pecten maximus*	M	Pectinidae	Frozen
11	Northern red shrimp	n.a.	*Pandalus borealis*	C	Pandalidae	Frozen
7	Pink prawn	Pacific Ocean and Indian Ocean	*Metapenaeus affinis/monoceros*	C	Penaeidae	Frozen
14	Whiteleg shrimp	n.a.	*Penaeus Vannamei*	C	Penaeidae	Frozen

*, only the common name is indicated in the label, but the scientific name is deducible in agreement with the Italian Ministerial Decree of the 14/01/2005.
n.a., not available

UNCORRECTED PROOF

Table 2. The forward and reverse primers that were designed for each of the selected barcode gene regions.

Marker	Primer name	Primer sequence (5'-3')	Ta (°C)	References
cox1	FishF2	TCGACTAATCATAAAGATATCGGCAC	60	Ward *et al*. (2005)
	FishR2	ACTTCAGGGTGACCGAAGAATCAGAA	60	Ward *et al*. (2005)
	LCO1490	GGTCAACAAATCATAAAGATATTGG	60	Folmer *et al*. (1994)
	HCO2198	TAAACTTCAGGGTGACCAAAAAATCA	60	Folmer *et al*. (1994)
16S-rDNA	16Sar-5′	CGCCTGTTTATCAAAAACAT	55	Palumbi (1996)
	16Sbr-3′	CCGGTCTGAACTCAGATCACGT	55	Palumbi (1996)
cob	GLUDG-l	TGACTTGAARAACCAYCGTTG	60	Palumbi (1996)
	CB3-H	GGCAAATAGGAARTATCATTC	60	Palumbi *et al.* (1991)

UNCORRECTED PROOF

Table 3. The local BLAST results from query sequences that were derived from the analyzed commercial seafood products

No.	Species declared in the label	cox1 GenBank	cox1 GenBank/Blast	cox1 E-value	cox1 Max ID	cox1 BOLD	cox1 BOLD/HMM	cox1 Similarity	cox1 Tree based identification**	16S-rDNA GenBank	16S-rDNA GenBank/Blast	16S-rDNA E-value	16S-rDNA Max ID	cob GenBank	cob GenBank/Blast	cob E-value	cob Max ID
2	*Perna canaliculus*	yes	*Perna canaliculus*	00:00	99	yes	*Perna canaliculus*	98.79	*Perna canaliculus*	yes	*Perna canaliculus*	8.00E-101	100	n.a.	n.d.		
3	*Merluccius gayi/productus*	yes	*Merluccius hubbsi*	00:00	99	yes	*Merluccius hubbsi*	99.5*	*Merluccius hubbsi*	yes	*Merluccius hubbsi*	00:00	100	yes	*Merluccius productus*	00:00	96
4	*Pleuronectes platessa*	yes	*Pleuronectes platessa*	00:00	99	yes	*Pleuronectes platessa*	99.67	*Pleuronectes platessa*	yes	*Pleuronectes platessa* (3)	00:00	100	yes	*Pleuronectes platessa*	00:00	99
5	*Merluccius hubbsi*	yes	*Merluccius hubbsi*	00:00	100	yes	*Merluccius hubbsi*	100*	*Merluccius hubbsi*	yes	*Merluccius hubbsi*	00:00	100	yes	*Merluccius hubbsi*	00:00	98
6	*Macruronus magellanicus*	yes	*Macruronus magellanicus*	00:00	98	yes	*Macruronus novaezelandiae*	99.54*	*Macruronus* spp.	yes	*Macruronus novaezelandiae*	00:00	100	yes	*Macruronus magellanicus* (11)	00:00	100
7	*Metapenaeus affinis/monoceros*	yes/n.a.	n.d.			n.a.	n.d.			n.a./yes	*Litopenaeus vannamei*	00:00	100	n.a.	*Macruronus spp* (11)	00:00	100
8	*Merluccius capensis/paradoxus*	yes/yes	*Merluccius paradoxus*	00:00	92	yes	*Merluccius paradoxus*	100*	*Merluccius paradoxus*	yes/yes	*Merluccius paradoxus*	00:00	100	yes/yes	*Merluccius paradoxus*	00:00	99
9	*Lates niloticus*	yes	*Lates niloticus*	00:00	100	yes	*Lates niloticus*	100*	*Lates niloticus*	yes	*Lates niloticus*	00:00	99	yes	*Merluccius hubbsi*	00:00	98
11	*Pandalus borealis*	n.a.	n.d.			n.a.	n.d.			yes	*Pandalus borealis*	00:00	97	n.a.	*Oncorhynchus mykiss*	00:00	100
12	*Oncorhynchus mykiss*	yes	*Oncorhynchus mykiss*	00:00	100	yes	*Oncorhynchus mykiss*	100	*Oncorhynchus mykiss*	yes	*Oncorhynchus mykiss*	00:00	99	yes	*Oncorhynchus mykiss*	00:00	100
13	*Paralichthys isosceles*	yes	*Xystreurys rasile*	00:00	99	yes	*Xystreurys rasile*	99.51*	*Xystreurys rasile*	n.a./yes	*Xystreurys liolepis*	00:00	96	n.a.	*Oncorhynchus mykiss*	00:00	100
14	*Penaeus vannamei*	yes	*Xystreurys rasile*	00:00	100	yes	*Xystreurys rasile*	100*	*Xystreurys rasile*	yes	*Penaeus Vannamei*	00:00	100	yes	*Oncorhynchus mykiss*	00:00	100
15	*Clupea harengus*	yes	*Clupea harengus*	00:00	100	yes	*Clupea harengus*	100*	*Clupea harengus*	yes	*Clupea harengus*	00:00	100	yes	*Oncorhynchus mykiss*	00:00	100
16	*Prionace glauca*	yes	*Prionace glauca*	00:00	100	yes	*Prionace glauca*	100*	*Prionace glauca*	yes	*Prionace glauca*	00:00	100	yes	*Prionace glauca*	00:00	100
17	*Xiphias gladius*	yes	*Xiphias gladius*	00:00	100	yes	*Xiphias gladius*	100*	*Xiphias gladius*	yes	*Xiphias gladius*	00:00	99	yes	*Xiphias gladius*	00:00	99
18	*Pecten maximus*	n.a.	n.d.			yes	n.d.			yes	*Pecten maximus*	00:00	99	n.a.	n.d.		
19	*Salmo salar*	yes	*Salmo salar*	00:00	99	yes	*Salmo salar*	100*	*Salmo salar*	yes	*Salmo salar*	00:00	100	yes	n.d.		
21	*Lophius piscatorius*	yes	*Lophius piscatorius*	00:00	100	yes	*Lophius piscatorius*	100*	*Lophius piscatorius*	yes	*Lophius piscatorius*	00:00	98	yes	*Solea solea*	4.00E-171	95
22	*Solea vulgaris/solea*	yes	*Solea solea*	00:00	99	yes	*Solea solea*	99.84	*Solea solea*	yes	*Solea solea*	00:00	99	yes	*Solea solea*	00:00	100
23	*Epinephelus malabaricus*	yes	*Epinephelus areolatus*	00:00	98	yes	*Epinephelus areolatus*	98.71	*Epinephelus areolatus*	yes	n.d.			yes	n.d.		
24	*Gadus macrocephalus*	yes	*Gadus macrocephalus*	00:00	100	yes	*Gadus ogac*	100*	*Gadus ogac*	yes	*Gadus macrocephalus* (4)	00:00	100	yes	*Gadus macrocephalus* (12)	00:00	99
25	*Octopus vulgaris*	n.a.	*Amphioctopus marginatus*	00:00	99	yes	*Amphioctopus marginatus*	100	*Amphioctopus marginatus*	yes	*Octopus spp.* (5)	00:00	99	yes	n.d.		
27	*Lates niloticus*	yes	*Lates niloticus*	00:00	100	yes	*Lates niloticus*	100*	*Lates niloticus*	yes	*Lates niloticus*	3.00E-133	95	yes	*Chelidonichthys lucernus*	00:00	96
28	*Perca fluviatilis*	yes	*Paralichthys* spp.	00:00	88	yes	*Paralichthys patagonicus*	100*	*Paralichthys patagonicus*	yes	*Paralichthys patagonicus*	00:00	100	yes	*Paralichthys olivaceus*	00:00	88
29	*Pangasius hypophthalmus*	yes	*Pangasius hypophthalmus*	00:00	100	yes	*Pangasius hypophthalmus*	100*	*Pangasius hypophthalmus*	yes	*Pangasius hypophthalmus* (6)	00:00	99	yes	*Pangasius hypophthalmus* (13)	00:00	99
30	*Thunnus albacares*	yes	*Thunnus albacares*	00:00	100	yes	*Thunnus obesus*	100*	*Thunnus* spp.	yes	*Thunnus albacares* (7)	00:00	99	yes	*Thunnus albacares*	00:00	100
31	*Thunnus albacares*	yes	*Thunnus albacares*	00:00	100	yes	*Thunnus* spp. (1)	100*	*Thunnus* spp.	yes	*Thunnus albacares* (7)	00:00	99	yes	*Thunnus albacares*	00:00	99
32	*Xiphias gladius*	yes	*Xiphias gladius*	00:00	99	yes	*Xiphias gladius*	100*	*Xiphias gladius*	yes	*Xiphias gladius*	00:00	99	yes	*Xiphias gladius*	00:00	99
	00:00	97	33 *Engraulis encrasicolus*	yes	*Thunnus albacares*	00:00	100	yes	*Thunnus* spp. (2)	99.84*	*Thunnus* spp.	yes	*Engraulis encrasicolus* (8)	00:00	99	yes	*Thunnus albacares*
	00:00	99	34 *Gadus morhua*	yes	*Gadus morhua*	00:00	98	yes	*Gadus morhua*	99.84*	*Gadus morhua*	yes	*Gadus morhua*	00:00	99	yes	*Gadus morhua*
	00:00	100	35 *Thunnus albacares*	yes	*Thunnus albacares*	00:00	100	yes	*Thunnus* spp. (2)	100*	*Thunnus* spp.	yes	*Thunnus albacares* (7)	00:00	99	yes	*Thunnus albacares*
	00:00	99	36 *Thunnus albacares*	yes	*Thunnus albacares*	00:00	100	yes	*Thunnus* spp. (2)	100*	*Thunnus* spp.	yes	*Thunnus albacares* (9)	00:00	99	yes	*Thunnus albacares*
	00:00	99	37 *Xiphias gladius*	yes	*Xiphias gladius*	00:00	100	yes	*Xiphias gladius*	100*	*Xiphias gladius*	yes	*Xiphias gladius*	00:00	99	yes	*Xiphias gladius*
almus (13)	00:00	99	50 *Pangasius hypophthalmus*	yes	*Pangasius hypophthalmus*	00:00	100	yes	*Pangasius hypophthalmus*	100*	*Pangasius hypophthalmus*	yes	*Pangasius hypophthalmus*	00:00	100	yes	*Pangasius hypophtl*
	00:00	98	51 *Pleuronectes platessa*	yes	*Pleuronectes platessa*	00:00	100	yes	*Pleuronectes platessa*	100	*Pleuronectes platessa*	yes	*Pleuronectes platessa* (10)	00:00	99	yes	*Pangasius spp.* (13)
			52 *Dosidicus gigas*	yes	*Dosidicus gigas*	00:00	99	yes	*Dosidicus gigas*	99.83	*Dosidicus gigas*	yes	*Dosidicus gigas*	00:00	98	yes	n.d.
			53 *Isurus oxyrhincus*	yes	*Isurus oxyrhincus*	00:00	99	yes	*Isurus oxyrhincus*	99.84	*Isurus oxyrhincus*	yes	*Isurus oxyrhincus*	00:00	98	yes	n.d.

n.d., not determined.

n.a., no sequence of the labelled species is available in the nucleotide databases.

**, the threshold divergence value to distinguish different species is 1%, speciemens with divergence value minor than 1% cluster toghether.

UNCORRECTED PROOF

Figure legends

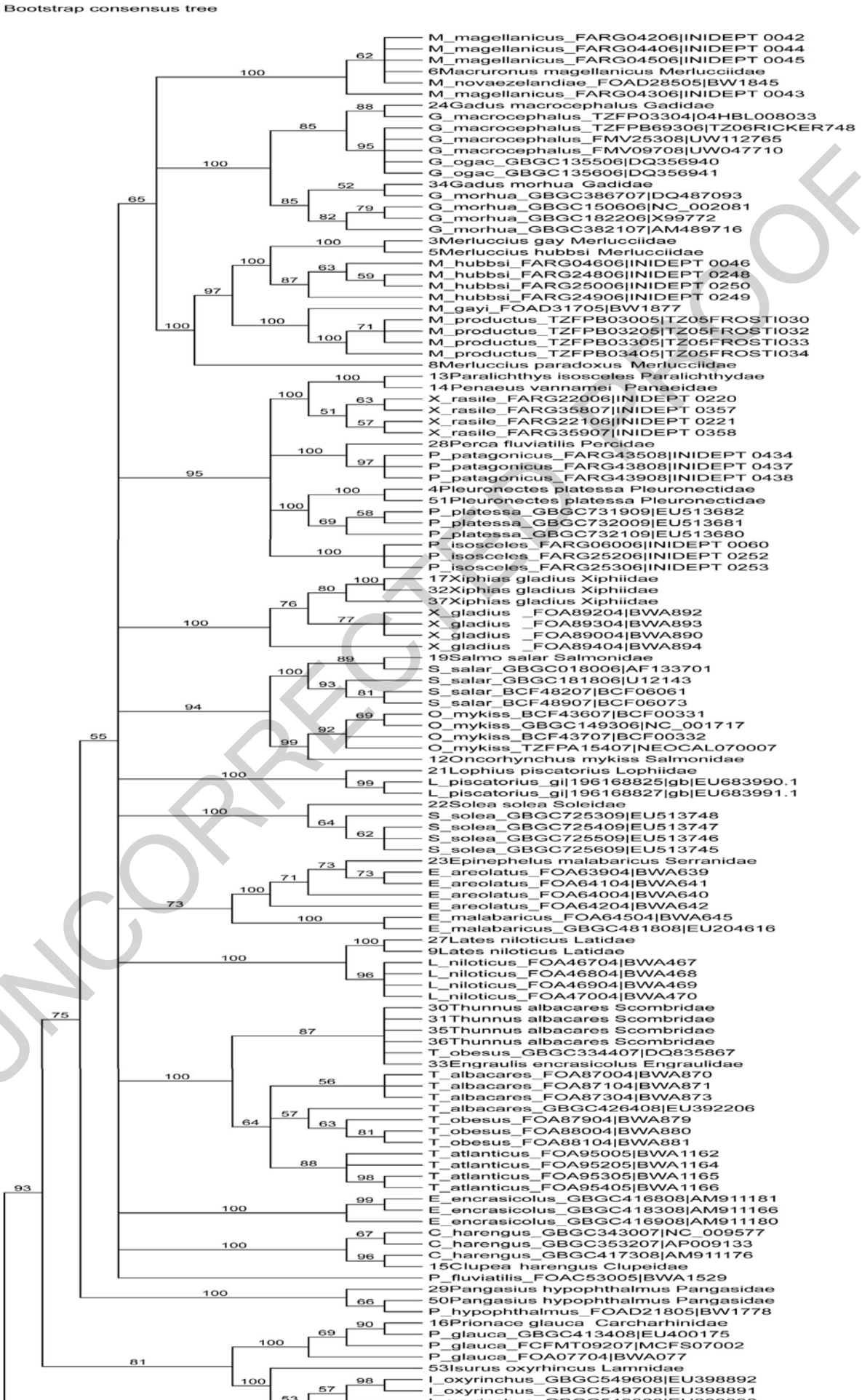
Figure 1. A Neighbor-Joining tree for species that correspond to the collected specimens, as well as the experimental *cox1* sequences that were obtained. The NJ tree was constructed from 104 mitochondrial *cox1* sequences that are available in the BOLD and GenBank databases. The numbers above the nodes represent bootstrap analysis after 1,000 replicates.

Supplementary materials

Figure 1S. The proportion of sequence accessions in GenBank and/or BOLD databases that are related to *cox1, nad1, cox2, cox3, 16S-rDNA* or *cob*.

Table 1S. The accession numbers of the sequences that were retrieved from the BOLD and GenBank databases and were used to construct the Neighbor-Joining trees.

UNCORRECTED PROOF



UNCORRECTED PROOF

Supplementary materials

Figure 1S. The proportion of sequence accessions in GenBank and/or BOLD databases that are related to *cox1*, *nad1*, *cox2*, *cox3*, *16S-rDNA* or *cob*.

Table 1S. The accession numbers of the sequences that were retrieved from the BOLD and GenBank databases and were used to construct the Neighbor-Joining trees.

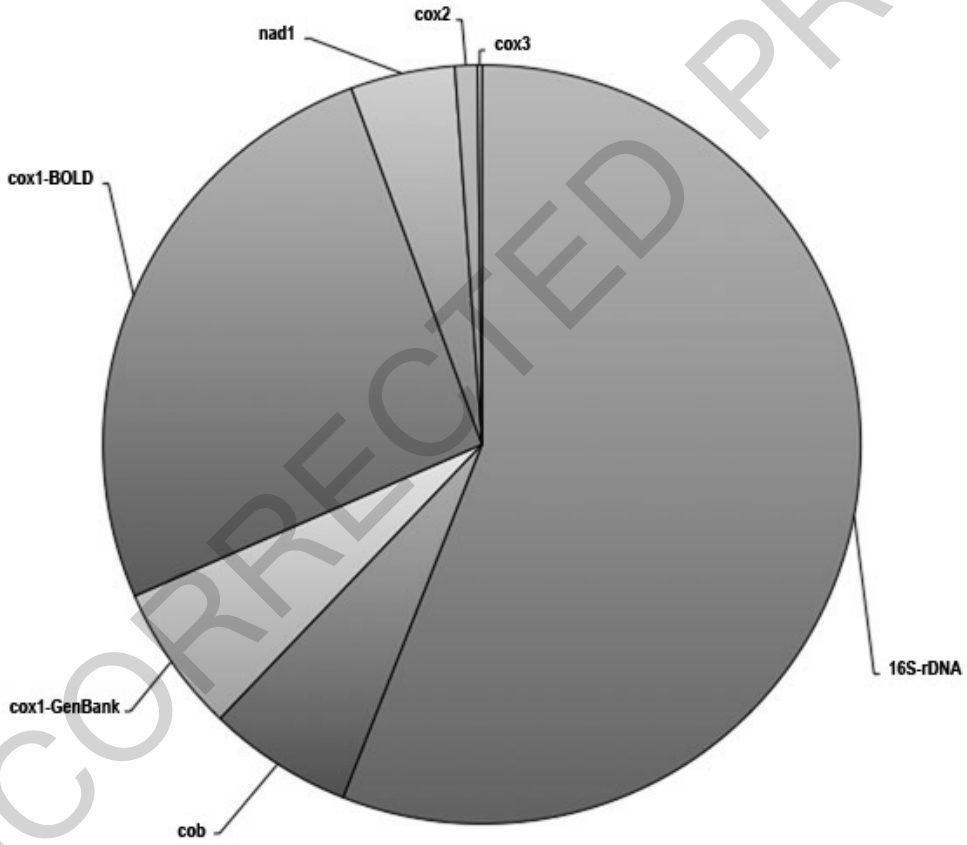


Fig. 1s

UNCORRECTED PROOF

Table 1 s

GBGC426408|EU392206; GBGC326507|DQ835949; GBGC326807|DQ835945; GBGC326207|DQ835953; GBGC426108|EU418252; GBGC326707|DQ835946; GBGC326607|DQ835948; GBGC326407|DQ835951; GBGC326307|DQ835952; GBGC326107|DQ835954; GBGC326007|DQ835955; GBGC325407|DQ835947; GBGC325307|DQ835950; WLIND46107|WLM461; FOA87204|BWA872; WLIND45907|WLM459; FOA87104|BWA871; FOA87004|BWA870; WLIND45707|WLM457; FOA86904|BWA869; FOA95405|BWA1166; FOA95305|BWA1165; FOA95205|BWA1164; FOA95005|BWA1162; FOA88104|BWA881; FOA88004|BWA880; FOA87904|BWA879; GBGC334407|DQ835867; FOA88304|BWA883; GBGC334107|DQ835870; GBGC334207|DQ835869; FOA88204|BWA882; GBGC334307|DQ835868; GBGC334007|DQ835871; FOA88804|BWA888; FOA88604|BWA886; FOA88904|BWA889; FOA88504|BWA885; FOA88704|BWA887; GBGC333407|DQ835877; GBGC333907|DQ835872; GBGC333507|DQ835876; GBGC333807|DQ835873; GBGC333607|DQ835875; GBGC080306|AY302574; GBGC165606|NC_004901; GBGC333207|DQ835879; GBGC333307|DQ835878; GBGC333707|DQ835874; FOA94805|BWA1160; FOA94705|BWA1159; FOA94605|BWA1158; FOA94505|BWA1157; GBGC004906|AB097669; GBGC338607|DQ835824; GBGC338807|DQ835822; GBGC339207|DQ835818; GBGC338707|DQ835823; GBGC339107|DQ835819; FOA86404|BWA864; GBGC339007|DQ835820; GBGC166806|NC_005317; GBGC005206|AB101291; GBGC338907|DQ835821; FOA86704|BWA867; FOA86504|BWA865; FOA86804|BWA868; FOA86604|BWA866; FOA88404|BWA884; FOA94405|BWA1156; FOA94205|BWA1154; GBGC181506|NC_008455; GBGC008706|AB185022; FOA94305|BWA1155FOA94105|BWA1153; FOA87804|BWA878; FOA87504|BWA875; FOA87604|BWA876; FOA87404|BWA874; FOA87704|BWA877; FOA95005|BWA1162; FOA95205|BWA1164; FOA95405|BWA1166; FOA95305|BWA1165; FOA89004|BWA890; FOA89404|BWA894; FOA89304|BWA893; FOA89204|BWA892; FOAD31705|BW1877; gi|166898013|gEU271893.1; TZFPB03405|TZ05FROSTI034; TZFPB03305|TZ05FROSTI033; TZFPB03205|TZ05FROSTI032; TZFPB03005|TZ05FROSTI030; FARG04606|INIDEPT0046; FARG25006|INIDEPT0250; FARG24906|INIDEPT0249; FARG24806|INIDEPT0248; gi|154761023|gb|EU074460.1; FOAD28505|BW1845; FARG04506|INIDEPT0045; |FARG04406|INIDEPT0044; |FARG04306|INIDEPT0043; FARG04206|INIDEPT0042; gi|154761019|gb|EU074458.1; gi|154761021|gb|EU074459.1; FARG04106|INIDEPT0041; gi|154761017|gb|EU074457.1; gi|154761015|gb|EU074456.1; gi|148374017|gb|EF609405.1; BCF43707|BCF00332; BCF43607|BCF00331; GBGC149306|NC_001717; TZFPA15407|NEOCAL070007; GBGC018006|AF133701; BCF48207|BCF06061; GBGC181806|U12143; BCF48907|BCF06073; FOA47004|BWA470; FOA46904|BWA469; FOA46804|BWA468; FOA46704|BWA467; GBGC382107|AM489716; GBGC386707|DQ487093; GBGC150606|NC_002081; GBGC182206|X99772; GBGC135406|DQ356938; gi|209366407|gb|FJ164619.1; gi|209366403|gb|FJ164617.1; GBGC135306|DQ356937; gi|209366405|gb|FJ164618.1; GBGC135606|DQ356941; GBGC135506|DQ356940; gi|124377051:54446994; GBGC732109|EU513680; GBGC732009|EU513681; GBGC731909|EU513682; FOAD21805|BW1778 ; FARG25306|INIDEPT 0253; FARG25206|INIDEPT 0252; FARG06006|INIDEPT0060; GBGC417308|AM911176; GBGC343007|NC_009577; GBGC353207|AP009133; FCFMT09207|MCFS07002; GBGC413408|EU400175; FOA07704|BWA077; GBGC725609|EU513745; GBGC725509|EU513746; GBGC725409|EU513747; GBGC725309|EU513748; GBGC481808|EU204616; FOA64504|BWA645; FOAC53005|BWA1529; GBGC418308|AM91116; GBGC416908|AM911180; GBGC416808|AM911181; GBGC549908|EU398889; GBGC549808|EU398890; GBGC549708|EU398891; GBGC549608|EU398892; FARG35907|INIDEPT0358; FARG35807|INIDEPT 0357; FARG22106|INIDEPT 0221; FARG22006|INIDEPT 0220; FOA64204|BWA642; FOA64104|BWA641; FOA64004|BWA640;

UNCORRECTED PROOF

FOA63904|BWA639; FARG43508|INIDEPT|0434; FARG43908|INIDEPT0438; FARG43808|INIDEPT0437; gi|196168825|gb|EU683990.1; gi|196168827|gb|EU683991.1; GBCPH77709|NC_009734; GBCPH41307|EU068697; GBCPH80109|FJ153075; GBCPH80209|FJ153074; GBCPH000106|AB052253; GBCPH70007|DQ683211; GBCPH70107|DQ683210; GBCPH70307|DQ683208; GBCMD96307|DQ534543; |GBCMD96207|NC_009626; |GBMLB172106|DQ343604; GBMLB172206|DQ343605.

UNCORRECTED PROOF