Evaluation of a fast method based on the presence of two restriction sites in the mitochondrial ND5 (mtND5) gene for the identification of *Scomber* species

Iciar Martinez¹,²,*, Amaya Velasco¹, Ricardo Pérez-Martín¹, Eskil Forås³, Michiaki Yamashita⁴ and Carmen G. Sotelo¹

¹Instituto de Investigaciones Marinas (IIM-CSIC), Eduardo Cabello 6, E-36208 Vigo, Spain.
²Norwegian College of Fishery Science, University of Tromsø, NO-9037 Tromsø, Norway.
³SINTEF Fisheries and Aquaculture, SINTEF Sealab, Brattørkaia 17C, NO-7010 Trondheim, Norway.
⁴National Research Institute of Fisheries Science, 2 12 4 Fukuura, Yokohama 236 8648, Japan

*Corresponding author: Iciar Martinez, Instituto de Investigaciones Marinas (IIM-CSIC), Eduardo Cabello 6, E-36208 Vigo, Spain Tel: +34-986 231930, Fax: +34-986 292 762, Email: iciarm@iim.csic.es.

KEYWORDS
Species identification, fish, mackerel, Scomber, PCR, FINS, mtDNA, RFLP-PCR, saba, caballa, scombrus, colias, japonicus, australasicus, traceability.

Acknowledgements

We would like to thank Dr Jun Sakai, Chief Researcher of the Food Marketing Research and Information Center (FMRIC), Tokyo, Japan, for his assistance during the work and for providing the products in the Japanese market. Susana Otero from the Instituto de Investigaciones Marinas-CSIC is acknowledged for excellent technical assistance. The work was financed by The Research Council of Norway, project "Traceability control mechanisms in the mackerel food chain between Norway and Japan- TraCtrolMae" (no. 25181-0704011000009).
Evaluation of a fast method based on the presence of two restriction sites in the mitochondrial ND5 (mt ND5) gene for the identification of *Scomber* species

**KEYWORDS**
Species identification, fish, mackerel, *Scomber*, PCR, FINS, mt DNA, RFLP-PCR, saba, caballa, scombrus, colias, japonicus, australasicus, traceability.

**ABSTRACT**
The purpose of this work was to evaluate the suitability of a method based on the presence of two restriction sites (for *Hae III* and *Hindf I*) in the mitochondrial NADH dehydrogenase subunit 5 (mt ND5) gene to identify *Scomber* species. The evaluation was performed on 144 reference and market samples by sequencing of the entire 505 bp fragment of the mt ND5 gene and of a 464 bp fragment of the Kocher fragment of the cytochrome b gene (mt Cytb).

Sequence analysis of any of the two fragments allows the identification of each of the four *Scomber* species, but *S. japonicus* and *S. colias* had the same restriction sites at the ND5 amplicon and would not have been differentiated by this analysis. Similarly, loss of the *Hae III* site in some *S. scombrus* individuals would have misidentified them as not being *Scomber*.

All the market products were correctly labeled except one acquired in Spain labeled as originating in the Atlantic and containing *S. japonicus*.

**INTRODUCTION**
Classically only three species were considered within the world-wide distributed genus *Scomber*: *S. scombrus* found mainly in the North Atlantic, Mediterranean and Black Sea; *S. australasicus* in the Pacific and Southeast Indian Oceans and *S. japonicus*, in warm and temperate waters of the Atlantic, Indian and Pacific Oceans and adjacent seas (Infante, et al., 2007 and references therein). However, significant differences detected between *S. japonicus* from the Atlantic and the Pacific in their morphology (for example *S. japonicus* from the Pacific had very few or no belly spots while *S. japonicus* from the Atlantic had a more prominent spot pattern and different scale size) and parasites (Matsui, 1967) led some taxonomist to propose the recognition of two species: *S. japonicus* in the Pacific and *S. colias*...
in the Atlantic. This proposal was been further substantiated by the analysis of the nuclear 5S rDNA (Infante et al., 2007) and of partial (Espíñeira et al., 2009) and entire mt DNA (Catanese et al., 2010a). It is therefore nowadays agreed that there are four species in the genus Scomber: S. scombrus and S. colias in the Atlantic Ocean and S. japonicus and S. australasicus in the Pacific Ocean. Analysis of the mt DNA was also used by Nesbø et al., (2000) to study the population structure of S. scombrus.

In order to ensure truthful labeling, analytical methods have been developed to identify the species of the ingredients contained in food products. It has been shown that DNA-based methods are usually the most reliable and easier to perform on a great variety of products, since the processing does not alter the diagnostic DNA sequences (Mackie et al., 1999; Martinez et al., 2003; Sotelo and Pérez-Martín, 2007), although it may render them shorter (Quinteiro et al., 1998). Different DNA fragments have been targeted for species identification and many laboratories use forensically informative sequences (Bartlett and Davidson, 1992) often from the mt DNA due to its resilience and high copy number in most tissues. In our lab, we use sequencing of a 464 bp (Burgener, 1997), or of a shorter 140 bp fragment for canned products (Quinteiro et al., 1998; Meyer, 1993) from the Kocher fragment of the mt cytochrome b (mt Cytb gene (Kocher et al., 1989). Recently published Scomber spp identification methods also target the mt DNA (Botti and Guiffra, 2010; Catanese et al., 2010b). However, faster methods not requiring sequencing are usually preferred and have led to the development of a multiplex PCR method targeting a 123 bp fragment of mitochondrial NADH dehydrogenase subunit 5 gene (mt ND5) with a positive control amplicon corresponding to the small 12S rRNA subunit, proposed by Catanese et al. (2010b) to differentiate the four Scomber species and to a restriction fragment polymorphism (RFLP) analysis of a 505 bp fragment using Hae III and Hinf I, also from the mt ND5, proposed in Japan to differentiate S. japonicus, S. australasicus and S. scombrus (Anon. 2007).

The work presented here is part of the research project TraCtrolMac which deals with the identification of traceability control mechanisms for S. scombrus exported from Norway to Japan. Japan is the largest importer of Norwegian mackerel: of a total of 265,000 tons of exported frozen round mackerel in 2010 for a value of 2.8 billion Norwegian kroner, Japan acquired 75,000 tons. Most of this mackerel, about 51,000 tons, was first exported to China where it was filleted and marinated before entering the Japanese market. The production chain of Norwegian mackerel to the Japanese market was mapped by Forås and Thakur (2010) and Thakur et al., (2011). The identification numbers used in the traceability chain could also be used as references for product information if an additional type of identifier, with which it
would not be possible to tamper, such as DNA profiles or diagnostic sequences, accompanied
to the product. Software based traceability solutions would enable a direct linking of the product
to the results of the genetic analysis and to the responsible laboratory. By law, seafood
products must be labeled with information regarding the species and geographic origin of
capture both in Europe (EU CR No 2065/2001 of 22 October 2001 laying down detailed rules
for the application of Council Regulation EC No 104/2000 which regards informing
consumers about fishery and aquaculture products) and in Japan (Law on Standardization and

It is obviously of the outmost importance that the analyses performed using different
methods and in different countries on the same samples produce consistent results.

The aim of this work is to evaluate the suitability of the analysis proposed in Japan
based on polymorphisms at two restriction sites in the *mt ND5* gene (Anon. 2007) for
*Scomber* species identification purposes.

**MATERIALS AND METHODS**

*Authentic mackerel species and commercial products:* 144 samples were analyzed that
included authentic reference samples (processed and unprocessed) and products of *S.
scombrus, S. colias, S. japonicus* and *S. australasicus*. The samples purchased in the Japanese
market were labelled with the generic term "saba" (mackerel) that refers to any *Scomber*
species, except for one product labelled as "masaba" that should contain *S. japonicus*. All the
products purchased in Japan and labelled as "of Norwegian origin" should contain *S.
scombrus*. The products purchased in Norway should all of them contain *S. scombrus*. The
correct names allowed to mackerel products in the Spanish market are regulated by law
(B.O.E., 2010): one of the three Spanish canned products was labelled as *caballa* (generic
term for all four species) from Cabo Verde, and should therefore contain *S. scombrus* or *S.
colias*. The other two were labelled with the term *caballa del sur*, that refers to both *S. colias*
or *S. japonicus* (BOE, 2010); one was from Perú, and should therefore contain *S. japonicus*
and the other from Isla Cristina (Spain) and should contain *S. colias*. The reference samples of
*S. colias* belonged to the reference collection of the Instituto de Investigaciones Marinas and
had been preserved frozen. All the other samples had been preserved in 96% ethanol until the
DNA was extracted. One sample of canned herring was included as a negative control and it
was correctly identified as *Clupea harengus*. 

4
DNA extraction, PCR amplification, sequencing and data treatment: DNA was extracted using the standard Wizard™ DNA Clean-up System (Promega). PCR amplifications were carried out (1) by the method described by Kocher et al., (1989) as modified Burgener (1997) to amplify a 464 bp fragment of the mt Cyb (Kocher fragment) using the primers H15149AD: 5’-GCICCTCARAATGAYATTTGTCCTCA-3’ and L14735: 5´-AAAAACCACCGTTGTTATTCAAACTA-3’ and (2) as described in the "Manual determination for the fish species Scomber" (Anon. 2007) to amplify the 505 bp fragment of the mt ND5 gene using the primers LSs1-LEU: 5´-ATCCGCTGGTCTTAGGAACC-3’ and HSs1-ND5: 5'-CCTTCTCAGCCGATAAATAGTT-3’. This method is based on the assumption that only three Scomber species need to be distinguished. Although the Norwegian canned products amplified satisfactorily both reactions, the three Spanish canned products did not and were analyzed following the protocol described by Quinteiro et al., (1998) using primers H15573- 5’-AAT AGG AAG TAT CAT TCG GGT TTG ATG -3’ and L15424-5’-ATC CCA TTC CAC CCA TAC TAC TC- 3’; corresponding to those described by Meyer (1993) that amplify a 176 bp suitable for heavily fragmented DNA. PCR reactions were performed in volumes of 25 μL using illustra TM puReTaq Ready-To-Go PCR beads (GE Healthcare) and 20 μl of PCR product were sequenced in an ABI PRISM 310 DNA Sequencer (Applied Biosystems). The diagnostic Hae III and Hinf I restriction sites described by Anon (2007) between the positions 243-244 and between 415-416 respectively of the mt ND5 amplicon were identified from the sequences. The data were processed using the software BIOEDIT (Hall, 1999) with CLUSTAL used to align the sequences (Thompson et al., 1994) and MEGA 4.0 to construct phylogenetic trees (Tamura et al., 2007) according to Tamura and Nei (1993), Saitou and Nei (1987) and Felsestein (1985).

RESULTS AND DISCUSSION

As already mentioned, the aim of this study was to evaluate the suitability of a fast method proposed in Japan and based on the presence of two polymorphic restriction sites in the mt ND5 gene for the identification of Scomber species. Our results target the food control authorities and research community responsible for providing consumers, the seafood industry and traders of mackerel worldwide with reliable, tested and validated methods for species identification. Mislabelling maybe very serious and it is important to differentiate the species from the Atlantic from those of the Pacific because they carry different hazards and constitute different potential risks. In the absence of food safety considerations however, the consumers’
usual preferences for their local products are reflected in higher market prices. Therefore, it is of the outmost relevance that all the stakeholders involved in the mackerel chain have the same data and validated official methods to refer to in case of disagreements, disputes or mislabellings.

Comparison of the results using Kocher and mt ND5 fragments: Fig. 1 shows the phylogenetic trees constructed with genetic distances among specimens of the 4 mackerel species with Kocher and ND5 sequences. Our results confirm the already reported (Infante et al., 2007) close relatedness of S. colias, S. japonicus and S. australasicus at the two loci tested, but the loci registered different variability depending on the species (see the number of haplotypes identified at each locus in Tables 1 and 2). Thus, even if the number of samples is small to extrapolate to real population variability, the Kocher fragment seemed to be more polymorphic in the species from the Atlantic, i.e., S. scombrus and S. colias while the mt ND5 seemed to show a higher number of polymorphisms in the species from the Pacific ocean, S. japonicus and S. australasicus, (Table 3). However, for each species and locus, there was one widely predominant haplotype with over 50% of the samples belonging to it; which were Kocher haplotypes number 1, 26, 27 and 36 and ND5 haplotypes number 1, 20 25 and 27 for S. scombrus, S. japonicus, S. colias and S. australasicus respectively (see Tables 1 and 2).

The RFLP analysis described by Anon. (2007) is based on two polymorphic sites in the mt ND5 sequence: one Hae III site between bp 243-244 of the amplicon present in S. scombrus and S. japonicus and absent in S. australasicus and a target site for Hinf I (bp 415-416) present only in S. scombrus. Sequence analysis of all the samples from this work showed that all the S. colias had the same sequence as S. japonicus at these two sites and would therefore have been misidentified as S. japonicus. In addition 3 haplotypes of S. scombrus corresponding to 4 samples would have remained unidentified due to the loss of the Hae III site: from the most frequent sequence GGCC to GGCA in haplotype 6 (n=2) to GACC in haplotype 8 (n=1) and to GGGC in haplotype 11 (n=1).

Identification of products from the Japanese, Norwegian and Spanish markets: None of the products purchased was labelled with the name of the species. Most of them used a generic term: saba in Japan, makrell in Norway and caballa or caballa del sur in Spain, followed by a region from which one should expect a given species, except in the case of Norway, where makrell is commonly understood as being only S. scombrus.

Sequence analysis of either the Kocher fragment or the mt ND5 allowed the easy identification of the mackerel products belonging to either of the four Scomber species; except, as expected, for some of the canned products that required the use of a shorter
diagnostic sequence. All the products were correctly labelled according to their claimed origin except the Spanish canned product from Cabo Verde that should have contained *S. colias* for its claimed Atlantic provenance but it was shown to be *S. japonicus*. Regarding the analysis of samples from the Japanese market, two of three products labelled with the generic "saba" term contained *S. australasicus* and the third *S. japonicus*. We could also identify that the three pieces of one of the Japanese products belonged at least to two different *S. australasicus* individuals.

An optimal analysis for species identification should be of high throughput, fast and easy to perform. However, the PCR-RFLP method examined here and proposed by Anon. (2007) is not optimal and would not have identified any *S. colias*. This is not surprising due to the very close genetic relationship between *S. japonicus* and *S. colias* but given the different geographic origin of these two species and the globalization of fisheries and trade, a method that permits the unequivocal identification of all four species should be implemented. An additional problem of RFLP analysis is that intraspecific variability affecting the restriction sites may lead to false positive or negative identifications. In the present work 4 samples of *S. scombrus* would have remained unidentified by the RFLP analysis.

As the methods here used, most of the advance methods published to identify *Scomber* and Scombrid species are based on sequence analysis (Infante et al., 2007; Espiñeira et al., 2009; Catanese et al., 2010a,b; Botti and Giuffra, 2010) and all of them are reliable for this purpose. However, simpler analyses need to be developed to diminish the costs both in terms of time and money that sequencing demands. Of the published fast methods, only the multiplex described by Catanese et al., (2010b) is able to discriminate the four species, but close examination of the results illustrates one of the facts usually encountered in multiplex analysis: in some samples the amplification of either the control of the specific band is weaker than it should have been and in others the presence of weak but detectable shadows may hinder a clear and undoubtful identification, for example in case the shadow represents a mixture of two species where one of them is present in much smaller amounts or due to the close relatedness between the species examined.

**CONCLUSION**

In conclusion, the two loci tested are suitable for species identification of *Scomber* species using sequence analysis, but an analysis based on the presence of the *Hae III* and *Hinf*
I restriction sites in the mt ND5 amplicon would misidentify all S. colias and would not be able to identify as Scomber some genotypes of true S. scombrus. Sequencing remains a demanding and time consuming task, and future improvements should focus in designing fast specific methods for each of the four species that increase the reliability of the analysis and decrease the number of steps involved, such as those based on the use of species-specific probes and/or Real-Time PCR techniques.

REFERENCES


Figure 1.- Dendograms obtained by Neighbor-Joining of the 40 haplotypes identified at the Kocher (upper) and the 38 haplotypes at the ND5 (bottom) loci by the Tamura-Nei distance measurement method.
TABLE 1. Haplotypes identified by the Kocher fragment. 138 samples were analyzed by this method.

<table>
<thead>
<tr>
<th>Species</th>
<th>Haplotype number</th>
<th>n of individuals in the haplotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. scombrus</td>
<td>1</td>
<td>48</td>
</tr>
<tr>
<td>S. scombrus</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>S. scombrus</td>
<td>3 to 5</td>
<td>2</td>
</tr>
<tr>
<td>S. scombrus</td>
<td>6 to 25</td>
<td>1</td>
</tr>
<tr>
<td>S. japonicus</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>S. colias</td>
<td>27</td>
<td>4</td>
</tr>
<tr>
<td>S. colias</td>
<td>28</td>
<td>2</td>
</tr>
<tr>
<td>S. colias</td>
<td>29 to 35</td>
<td>1</td>
</tr>
<tr>
<td>S. australasicus</td>
<td>36</td>
<td>13</td>
</tr>
<tr>
<td>S. australasicus</td>
<td>37</td>
<td>3</td>
</tr>
<tr>
<td>S. australasicus</td>
<td>38 to 40</td>
<td>1</td>
</tr>
</tbody>
</table>

TABLE 2. Haplotypes identified by the ND5 fragment. 132 samples were analyzed by this method.

<table>
<thead>
<tr>
<th>Species</th>
<th>Haplotype number</th>
<th>n of individuals in the haplotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. scombrus</td>
<td>1</td>
<td>41</td>
</tr>
<tr>
<td>S. scombrus</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>S. scombrus</td>
<td>3 and 4</td>
<td>5</td>
</tr>
<tr>
<td>S. scombrus</td>
<td>5 and 6</td>
<td>2</td>
</tr>
<tr>
<td>S. scombrus</td>
<td>7 to 19</td>
<td>1</td>
</tr>
<tr>
<td>S. japonicus</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>S. japonicus</td>
<td>21</td>
<td>6</td>
</tr>
<tr>
<td>S. japonicus</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>S. japonicus</td>
<td>23 and 24</td>
<td>1</td>
</tr>
<tr>
<td>S. colias</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>S. colias</td>
<td>26 and 27</td>
<td>1</td>
</tr>
<tr>
<td>S. australasicus</td>
<td>28</td>
<td>8</td>
</tr>
<tr>
<td>S. australasicus</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td>S. australasicus</td>
<td>30 to 38</td>
<td>1</td>
</tr>
</tbody>
</table>

TABLE 3. Comparison of the number of haplotypes identified in the four Scomber species at the Kocher and ND5 loci

<table>
<thead>
<tr>
<th>Species</th>
<th>Kocher n fish</th>
<th>Kocher n haplotypes</th>
<th>ND5 n fish</th>
<th>ND5 n haplotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. scombrus</td>
<td>79</td>
<td>25</td>
<td>75</td>
<td>19</td>
</tr>
<tr>
<td>S. japonicus</td>
<td>27</td>
<td>1</td>
<td>26</td>
<td>5</td>
</tr>
<tr>
<td>S. colias</td>
<td>13</td>
<td>9</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>S. australasicus</td>
<td>19</td>
<td>5</td>
<td>19</td>
<td>11</td>
</tr>
</tbody>
</table>