Food Control 60 (2016) 394–400

Contents lists available at ScienceDirect

Food Control

journal homepage: www.elsevier.com/locate/foodcont

Molecular testing on sardines and rulings on the authenticity and nutritional value of marketed fishes: An experience report in the state of Rio de Janeiro, Brazil

Raphael Leonardo [a], Raquel Soares Casaes Nunes [a], Maria Lúcia G. Monteiro [b], Carlos Adam Conte-Junior [b], Eduardo Mere Del Aguila [a], Vânia M.F. Paschoalin [a, *]

[a] Universidade Federal do Rio de Janeiro, Instituto de Química, Av. Athos da Silveira 149, Cidade Universitária, 21949-909, Rio de Janeiro, RJ, Brazil
[b] Universidade Federal Fluminense, Faculdade de Medicina Veterinária, Departmento de Tecnologia de Alimentos, Rua Vital Brazil Filho 64, 24230340, Niterói, RJ, Brazil

ARTICLE INFO

Article history:
Received 23 April 2015
Received in revised form
3 August 2015
Accepted 4 August 2015
Available online 10 August 2015

Keywords:
Sardines
PCR-RFLP
Sequencing
Authenticity
Proximate composition
Fraud

ABSTRACT

Members of the *Clupeidae* family are fish sold and consumed worldwide. Sardine species are the main representatives of this family, due to their nutritional qualities and contribution to the delivery of a wide variety of products and byproducts. The authenticity of seafood, due to high costs of analytical methods and sometimes limited availability, is a problem in fish trade. The authenticity of sardines marketed in the State of Rio de Janeiro, was evaluated by sensitive and unequivocal PCR-based techniques, such as RFLP and DNA sequencing. The *CYTB* mitochondrial gene was used for the screening of 170 sardine samples collected from markets, fish stores, street markets and canned samples from sardine factories. Sixty per cent of the collected fish were identified as *Sardinella aurita* (18.8%), *Sardina pilchardus* (25.9%), *Sardinops sagax* (2.9%), *Sardinops caeruleus* (0.6%) and *Opisthonema oglinum* (11.8%). The fraud samples were identified as *Clupea harengus* (4.7%), *Brevortia aurea* (21.2%), *Centengraulis edentulus* (6.5%) and *Scomber japonicus* (7.6%). *Sardinella brasiliensis*, considered the most abundant species on the south-eastern coast, was not found among the collected samples. The phylogenetic analysis of the marketed sardines showed that *S. sagax*, *O. oglinum*, *S. pilchardus* and *S. caeruleus* and fraud species as *S. japonicus* and *B. aurea* were clustered with a genetic distance of 0.1. A secondary cluster grouped only fraud species, such as *C. harengus* and *C. edentulus*, with a genetic distance <0.1. *S. aurita* appeared isolated with a genetic divergence >0.6. Our study also observed that theses frauds negatively changed the nutritional value of the product. Fraudulent species, such as *C. harengus*, *B. aurea*, *H. clupeola* and *C.* edentulus have lower protein content than authentic sardines. According to the Brazilian legislation with regard to labeling requirements, 40% of the samples were considered frauds by unequivocal molecular analysis.

1. Introduction

The world fish production increased from 9.9 tons in 1960 to 19.2 tons in 2012, mainly due to the increased demand for healthy and highly nutritious food, as well as due to the increase of purchasing power of developing countries (FAO, 2014).

Seafood substitution is a form of economic deception and has been prohibited through both domestic and international regulatory labeling laws (Martinez, James, & Loreal, 2005; Moretti, Turchini, Bellagamba, & Caprino, 2003; USFDA, 2006). Identification of fish species is also important to ascertain commercial frauds, mainly performed by replacing valuable species with others of lower value, especially in very transformed foodstuffs (breaded fillets) (Ardura, Planes, & Garcia Vazquez, 2011).

Fish belonging to the Clupeidae family comprise actinopterygii fish of high economic interest to the seafood industry, sold raw or processed as a wide variety of products and by-products. The sardine is the main species of this family (Herrero, Lago, Vieites, & Espinera, 2011) marketed fresh in fishmonger shops, fishing stores and street markets, with the canned product being considered the main commercial product (Granada, Flick, & Roy, 2012). Sardines

* Corresponding author. Instituto de Química, Universidade Federal do Rio de Janeiro, Av. Athos da Silveira Ramos 149 – Centro de Tecnologia – Bloco A, sala 545, Cidade Universitária, 21941-909 Rio de Janeiro, RJ, Brazil.
E-mail address: paschv@iq.ufrj.br (V.M.F. Paschoalin).

http://dx.doi.org/10.1016/j.foodcont.2015.08.004
0956-7135/

(*Sardina* and *Sardinops*) are distributed in most of the temperate boundary current systems and *Sardinella* are found in subtropical and tropical zones (Ganias, 2015). Species recognition is essential for studies involving the extinction of some genetic variants and subsequent loss of intra-specific diversity, with unpredictable effects on species biodiversity (Ardura et al., 2011). Traditionally, fish identification is usually based on morphological and histological characteristics, but, primarily, external morphological characters. However, even within species, fish may exhibit morphological plasticity (Bottero, 2011).

Additionally, species recognition is also important in order to guarantee the authenticity of the fish species and fish products on sale in several types of markets. However, when morphological characteristics, such as shape, size or fish appearance, are obliterated during the processing phase, fish species identification becomes jeopardized and the fish, in fact, may not even be recognized by consumers. The growth in the marketing of filleted fish, both fresh or frozen, processed or not, is not always accompanied by captive production techniques of the species, which enables frauds to occur, by swapping high-value fish for other, low-cost species, that may present inadequate nutritional characteristics (Infante, Blanco, Zuasti, Crespo, & Manchado, 2007).

Currently, fish identification (certification) is needed in order to detect changes or frauds in fish trading in Brazil and in the international market, to avoid substitution of certain fish for higher availability and lower commercial value products (ECLAC, 2006). In Brazil, the official regulation determines that the product label describe the marketed fish species. The authenticity of marketed fish species has become a major challenge for the fish processing trade and fishing industry (Rasmussen and Morrissey, 2011).

Fish popularly called sardines are highly appreciated in Brazil, due to their low-cost and nutritional value, and because they are considered fish that display low heavy metal retention, since they are short lived (Tarley, Coltro, Matsushita, & de Souza, 2001), when compared with tuna, bonito and chub mackerel (Mol, 2011).

The term “sardine” in the present study shall be used exclusively for *Sardinella brasiliensis*, *Sardinella aurita* and *Sardina pilchardus*; the term “Pacific sardine” shall be used exclusively for *Sardinops sagax*, *Sardinops melanostictus*, *Sardinops neopilchardus*, *Sardinops caeruleus*, and the term “slab sardine” shall be used exclusively for *Opisthonema oglinum*. All these species are considered authentic by the current Brazilian regulation. The legislation is also applicable to canned sardines, which are intended for national and/or international trade, and the marketing of any other kind of species from the Clupeidae family under the name “sardine” is considered fraudulent as established by the technical resolution from the Ministry of Agriculture, Livestock and Supply (MAPA) Portaria 406, article 3 from August 10, 2010) (MAPA, 2010).

The municipalities of Rio de Janeiro and Niterói, located on the coastline of the state of Rio de Janeiro, are the two largest metropolitan areas of the region, where the fish distributed in the state are marketed. The availability of each sardine species is not known, and although there is a general suspicion regarding fraudulent marketing of Clupeidae family fish species, especially during the off-season between November and February and June–July, no scientific evidence is available.

The official methodologies to lend authenticity to fish species are based on the analysis of fish muscle proteins (AOAC, 2012). Although most of these methods show reliable results, they are not suitable for routine analyses in some cases, such as to differentiate between closely related species or products subjected to heat processing, since this process results in altered biological activity due to protein denaturation (Rasmussen & Morrissey, 2011). Molecular tests, such as PCR, are currently available. These methods are able to discriminate sardines from other members of the Clupeidae family (Herrero et al., 2011). PCR-based methods, when associated to nucleotide sequencing, allow for the unequivocal identification of the fish species present in food matrices, even after several processing steps, such as industrial sterilization (Teletchea, 2009). CYT B or COX1 are commonly used as target sequences for fish identification. Recently, a study developed an useful complementary approach to *COX1* barcode fragment sequencing, leveraging the ability to use the extensive fish barcoding sequence databases for primer development and restriction enzyme selection (Mueller et al., 2015) and the development of real-time or quantitative PCR (qPCR) methods in the field of fish DNA detection are focused on the use of mitochondrial DNA sequences as targets (Prado, Boix, & von Holst, 2013) or nuclear (Hird et al., 2011).

The aim of the present study was to identify the fish species names *“sardine”* traded in markets, fishing-warehouses, street markets and in industrial plants that produce canned sardines in the state of Rio de Janeiro, Brazil, by PCR tests targeting the *CYTB* associated to RFLP assays and nucleotide sequencing. The frequency of frauds in each market was calculated and the nutritional values of true and false sardines were compared.

2. Material and methods

2.1. Specimen collection

One hundred and seventy (170) samples of whole or eviscerated fish sold as sardines (fresh or frozen) were collected at different markets, fishing warehouses, street markets and in an industrial plant that produces canned sardines. All fishes were traded in the municipalities of Rio de Janeiro and Niterói, Brazil. Reference fishes were identified based on their external morphological characteristics by an experienced veterinarian. The collected samples were stored in polyethylene bags, duly sealed to ensure sample inviolability. The samples were transported in a portable cooler containing ice to the laboratory and stored at −80 °C until processing. Fishes collected at the industrial plant were sampled just before canning.

2.2. DNA extraction

Mitochondrial DNA templates for PCR tests were prepared from 30 mg of sardine muscle tissues using an automatic Maxwell® 16 extractor (Promega, Madison, WI, USA), with the Maxwell® 16 Tissue DNA Purification Kit (Promega, Madison, WI, USA) following the manufacturer's instructions. DNA was quantified using the Qubit fluorimeter (Invitrogen™, Grand Island, New York, USA) and Qubit assays kit.

2.3. Amplification of the mitochondrial cytochrome b (CYTB) gene by PCR tests

A 147 bp fragment of *CYTB* gene was amplified using a primer pair (C-CB285 dF, CGCCCACATTGGNCGAGG and C-CB431R, GTGGCCCCTCAGAAGGACATTTGGCC) (Jérôme, Lemaire, Bautista, Fleurence, & Etienne, 2003). PCR mixtures contained 100 ng of DNA template, 10 pmoles of each primer (Integrated DNA Technologies Inc., Coralville, Iowa, USA), 1.5 mM $MgCl_2$ (Invitrogen™, Grand Island, New York, USA), 0.2 mM each dNTP (Fermentas, CA, USA), 1× of buffer (Invitrogen™, Grand Island, New York, USA), 1.5 U of Taq DNA polymerase (Invitrogen™, Grand Island, New York, USA) and nuclease-free sterile water in a final reaction volume of 50 μL.

DNA amplification was performed using the Veriti® 96-Well Thermal Cycler (Applied Biosystems Inc., Foster City, California, USA) under the following conditions: 95 °C for 5 min followed by 35 cycles of 95 °C for 30 s, 55 °C for 40 s and 72 °C for 40 s and a step

of 72 °C for 7 min. Amplicons were analyzed on 2% agarose gels. The electrophoresis was run with 1× TAE buffer (Tris-Acetic acid-EDTA) at 100 V, 200 mA for 1 h and visualized by staining with GelRed™ Nucleid Acid Gel Stain (Biotium, Hayward, CA, USA). Gels were documented by using the Bio-Imaging System (BioAmerica Inc., TelAviv, Israel).

2.4. PCR-RLFP and DNA sequencing

The 147 bp amplicons were purified directly from the PCR mixture by the GFX PCR DNA and Gel Band purification kits (GE Healthcare Life Science Inc., Little Chalfont, Buckinghamshire, UK) following the manufacturer's instructions and quantified using the Quibit fluorimeter (Invitrogen™, Grand Island, New York, USA) and Qubit assays kit. Aliquots of the purified 147 bp amplicons were used for the RFLP and nucleotide sequencing analyses.

DNA digestion by *Mnl*1 and *Hinf*1 endonucleases (Thermo Fisher Scientific, Sunnyvale, CA, USA) was performed in 30 μL containing 200 ng of the purified 147 bp amplicon, 2 U of *Mnl*1 and 1 U of *Hinf*1 and 1× digestion Tango buffer (Thermo Fisher Scientific, Sunnyvale, CA, USA). The reaction was run at 37 °C for 2 h and then inactivated at 65 °C for 20 min. Restriction fragments were resolved on a 12% electrophoresis polyacrylamide gel (29 acrylamide: 1 bisacrylamide, 4.93 ml H2O, 1 ml TBE 10×, 70 μl APS 10%, 35 μl TEMED) run in 1× Tris-Borate-EDTA (TBE) buffer at 100 V, 200 mA for 2 h. Gels were stained by GelRed™ Nucleid Acid Gel Stain (Biotium, Hayward, CA, USA). The length of the fragments was estimated by comparing with a 10 bp DNA ladder (Promega, Madison, WI, USA). Reference fishes were verified by their RFLP profile.

Amplicons that generated distinct restriction digestions profiles, equal or different from reference fishes, were grouped and selected for sequencing by the Sanger method. DNA sequencing was performed using 3 ng of the purified amplicons, 4 μL of BigDye mix, 2 μL sequencing buffer 5×, 3.2 pmol/reaction of primer (C-CB285dF/C-CB431R primers) and 10 ng of each amplicon to a final volume of 20 μL with sterile ultra-pure water and this reaction was run at the same conditions of amplification for each primers. The PCR products were purified according to the BigDye™ X Terminator Purification kit protocol (Applied Biosystems, California, USA) and sequenced in a sequencer 3130 Genetic Analyzer (Applied Biosystems, California, USA).

2.5. Analysis of the nucleotide sequences using bioinformatics tools

The interpretation of the sequencing chromatograms verified the quality of sequences, accepting a quality Phred >20 and the similarities between those and the genome sequences at GenBank using the BLAST software (http://blast.ncbi.nlm.nih.org). The results were aligned using the ClustalW software and identified through the BOLD Identification System (www.boldsystems.org). Reference fishes identification were confirmed by nucleotide sequencing.

2.6. Phylogenetic analyses

Phylogenetic relationships among sardine samples were performed by the alignment of sequences in the Clustal X 2.0 software (Larkin et al., 2007). The phylogenetic trees were constructed using the software Mega 6.0 and UPGMA methods (Sneath & Sokal, 1973).

2.7. Proximate composition of distinct species

The proximate composition (moisture, protein, lipid and ash content) was determined according to AOAC recommendations (AOAC, 2012). Briefly, moisture was determined by drying the samples at 100–105 °C until constant weight, protein content was estimated by the Kjeldahl technique using 6.25 as the correction factor, ash content was determined after incineration at 550 °C in a muffle furnace and lipid content was extracted by petroleum ether extraction using a Soxhlet apparatus (AOAC, 2012).

2.8. Statistical analyses

Triplicate analyses of the proximate compositions of each sample were performed for each treatment and the mean value ± SD was calculated. Statistical analyses were performed using one-way ANOVA at 95% of confidence level to compare mean values among the treatments. Data was further analyzed by the Tukey test when means were considered different ($P < 0.05$). Analyses were performed using the XLSTAT version 2012.6.08 software package (Addinsoft, Paris, France).

3. Results and discussion

3.1. Identification of sardine species by molecular tests

The reference fish members of the Clupeidae family, *S. aurita, S. pilchardus, S. sagax, S. caeruleus*, and *O. oglinum*, were confirmed by PCR using the primer set designed for the *CYTB* gene. All of them produced a single 147 bp amplicon characteristic of this family. The specificity of the primer pair was also tested against a DNA template from *Salmo salar* (salmon), a phylogenetically unrelated fish, obtaining several fragments, but not the 147 pb of Cupleidae members.

All the sardine samples, both eviscerated or in fillet form and with preserved morphological features, fresh or frozen, sampled from different markets, fishing warehouses, street markets and in an industrial plant that produces canned sardines, showed a band of the expected 147 bp fragment size (Fig. 1).

To identify and distinguish between the different sardine species, the samples were screened by PCR-RFLP (Fig. 2). The digestion profile of the 147 bp amplicon was compared to the profile obtained from the reference species, which can be marketed as sardines, as set by the Ministry of Agriculture, Livestock and Supply (MAPA - Ministério da Agricultura, Pecuária e Abastecimento) (Fig. 2) and to the expected fragments obtained by the *in silico* (theoretical) analysis of the restriction sites for *Mnl*1 and *Hinf*1 of the 147 bp fragment (Table 1). From the nucleotide sequence of the 147 bp fragment available in GenBank, the double digestion of *S. pilchardus* by *Hinf*1 and *Mnl*1

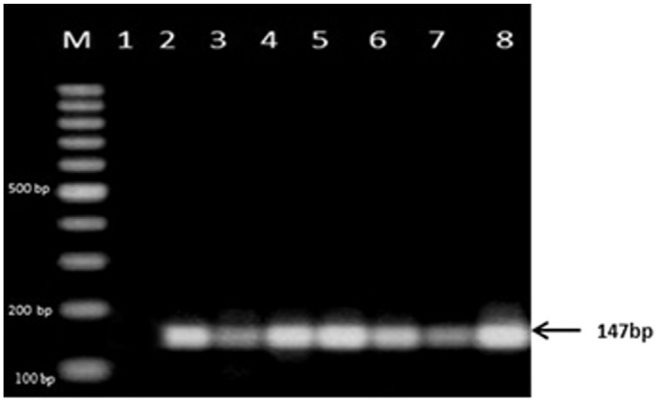

Fig. 1. Primer specificity for detecting sardine species. PCR products (5 μL) were loaded on a 2% agarose gel stained with GelRed. Lanes: M, 100 bp DNA ladder; lane 1, negative control; lane 2, *S. pilchardus* (positive control); lanes 3–5, sardine samples from the fish market; lanes 6–8, sardine samples from the industrial plant.

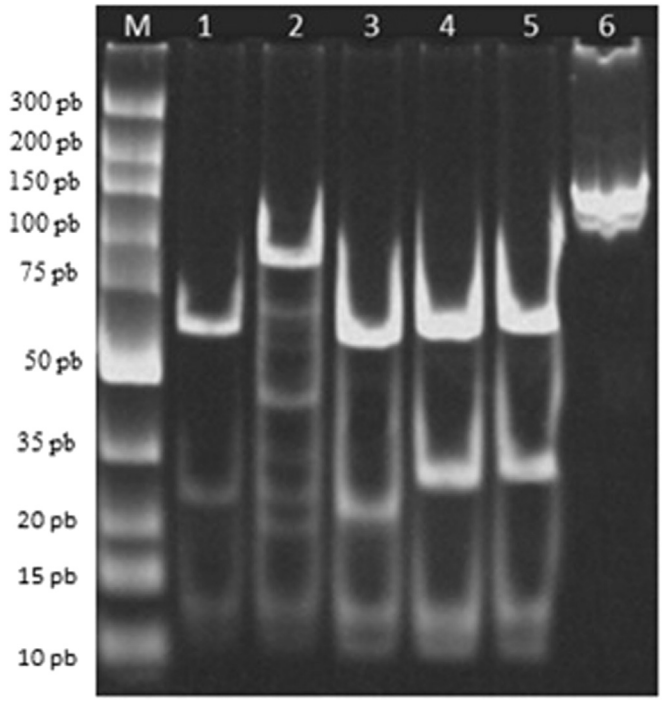

Fig. 2. Restriction fragment pattern of the 147 bp amplicon of the *CYTB* gene from sardine species (PCR-RFLP). The 147 bp amplicons were digested by restriction endonucleases *MnlI* and *HinfI*. Lane: M, 10 bp DNA ladder (Promega, Madison, WI, USA); lane 1, *Sardinella aurita*; lane 2, *Opisthoanema oglinum*; lane 3, *Sardina pilchardus;* lane 4, *Sardinops caeruleus;* lane 5, *Sardinops sagax* and lane 6, undigested product.

Table 1
PCR-RFLP theoretical analysis of the 147 bp amplicons from the *CYTB* gene sequence of fish species digested by *Hinf*1, by *Mnl*1 or by both endonucleases. The grey shade is the name of each column.

Species	*Hinf*1	*Mnl*1	*Hinf*1/*Mnl*1
Sardinella aurita	–	79, 57, 27, 25, 18 bp	–
Sardina pilchardus	×	78, 35, 30, 18 bp	72.6 bp
Sardinops caeruleus	×	79, 41, 18 bp	×
Opisthoanema oglinum	×	34, 25, 24, 18, 18 bp	×
Sardinops sagax	×	49, 18, 18 bp	×
Clupea harengus	–	70, 59, 11 bp	×
Brevoortia aurea		84, 56, 18 bp	×
Cetengraulis edentulus		97, 28 bp	×
Scomber japonicus		66, 36, 29, 27 bp	

In silico analyses of the 147 bp digestion by *Hinf*1, *Mnl*1 and *Hinf*1/*Mnl*1 endonucleases using the NEBcutter V2.0 software (New England BioLabs® Inc.).

endonucleases was expected to produce a typical profile with two major bands, of 78 and 35 bp (Table 1). *S. sagax* and *S. caeruleus* restriction profiles should be identical, formed by two 49 and 18 bp fragments (not discriminated in the polyacrylamide gel).

The polymorphism differences observed between the samples were sufficient to distinguish and classify between true and false sardine species (Fig. 2). *S. aurita* and *S. pilchardus* restriction profiles were very similar.

Among the sardine species, *O. oglinum* showed higher polymorphism when compared to *S. brasiliensis*. The latter is endemic to the Brazilian coast, geographically isolated from other species of the genus. *S. brasiliensis* was not found among the samples collected for this study due to over-exploitation of this species in Brazil, as fishing takes place both during the closed and open season.

The same type of polymorphism was found in two species, *Brevoortia aurea* and *Scomber japonicus*, making it impossible to tell them apart. As a result of the similarity in the restriction profiles, there was ambiguity in the identification of different species and because of this, sequencing of the 147 bp fragment was conducted.

The identity of the samples was confirmed by sequencing the 147 bp fragment of the *CYTB* gene (Table 2), where 102 of the 170 samples (60%) identified as *S. pilchardus* (25.9%) and *S. aurita* (18.8%), known as sardines; *S. caeruleus* (0.6%) and *S. sagax* (2.9%), known as the Pacific sardine and *O. oglinum* (11.8%), known as the Slab-sardine. These four species can be called sardines, according to Brazilian current legislation (MAPA, 2010), while *S. aurita* is considered fraudulent in Europe, being one of the main species used in place of *S. pilchardus,* especially in processed products (Herrero et al., 2011).

The following species considered fraudulent were identified in 68 of 170 samples (40%): *Clupea harengus* (4.7%), *B. aurea* (21.2%), *Centengraulis edentulus* (6.5%), and S. *japonicus* (7.6%). Thus, 60% of fish marketed as sardines in the state of Rio de Janeiro were identified as sardine species. However, in some cases, the names of the products were incorrect, since all the Pacific and Slab sardine samples were marketed under the name sardine, and the term sardine, in Brazil, is reserved for exclusive use for denomination of *S. brasiliensis*, *S. pilchardus* and *S. aurita*. This misnomer characterizes fraud by species falsification.

The DNA sequencing confirmed the species identified by the RFLP analyses, even in canned fish samples, a processed material not always suitable for molecular analyses. Nucleotide sequencing analysis has also been used in different countries for identification of certain fish species (Ardura et al., 2011; Cutarelli et al., 2014).

S. aurita, that can be marketed as a sardine in Brazil, but not in Europe, corresponded to 18.8% of the samples. Because of the difficulty in differentiating between the two species, *S. aurita* and *S. pilchardus*, by PCR-RFLP, it would be necessary to confirm the identity of the species by DNA sequencing in order to securely export this fish to countries belonging to the European Community.

When it is compared the authentic samples between them, 43.1% of samples were *S. pilchardus* and between the fraud samples, it was identified 52.9% of *S. aurea*. *S. pilchardus* was the most prevalent species of sardines found (43%), and most of them were imported from Morocco and sampled from the canned production line. The label description was consistent with the actual product.

None *S. brasiliensis* samples were identified among the samples collected. The nucleotide sequencing analysis of sardines pre-identified as *S. brasiliensis* indicated they were in fact *S. aurita*, demonstrating that identification based on external morphological fish characteristics is flawed and confuses even for experienced professionals from the fish sector.

From the 82 samples from industrial plants in Rio de Janeiro, 8 samples were identified as fraudulent (9.8%). From the 40 samples from the industrial plant that imports the fish from Morocco, 3 samples were identified as fraudulent (7.5%), not corresponding to the species declared on the product label, but the fish species, however, are considered authentic in Brazil. Among the 22 canned sardine samples from industrial plants that acquire the fish captured in the coastline of the state of Rio de Janeiro, a higher fraud frequency was detected – 23%.

When it was considered fraud by mislabeling, *S. pilchardus* was replaced by others species as *S. aurita*, S. japonicas, *S. caeruleus* and *S. sagax*, where 75% of samples were identified as *S. pilchardus*, 10% as *S. aurita*, 7.5% as *S. japonicas*, 2.5% as *S. caeruleus* and 5% as *S. sagax*. *B. aurea* was the predominant species in fraud replacements (53%), but 100% of sardines trade in fishmongers.

Free markets and fishing cooperative sales outlet were found to fraud 84% and 55%, of the sardine traded, respectively. Sardine was replaced mainly by *B. aurea* in free markets and by *C. edentulus* in fish cooperatives.

Table 2
Sardine species identified by nucleotide sequencing of the 147 bp amplicon from the *CYTB* gene. Bold signifies the total samples.

Identified species	Common name	Acession number GenBank	Identity	Number of samples
S. aurita	Sardine	EU552619.1	99%	32
S. pilchardus	Sardine	FR851430.1	100%	44
S. caeruleus	Pacific sardine	AF472585.1	97%	01
S. sagax	Pacific sardine	AF472586.1	90%	05
O. oglinum	Slab sardine	EU552620.1	97%	20
				102
C. harengus	Herring	AF472580	100%	8
B. aurea	Yellow tail	EP564676.1	97%	36
C. edentulus	Anchovy (*manjuba boca torta*)	GU357645.1	99%	11
S. japonicus	Mackerel	KF264282.1	100%	13
				68

One hundred and seventy samples traded in the State of Rio de Janeiro, were collected in markets, fishing warehouses, street markets and in industrial plants that produce canned sardines. Authentic species are shaded and the fraudulent species are unshaded.

3.2. Phylogenetic relations among the Clupeidae family members marketed as sardines in Rio de Janeiro

The first phylogenetic tree was constructed to evaluate only the five authentic species of sardines (Fig. 3A). *S. pilchardus*, *S. sagax*, *S. aurita* and *S. caeruleus* were grouped in one cluster, while *O. oglinum* formed a separate cluster. The previous polymorphism analysis by PCR-RFLP had already indicated this result, since the restriction profile of the 147 bp amplicon for *O. oglinum* was different from that obtained for *S. pilchardus*, *S. sagax*, *S. aurita* and *S. caeruleus* (Table 2). *S. sagax* and *S. caeruleus* belong to the same genus, but were grouped in different subclusters and are phylogenetically distant.

The second phylogenetic tree was constructed with all the fish species identified from samples collected in the state of Rio de Janeiro. The fish were grouped into two clusters and one clade containing *S. aurita*, which showed the highest phylogenetic distance between all species commercialized in the cities of Niterói and Rio de Janeiro. The divergence found among the sardines and *S. aurita* may justify the reason that this species is not allowed to be marketed as a sardine in Europe (Fig. 3B).

The main cluster presented subclusters containing true species such as *S. sagax*, *O. oglinum*, *S. pilchardus* and *S. caeruleus* and fraudulent species such as *S. japonicus* and *B. aurea*, with a genetic distance between species of up to 0.1. The secondary cluster only grouped the fraudulent species *C. harengus* and *C. edentulus*, with a genetic distance below 0.1. The genetic distance between the primary cluster and the secondary cluster was 0.4. *S. aurita* was presented isolated, with a genetic divergence greater than 0.6 in relation to the other species from the different clusters of the phylogenetic tree.

A previous study demonstrated that *S. sagax* and *S. caeruleus* are genetically distant over 20%, which corresponds to a divergence between the Clupeidae family genus (Jerome et al., 2003). It seems

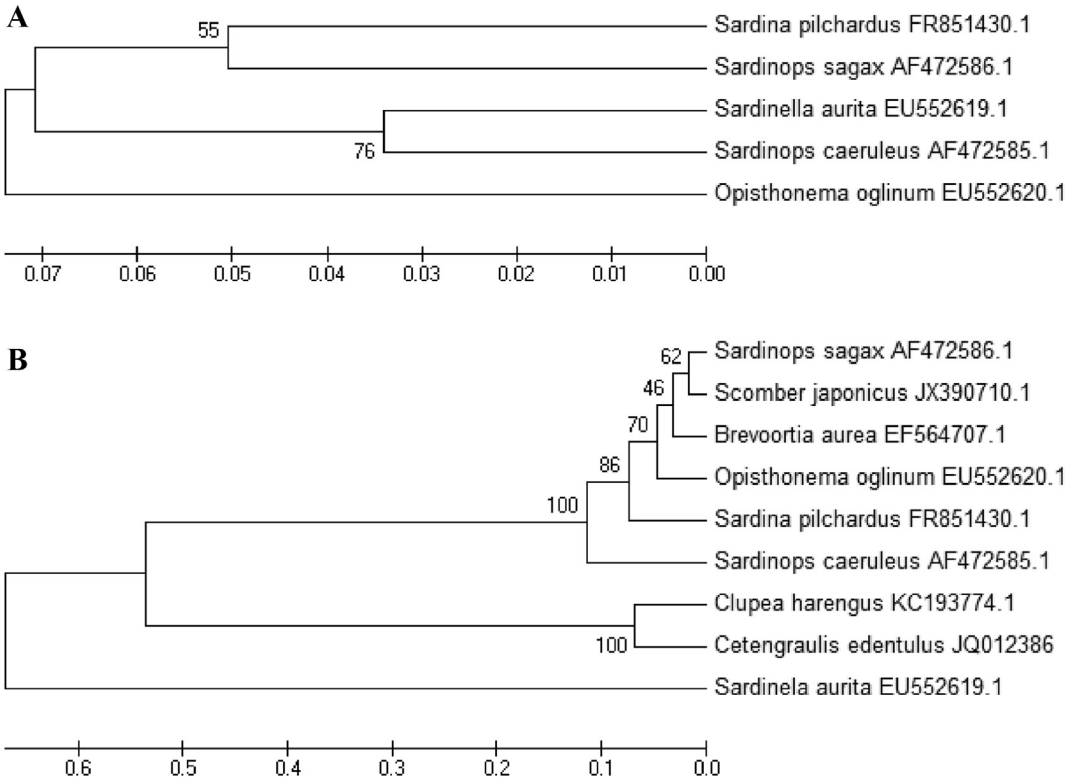

Fig. 3. A: Neighbor-joining tree of the genetic relationships among the true sardine species identified by the 147bp nucleotide sequencing. B: Neighbor-joining tree of genetic relationships obtained from the 147 bp nucleotide sequencing from true and false sardines traded in the State of Rio de Janeiro. True sardines: *S. aurita* (n = 32), *S. pilchardus* (n = 44), *S. caeruleus* (n = 1), *S. sagax* (n = 5), *O. oglinum* (n = 20); False sardines: *C. harengus* (n = 8), *B. aurea* (n = 36), *C. edentulus* (n = 11), *S. japonicus* (n = 13).

Table 3
Proximate composition of sardines identified by nucleotide sequencing of the 147 bp amplicon from the *CYTB* gene.

Parameters (g/100 g)				
Species	Protein	Lipid	Ash	Moisture
S. aurita	$20.65^{ab} \pm 0.08$	$4.63^{ab} \pm 0.10$	$1.21^{c} \pm 0.05$	$73.12^{bc} \pm 0.66$
S. pilchardus	$22.04^{a} \pm 0.64$	$2.21^{bc} \pm 0.03$	$0.45^{c} \pm 0.02$	$74.48^{bc} \pm 0.99$
S. caeruleus	$20.25^{ab} \pm 0.37$	$6.52^{a} \pm 0.33$	$0.74^{c} \pm 0.03$	$72.12^{bc} \pm 0.44$
O. oglinum	$14.94^{c} \pm 0.71$	$5.36^{a} \pm 1.33$	$1.53^{bc} \pm 0.54$	$71.97^{bc} \pm 0.47$
C. harengus	$11.26^{d} \pm 0.41$	$7.05^{a} \pm 1.20$	$1.07^{c} \pm 0.08$	$71.52^{c} \pm 1.07$
B. aurea	$11.57^{d} \pm 0.58$	$2.46^{bc} \pm 0.13$	$1.37^{c} \pm 0.30$	$75.64^{b} \pm 0.20$
C. edentulus	$12.12^{d} \pm 1.12$	$0.46^{c} \pm 0.16$	$2.90^{a} \pm 0.38$	$79.83^{a} \pm 1.35$
S. japonicus	$13.16^{cd} \pm 0.42$	$6.53^{a} \pm 0.57$	$0.62^{c} \pm 0.02$	$74.70^{bc} \pm 2.26$

Values are displayed as mean ± SD. These experiments were carried out in triplicate. Different letters in the same column indicate significant differences ($P < 0.05$). Authentic species are shaded and fraudulent species are unshaded.

likely that one of these species should be attributed to a distinct genus. For this reason, homology studies with other mitochondrial genes should be considered.

The divergence between *S. aurita* and *S. caeruleus* was greater than 0.5, differing from the results reported by Jérôme et al., 2003. *C. harengus* and *O. oglinum* belong to the same subfamily, but *C. harengus* is considered a fraudulent species in Brazil. In the present study these species were grouped into different clusters. Zhang (2011) described the proximity of clusters containing species of the Scombridae and Clupeidae genera (*S. japonicus* and *S. aurita*), where *S. japonicus* is considered a fraudulent species in Brazil. In the present study, these species were grouped into distinct clusters.

O. oglinum and *S. pilchardus,* that belong to the Clupeinae subfamily, were grouped in the same cluster as already discussed, and *B. aurea* (considered a fraudulent species), belonging to the Alosiae subfamily, was grouped near the Clupeinae subfamily cluster, while *C. harengus*, that belongs to the Clupeinae subfamily, is phylogenetically distant from these two clusters, confirming previously reported data (Li & Ortí, 2007).

3.3. Proximate composition of true and fraudulent sardines

The chemical composition regarding proteins, lipids and moisture of different sardine species, authentic or fraudulent, was evaluated (Table 3) and compared. *S. pilchardus*, *S. aurita* and *S. caeruleus* had higher protein content ($P < 0.05$) when compared to the other species identified in this study. The identified species can be divided into four groups considering protein content: the first group had content ranging from 20.25 ± 0.37 to 22.04 ± 0.64 g/100 g comprising *S. pilchardus, S. aurita* and *S. caeruleus*; the second group was represented by *S. brasilliensis*, with 18.49 ± 0.18/100 g; the third group had 13.16 ± 0.42 and 14.94 ± 0.71 g/100 g for *O. oglinum* and *S. japonicas,* respectively, and the fourth group ranged from 11.26 ± 0.41 to 12.12 ± 1.12 g/100 g including *C. edentulus, H. clupeola, B. aurea* and *C. harengus*.

Among the species considered authentic, *S. pilchardus* presented protein content higher than that of *S. brasilliensis* and all species identified as fraudulent had lower protein content when compared to the authentic species, with the exception of *O. oglinum*, which did not present any difference ($p > 0.05$) when compared to *S. japonicus*, considered fraudulent.

No difference was observed for moisture content ($P < 0.05$) between *S. brasilliensis*, *S. aurita, S. pilchardus, S. caeruleus* and *O. oglinum* when compared to *C. harengus, S. japonicus* and *B. aurea*. *H. clupeola* and *C. edentulus* showed the highest moisture content compared to the other species.

With respect to lipid content, *S. caeruleus, O. oglinum, S. aurita, C. harengus* and *S. japonicus* showed superior lipid content compared to the other species ($P > 0.05$). *S. brasilliensis, S. pilchardus, B. aurea, C. edentulus* and *H. clupeola* had lower lipid content ($P > 0.05$).

Regarding ash content, *C. edentulus* and *H. clupeola* showed the highest amount of inorganic residue ($P < 0.05$) when compared to the other species.

The proximate composition of sardines has been already evaluated in several studies: *O. oglinum* presented values of 10.18 g/100 g protein, 9.3 g/100 g fat, 1.73 g/100 g of ash and 71.13 g/100 g of moisture (Fernandes et al., 2014); *S. brasiliensis* had values of 22.0 g/100 g protein, 10.96 g/100 g fat, 2.96 g/100 g of ash and 63.85 g/100 g moisture (Tarley, Visentainer, Matsushita, & Souza, 2004) and *S. pilchardus* of 17.8, 12.2, 1.1 and 67.7 g/100 g of protein, fat, ash and moisture, respectively (Ozogul, Polat, & Ozogul, 2004). The difference in the results of the present study can be explained by exogenous (type and place of capture, season, environmental conditions, diet) and endogenous (specific physiological characteristics of the species) factors that contribute to changes in nutritional composition (Boran & Karaçam, 2011). Variations in lipid content have been reported previously, of 3.88–11.86 g/100 g, from February to December, for *S. pilchardus* (Zlatanos & Laskaridis, 2007). The present study found significant differences ($P < 0.05$) between the several sardine species. This difference was also observed in other studies, which reported 19.8, 13.3, 9.64 and 6.94 g/100 g lipid content for *S. brasilliensis, O. oglinum, B. aurea* and *S. japonicus*, respectively (Visentainer et al., 2007). A higher moisture content in the fraudulent samples was also observed when compared to the true species, which means that, by acquiring these species, the consumer would be buying a product with less mass than expected if it were an authentic sardine.

4. Conclusions

With the results of the present study, we conclude that molecular identification techniques are needed to ensure proper labeling of marketed fish, to facilitate the detection of fraud, prevent negative effects on the fishing industry and, in particular, ensure consumer rights. The *CYTB* gene 147 bp fragment was not enough to identify sardine species by RFLP, but is useful in species identification, allowing for unequivocal fish speciation by the nucleotide sequencing of this small fragment.

The PCR-RFLP method showed to be impractical when a large number of samples must be analyzed, requiring good data handling capacity to differentiate the restriction profiles. The most direct means of obtaining unequivocal information from PCR products is by sequencing, which allows for the unambiguous identification of sardine species caught and marketed in the State of Rio de Janeiro, Brazil, indicating that, if used by regulatory agencies, they would allow for surveillance in all lines of fish marketing, preventing fraud, ensuring internal consumer rights to choose and preserving the potential of fish exports for international consumer markets.

Acknowledgments

The authors acknowledge the financial support from Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ, Rio de Janeiro, Brazil; with grant Prioridade Rio 2014 – E-26/010.001968/2014), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasília, Brazil) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brasília, Brazil).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.foodcont.2015.08.004.

References

AOAC. (2012). In G. W. Latimer (Ed.), *Official methods of analysis of the association of official analytical chemists* (19th ed.). Arlington: AOAC Inc.

Ardura, A., Planes, S., & Garcia Vazquez, E. (2011). Beyond biodiversity: fish metagenomes. *PLoS ONE, 6*, 22592.

Boran, G., & Karaçam, H. (2011). Seasonal changes in proximate composition of some fish species from the Black Sea. *Turkish Journal of Fisheries and Aquatic Sciences, 11*, 01–05.

Cutarelli, A., Amoroso, M. G., De Roma, A., Girardi, S., Galiero, G., Guarino, A., et al. (2014). Italian market fish species identification and commercial frauds revealing by DNA sequencing. *Food Control, 37*, 46–50.

ECLAC (Economic Commission for Latin American and the Caribbean). (2006). United Nations Publication.

Fernandes, E. C., Vasconcelos, S. A. M., Ribeiro, A. M., Sarubbo, A. L., Andrade, C. A. S., & Filho, M. B. A. (2014). Nutritional and lipid profiles in marine fish species from Brazil. *Food Chemistry, 160*, 67–71.

Food and Agriculture Organization of the United Nations – FAO. (2014). *The state of world fisheries and aquaculture–Opportunities and challenges* (p. 3). Rome: FAO Fisheries and Aquaculture Department, FAO.

Ganias, K. (2015). *Biology and ecology of Sardines and Anchovies*. CRC Press, Taylor & Francis Group.

Granada, A. H., Flick, G. J., & Roy, E. M. (2012). *The seafood Industry: Species, processing and safety* (2nd ed., p. 488). Oxford, UK: Blackwell Publishing Ltd..

Herrero, B., Lago, C. F., Vieites, M. J., & Espinera, M. (2011). Development of a rapid and simple molecular identification methodology for true sardines (*Sardina pilchardus*) and false sardines (*Sardinella aurita*) based on the real-time PCR technique. *European Food Research and Technology, 233*, 851–857.

Hird, H. J., Chisholm, J., Kaye, J., Colyer, A., Hold, G., Conyers, C. M., et al. (2011). Development of real-time PCR assays for the detection of Atlantic cod (*Gadus morhua*), Atlantic salmon (*Salmo salar*) and European plaice (*Pleuronectes platessa*) in complex food samples. *European Food Research and Technology, 234*, 127–136.

Infante, C., Blanco, E., Zuasti, E., Crespo, A., & Manchado, M. (2007). Phylogenetic differentiation between Atlantic Scomber colias and Pacific Scomber japonicus based on nuclear DNA sequences. *Genetica, 130*, 1–8.

Jérôme, M., Lemaire, C., Bautista, J. M., Fleurence, J., & Etienne, M. (2003). Molecular phylogeny and species identification of sardines. *Journal of Agricultural and Food Chemistry, 51*, 43–50.

Li, C., & Ortí, G. (2007). Molecular phylogeny of Clupeiformes (Actinopterygii) inferred from nuclear and mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution, 44*, 386–398.

MAPA. (2010). *Plano Agrícola e Pecuário 2013/2014. Ministério da Agricultura, Pecuária e Abastecimento*. Secretária de Política Agrícola. Portaria 406, articulo 3.

Martinez, I., James, D., & Loreal, H. (2005). *Application of modern analytical techniques to ensure seafood safety and authenticity*. FAO Fisheries Technical Paper 455 (p. 73). Rome, Italy: FAO.

Mol, S. (2011). Levels of heavy metals in Canned Bonito, Sardines, and Mackerel produced in Turkey. *Biological Trace Element Research, 143*, 974–982.

Moretti, V. M., Turchini, G. M., Bellagamba, F., & Caprino, F. (2003). Traceability issues in fishery and aquaculture products. *Veterinary Research Communications, 27*(Suppl. 1), 497–505.

Mueller, S., Handy, S. M., Deeds, J. R., George, G. O., Broadhead, W. J., Pugh, S. E., et al. (2015). Development of a COX1 based PCR-RFLP method for fish species identification. *Food Control, 55*, 39–42.

Ozogul, F., Polat, A., & Ozogul, Y. (2004). The effects of modified atmosphere packaging and vacuum packaging on chemical, sensory and microbiological changes of sardines (*Sardina pilchardus*). *Food Chemistry, 85*, 49–57.

Prado, M., Boix, A., & von Holst, C. (2013). Development of a real-time PCR method for the simultaneous detection of mackerel and horse mackerel. *Food Control, 34*, 19–23.

Rasmussen, R. S. H., & Morrissey, M. T. (2011). Advances in DNA-based techniques for the detection of seafood species substitution on the commercial market. *Journal of laboratory Automation, 16*, 308–321.

Tarley, C. R. T., Coltro, W. K. T., Matsushita, M., & de Souza, N. D. (2001). Characteristic levels of some heavy metals from Brazilian canned sardines (*Sardinella brasiliensis*). *Journal of Food Composition and Analysis, 14*, 611–617.

Tarley, R. T. C., Visentainer, V. J., Matsushita, M., & Souza, E. N. (2004). Proximate composition, cholesterol and fatty acids profile of canned sardines (*Sardinella brasiliensis*) in soybean oil and tomato sauce. *Food Chemistry, 88*, 1–6.

Teletchea, F. (2009). Molecular identification methods of fish species: reassessment and possible applications. *Reviews in Fish Biology and Fisheries, 19*, 265–293.

[USFDA]. (2006). *U.S. Food and Drug Administration Center for Food Safety and Applied Nutrition*. Seafood substitution. College Park, MD: U.S. Food and Drug Administration Center for Food Safety and Applied Nutrition. Available from http://www.cfsan.fda.gov/%7efrf/econ.html Accessed June, 2015.

Visentainer, V. J., Noffs, A. D., Carvalho, O. P., Almeida, V. V., Oliveira, C. C., & Souza, E. N. (2007). Lipid content and fatty acid composition of 15 marine fish species from the Southeast Coast of Brazil. *Journal of American Oil Chemists' Society, 84*, 543–547.

Zhang, J. (2011). Species identification of marine fishes in China with DNA barcoding. *Evidence-Based Complementary and Alternative Medicine*, 10. Article ID 978253.

Zlatanos, S., & Laskaridis, K. (2007). Seasonal variation in the fatty acid composition of three Mediterranean fish – sardine (*Sardina pilchardus*), anchovy (*Engraulis encrasicholus*) and picarel (*Spicara smaris*). *Food Chemistry, 103*, 725–728.