Specific identification of western Atlantic Ocean scombrids using mitochondrial DNA cytochrome C oxidase subunit I (COI) gene region sequences

MA Paine
*Virginia Institute of Marine Science*

Jan McDowell
*Virginia Institute of Marine Science*

John E. Graves
*Virginia Institute of Marine Science*

Follow this and additional works at: https://scholarworks.wm.edu/vimsarticles

Part of the Aquaculture and Fisheries Commons, and the Marine Biology Commons

**Recommended Citation**
https://scholarworks.wm.edu/vimsarticles/1505

This Article is brought to you for free and open access by the Virginia Institute of Marine Science at W&M ScholarWorks. It has been accepted for inclusion in VIMS Articles by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
SPECIFIC IDENTIFICATION OF WESTERN ATLANTIC OCEAN SCOMBRIDS USING MITOCHONDRIAL DNA CYTOCHROME C OXIDASE SUBUNIT I (COI) GENE REGION SEQUENCES

Melissa A. Paine, Jan R. McDowell, and John E. Graves

ABSTRACT

Identification of scombrids (tunas, mackerels, bonitos, etc.) is difficult when morphological characters are ambiguous or missing, such as with early life history stages or tissues found in the stomachs of predators. The mitochondrial cytochrome c oxidase subunit I (COI) gene region was evaluated as a molecular marker for the specific identification of the 17 members of the family Scombridae common to the western Atlantic Ocean. A 950 base pair region in the COI gene was sequenced from up to 20 individuals of each species, and suites of nucleotide polymorphisms that unambiguously distinguish among these scombrid species were identified. A shorter 250 base pair fragment of COI proved to be sufficient for species identification and was better suited for analyzing degraded tissue samples. Scombrid larvae collected in the Florida Straits and scombrid remains in the stomachs of large pelagic predators were used to demonstrate the utility of both the long and short COI fragments.

Members of the family Scombridae (tunas, mackerels, bonitos, etc.) are important components of pelagic ecosystems, with several species supporting large commercial and recreational fisheries throughout the world’s oceans. Proper identification of these species at all life stages and in various conditions, even as degraded stomach contents, is essential to better understand early life history characteristics and ecological relationships in the pelagic ecosystem, and to enable effective management. In addition, specific identification of processed tissues or fillets is necessary for enforcement of fisheries management regulations.

While specific identification of adult scombrids is essentially unambiguous (Collette and Nauen, 1983), identification is problematic in situations where morphological characters are difficult to interpret (early life history stages) or missing (fillets, digested stomach contents). Identification of early life history stages of scombrids has been especially challenging. Scombrid eggs, larvae (especially those of the genus Thunnus), and juveniles generally cannot be distinguished unambiguously based solely on morphology (Richards et al., 1990).

Molecular markers can provide a means for positive identification when morphological identification is uncertain. Various molecular markers have been used to identify fish eggs and larvae including allozymes (Morgan, 1975), polymerase chain reaction (PCR)/restriction fragment length polymorphism (RFLP) analysis (Daniel and Graves, 1994; McDowell and Graves, 2002), multiplex PCR (Rocha-Olivares, 1998; Hyde et al., 2005) and sequencing (Hare et al., 1994; Kirby and Reid, 2001; Perez et al., 2005). Many of these techniques have been used to identify scombrids. Allozymes have been successfully used to discriminate between early juveniles of bigeye tuna Thunnus obesus (Lowe, 1839) and yellowfin tuna Thunnus albacares (Bonnaterre, 1788) (Graves et al., 1988) as well as between adult Pacific northern bluefin tuna Thunnus thynnus orientalis Serventy, 1956 and southern bluefin tuna Thunnus maccoyii Castelnau, 1872 (Ward, 1995). Several studies have used PCR/
RFLP analysis to identify species of the scombrid tribes Thunnini and Sardini (Chow et al., 2003) as well as eight species of the genus *Thunnus* (Chow and Inoue, 1993). In addition, sequencing of a mitochondrial gene region has been used to identify *Thunnus* species (Bartlett and Davidson, 1991; Ram et al., 1996; Quintero et al., 1998; Terol et al., 2002; Ward et al., 2005). While each of these techniques has advantages and disadvantages, sequencing provides the highest level of resolution as it shows genetic differences at the nucleotide level. A few studies have used sequence analysis to identify scombrids, but these investigations were limited as they only distinguished between a few species, used a region that revealed considerable intraspecific variation, had limited sample sizes, or encountered problems with non-specific amplification (Bartlett and Davidson, 1991; Ram et al., 1996; Quintero et al., 1998; Terol et al., 2002; Ward et al., 2005). Additionally, molecular techniques have only recently appeared promising as an answer for the difficulties associated with identifying degraded remains in stomachs (Harper et al., 2005; Gorokhova, 2006), and sequencing has been used to identify gut contents of marine invertebrates (Blankenship and Yayanos, 2005) and large pelagic fishes (Smith et al., 2005).

The mitochondrial genome has been preferred for analysis in many genetic studies as it has a high number of copies per cell, which facilitates PCR amplification, and because the presence of a single allele makes it possible to sequence products directly (Avise, 1994). Many mitochondrial gene regions (cytochrome b, ND4, 16S, COI) have been successfully used for fish identification (Bartlett and Davidson, 1991; McDowell and Graves, 2002; Hyde et al., 2005; Lopez and Pardo, 2005). These gene regions display different levels of genetic variation as a result of different evolutionary rates. While variation is necessary to highlight interspecific differences, too much variation can hinder primer design. Because of this, the use of a conserved region is advantageous for effective amplification across many species.

One of the most conserved protein coding genes in the mitochondrial (mt) genome is cytochrome c oxidase subunit I (COI) (Brown, 1985). COI is critical for cellular energy production and this functional importance constrains its evolution (Rawson and Burton, 2002). The high level of conservation of COI allows for the design of a unique primer pair that successfully amplifies the same fragment across the diverse members of the Scombridae. Previous work has taken advantage of COI for broad taxonomic studies (11 invertebrate phyla, Folmer et al., 1994; 11 animal phyla, Hebert et al., 2003), but COI has also been useful to distinguish closely related genera in species identification (three copepod genera, Bucklin et al., 1999). Efforts to use a segment of DNA as a barcode of identity have successfully employed COI to identify various taxa including fish species (Steinke et al., 2005; Ward et al., 2005). Because COI is informative for distinguishing species across and within many different taxa, it is well suited for identification across a family as diverse as the Scombridae. In this study, a molecular key is developed based on the mitochondrial COI region for the specific identification of the 17 scombrids present in the western Atlantic Ocean.

**Materials and Methods**

Tissue samples were obtained from up to 20 specimens of each of the 17 scombrid species common to the western Atlantic Ocean: *Acanthocybium solandri* (Cuvier, 1831); *Auxis rochei* (Risso, 1810); *Auxis thazard* (Lacépède, 1800); *Euthynnus alletteratus* (Rafinesque, 1810); *Katsuwonus pelamis* (Linnaeus, 1758); *Sarda sarda* (Bloch, 1793); *Scomber colias* Gme-
lin, 1789; Scomber scombrus Linnaeus, 1758; Scomberomorus brasiliensis Collette, Russo and Zavalla-Camin, 1978; Scomberomorus cavalla (Cuvier, 1829); Scomberomorus maculatus (Mitchill, 1815); Scomberomorus regalis (Bloch, 1793); Thunnus alalunga (Bonnaterre, 1788); Thunnus albacares (Bonnaterre, 1788); Thunnus atlanticus (Lesson, 1830); Thunnus obesus (LOWE, 1834) and Thunnus thynnus (Linnaeus, 1758). All specimens were identified based on morphological characters (Collette and Nauen, 1983). Tissue samples were either stored in DMSO buffer (Seutin et al., 1991) or frozen. Published COI sequences of A. thazard and A. rochei (Infante et al., 2004) were used to supplement the number of samples for these species. Collection information is provided in Table 1.

To evaluate the efficacy of COI as a marker to identify scombrids, specimens of larval scombrids stored in ethanol were obtained from D. Richardson and R. Cowen, Rosenstiel School of Marine and Atmospheric Science, University of Miami. In addition, stomach content samples containing putative scombrids were collected from blue marlin and white marlin caught in recreational fishing operations out of Cape May, NJ, USA and La Guaira, Venezuela. Putative scombrids were removed from the marlin stomachs dockside and rinsed with water. Either a muscle sample was removed and placed in DMSO buffer (Seutin et al., 1991) or the whole fish was frozen until analysis.

Total genomic DNA was extracted from adult tissues of known scombrid species using a standard phenol/chloroform isolation protocol (modified from Sambrook and Russell, 2001). A series of extractions was performed on each sample using equilibrated phenol, followed by phenol: chloroform: isoamyl alcohol (25:24:1) and finally, chloroform: isoamyl alcohol (24:1). Following extraction, DNA was precipitated with ethanol. For larval fishes, one eyeball (right eyeball when available) was removed and rinsed with distilled water. DNA was extracted from this tissue using proteinase-K and Chelex beads (Bio-Rad Laboratories, Hercules, CA) (Estoup et al., 1996). Each larva was photographed using a digital camera attached to a stereomicroscope via a phototube, capturing as much detail as possible for future morphological or meristic analysis.

Primers that amplify the COI gene region across the scombrid family were designed using conserved regions of seven scombrid COI sequences (A. rochei, A. thazard, E. alletteratus, K. pelamis, S. scombrus, T. alalunga, and T. thynnus) available through GenBank (National Center for Biotechnology Information). Two sets of primers were developed that amplify a ~950 base pair (bp) fragment (long fragment) of the COI gene and a ~250 bp fragment (short fragment) located within the 950 bp fragment:

- **950 bp fragment:** LCOI 121 CTA AGC CAA CCA GGT GCC CTT CT
  HCOI 1199 AAT AGT GGG AAT CAG TGT ACG A
- **250 bp fragment:** LCOI 646 AAT ACA ACC TTC TTC GAC C
  HCOI 947 GTT GGA ATT GCG ATA ATC

(The number in the primer name designates the position of the 5’ end within the COI gene (1550 bp)). All primers were ordered from Invitrogen Corporation (Carlsbad, California).

Polymerase chain reactions were performed on each known or putative scombrid sample. Each 25 µl reaction consisted of 0.25 µl template DNA, 2.5 µl 10× PCR Buffer plus magnesium (QIAGEN, Inc., Valencia, CA), 0.5 µl 10 mM dNTP (QIAGEN), 0.25 µl forward primer (100 pm µl⁻¹), 0.25 µl reverse primer (100 pm µl⁻¹), 0.125 µl Taq DNA polymerase (QIAGEN), 5.0 µl BSA (bovine serum albumin) (1 mg mL⁻¹) and 16.125 µl sterile water (modified from McDowell and Graves, 2002). Amplifications using Chelex extractions contained 2.5 µl DNA template. Reactions were carried out in an MJ Research Corporation PTC-200 Peltier thermal cycler (Watertown, MA) under the following conditions: initial denaturation at 94 °C for 4 min, followed by 35 cycles of 94 °C for 1 min, 57 °C for 1 min, 72 °C for 2 min, a final extension at 72 °C for 5 min, and final hold at 4 °C. Amplifications done with the LCOI 646/ HCOI 1085 primers used an annealing temperature of 54 °C, but were otherwise run using the same conditions.
Table 1. Scombrid reference sample collection location and sample size.

<table>
<thead>
<tr>
<th>Species name</th>
<th>Common name</th>
<th>Abbreviation</th>
<th>Catch location (no. of samples), date collected</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acanthocybium solandri</td>
<td>Wahoo</td>
<td>ASOL</td>
<td>Cape May, New Jersey (12), 2004; Gulf of Mexico (5); Isla Mujeres, Mexico (1)</td>
<td>18</td>
</tr>
<tr>
<td>Auxis rochei</td>
<td>Bullet tuna</td>
<td>AUXR</td>
<td>Virginia (4), 2004; Philippines (1), 2003; Pacific (3) and Atlantic/Mediterranean (8) from Infante et al., 2004.</td>
<td>16</td>
</tr>
<tr>
<td>Auxis thazard</td>
<td>Frigate tuna</td>
<td>AUTC</td>
<td>Panama, Pacific (5), 2004; Philippines (1), 2003; Seychelles (1), 2002; Spain (3) from Infante et al., 2004.</td>
<td>10</td>
</tr>
<tr>
<td>Euthynnus alleteratus</td>
<td>Little tunny</td>
<td>EUTH</td>
<td>Isla Mujeres, Mexico (10)</td>
<td>10</td>
</tr>
<tr>
<td>Katsuwonus pelamis</td>
<td>Skipjack tuna</td>
<td>SKJT</td>
<td>Virginia (16), 2004; Hawaii (3), 2004</td>
<td>19</td>
</tr>
<tr>
<td>Sarda sarda</td>
<td>Atlantic bonito</td>
<td>SARD</td>
<td>Mid-Atlantic Bight (1), 1994; Mediterranean (2), 2002</td>
<td>3</td>
</tr>
<tr>
<td>Scomber colius</td>
<td>Chub mackerel</td>
<td>CHMK</td>
<td>Gulf of Mexico (10); Virginia (4), 2004; Argentina (3); Ivory Coast (3), 1992</td>
<td>20</td>
</tr>
<tr>
<td>Scomber scombrus</td>
<td>Atlantic mackerel</td>
<td>ATLM</td>
<td>Massachusetts (10); Plymouth, England (10)</td>
<td>20</td>
</tr>
<tr>
<td>Scomberomorus brasiliensis</td>
<td>Serra Spanish mackerel</td>
<td>SCBR</td>
<td>Brazil (2); Trinidad (1); Saba (1), 2006</td>
<td>4</td>
</tr>
<tr>
<td>Scomberomorus cavalla</td>
<td>King mackerel</td>
<td>KMCK</td>
<td>Panama City, Florida (2), 1994; South Carolina (2), 1996; Virginia (3), 2005; Maryland (1), 2005; Florida Keys (1), 2006</td>
<td>9</td>
</tr>
<tr>
<td>Scomberomorus maculatus</td>
<td>Spanish mackerel</td>
<td>SPMK</td>
<td>South Carolina (13), 1997; Chesapeake Bay (5), 1994; Gulf of Mexico (2)</td>
<td>20</td>
</tr>
<tr>
<td>Scomberomorus regalis</td>
<td>Cero</td>
<td>CERO</td>
<td>Bahamas (2); Colombia (1); Florida (1)</td>
<td>4</td>
</tr>
<tr>
<td>Thunnus alalunga</td>
<td>Albacore</td>
<td>ALBC</td>
<td>Mid-Atlantic Bight (3), 2004; Atlantic (12); North Carolina (2), 2005</td>
<td>17</td>
</tr>
<tr>
<td>Thunnus albacares</td>
<td>Yellowfin tuna</td>
<td>YFT</td>
<td>Mid-Atlantic Bight (14), 2004; Hawaii (4), 2004</td>
<td>18</td>
</tr>
<tr>
<td>Thunnus atlanticus</td>
<td>Blackfin tuna</td>
<td>BLKF</td>
<td>Isla Mujeres (6), 2004; Florida Straits (5), 2004; North Carolina (7) 2005</td>
<td>18</td>
</tr>
<tr>
<td>Thunnus obesus</td>
<td>Bigeye tuna</td>
<td>BET</td>
<td>Florida Straits (14), 2004; Hawaii (4), 2004</td>
<td>18</td>
</tr>
<tr>
<td>Thunnus thynnus</td>
<td>Atlantic bluefin tuna</td>
<td>BLFT</td>
<td>Ocean City, Maryland (14), 2003; Spain (2), 1998; Mediterranean (2), 1998; New Jersey (2)</td>
<td>18</td>
</tr>
</tbody>
</table>
Sequencing was performed on either gel-based or capillary-based automated sequencers. For gel-based sequencing, purified PCR (using ExoSAP; USB Corporation) products were cycle sequenced using a Thermo Sequenase Primer Cycle Sequencing Kit (Amersham Biosciences, Piscataway, NJ) and loaded onto a Li-Cor NEN IR² 4200 global sequencing system (Li-Cor, Lincoln, NE). The sequencing program eSeq version 2.0 was used to read sequences and to check base calls. For capillary-based sequencing, purified PCR products were cycle sequenced using a 1/8 dilution of the manufacturer’s (Applied Biosystems BigDye) sequencing reaction protocol for a 5 µl reaction: 0.25 µl BigDye reagent, 0.875 µl 5x BigDye Buffer, 0.32 µl primer, 1.0 µl template (10–40 ng for 1000 bp product), 2.55 µl water. The sequencing reaction products were loaded onto an ABI 3100 capillary sequencer (Applied Biosystems, Foster City, CA) and analyzed using the program Sequencing Analysis 5.1.1.

Both primer pairs successfully amplified samples taken from known scombrids. The longer fragment was used to generate information on the reference samples to identify those sites that discriminated between species. The internal primer pair was designed after the sequence of the long fragment was known, in order to amplify the shortest possible fragment that included informative sites. For the “unknown” samples (larvae, stomach contents), PCR was performed using the shorter primer pair. In cases where the shorter primer pair did not generate a PCR product, universal COI primers were used as a positive control. If the universal primers generated an amplicon, the sample was inferred to be a non-scombrid. If the universal primers did not result in a successful amplification, the sample was considered too degraded for analysis. The universal primers were designed using Primaclade (Gadberry et al., 2005) with known COI sequences of several marine fish species in GenBank: Mola mola Linnaeus, 1758, Trachurus trachurus Linnaeus, 1758, Hoplostethus japonicus Hilgendorf, 1879, Antigonia capros Lowe, 1843, Enemmlichthys struhsakeri Heemstra and Randall, 1977, A. thazard, E. alletteratus, T. alalunga, K. pelamis, S. scombrus, and T. thynnus. The universal COI primers used were: COIuniv113Fdeg (5´gRg CNg ARC TAA gyC AAC C3´) and COIuniv697R (5´CCR AAg AAT CAg AAB AgR Tg3´).

All sequences were edited using Sequencher version 4.2.2 (Gene Codes Corp.). Edited reference sequences (long fragment) of each species were aligned using the Clustal W program in MacVector version 7.2 (Accelyrs Inc.) to assess intraspecific variation. A consensus sequence of all haplotypes was generated for each species and these representative sequences were aligned to visualize informative interspecific differences using the program MEGA version 3.0 (Kumar et al., 2004). MacClade v. 4.07 (Maddison and Maddison, 2005) was used to assess variability at each base position. For unknown samples, the species identity was inferred by noting where the sample sequence clustered in a UPGMA tree using absolute number of differences between the consensus sequences.

Preliminary analyses suggested either misidentification or introgression in one sample identified as S. regalis based on morphology. To discriminate between misidentification and introgression between S. maculatus and S. regalis, the nuclear ITS-1 region was analyzed in four samples of each of these two species using the primers F-ITS-1 (5´GAG GAA GTA AAA GTC GTA ACA AGG3´) and 5.8SR2 (5´GTG CGT TCG AAR KGT CGA TGA TCA TCA AT3´) (K. Johnson, Virginia Institute of Marine Science, unpubl.). PCR products were cloned into the pCR4-TOPO vector (Invitrogen Corporation, Carlsbad, California) and three clones from each sample were sequenced. This fragment was amplified and sequenced as previously described for the long COI fragment, using the capillary sequencer except 5 µl of Q solution (Qiagen, Valencia, California) was used in the 25 µl reaction and the annealing temperature was 45 °C. The restriction enzyme Sma I (New England BioLabs, Ipswich, Massachusetts) was found to distinguish between S. maculatus and S. regalis based on sequence alignment information. Samples were subsequently distinguished using a 15 µl restriction digestion reaction using 0.5 µl Sma I, 1.5 µl 10× NEB4 10× reaction buffer, 11 µl deionized sterile water, and 2 µl PCR product, digested at 25 °C for 1 hr, then 65 °C for 20 min.
Results

Two amplicons were generated in this study, a long (945 bp) and a short (264 bp) fragment of the COI gene. Within the long fragment there were 279 (30%) variable sites and in the short fragment there were 64 (24%) variable sites. The vast majority (93.7%) of substitutions occurred at the third codon position, while only 5.7% occurred at the first position and 0.7% at the second position. There were no insertions or deletions within the COI regions analyzed.

The long fragment of COI exhibited a wide range of differences between the 17 species analyzed (GenBank Accession nos DQ835818–DQ835957; EF379124). The number of nucleotide differences between consensus sequences of each species ranged from 2 base changes (between T. obesus and T. albacares or T. atlanticus) to 152 (between S. scombrus and S. cavalla) (Table 2). The differences between species in the short fragment ranged from 1 base change (between T. obesus and T. atlanticus and between S. maculatus and S. regalis) to 48 (between S. colias and A. thazard). Within species, variation of the long fragment ranged from 0 in S. brasiliensis and S. regalis to 26 variable sites within S. sarda. Reference samples of K. pelamis, T. albacares, T. obesus, A. rochei, and A. thazard included Atlantic and Pacific individuals, and thus the intraspecific variation observed in these species encompassed any inter-oceanic differences.

The positively identified samples of a species consistently grouped together in a UPGMA tree of all COI sequences in this study. A consensus sequence was generated for each species to serve as a representative of that species in a reference UPGMA tree (Fig. 1). A single S. regalis sample had a COI sequence that was more similar to S. maculatus than S. regalis. To distinguish between misidentification and introgression, the nuclear ITS-1 region was sequenced (GenBank Accession nos EF379101–EF379123). The ITS-1 sequence data indicated that the anomalous sample was indeed S. regalis, suggesting introgression. The COI sequence of this sample was not included in the consensus sequence of S. regalis. The differences between these two species in the ITS-1 region are shown in Figure 2. From the ITS-1 sequence information, a restriction fragment length polymorphism was found between S. regalis and S. maculatus whereby the restriction enzyme SmaI would create 250 bp and 600 bp bands for S. regalis and 280 bp and 600 bp bands for S. maculatus (Fig. 3).

From the consensus sequences, an unambiguous molecular key was developed that allows identification of all 17 western Atlantic scombrids. Positions at which a species has a consistent, unique combination of nucleotide base pairs are indicated in Figure 4. The shorter COI fragment also provided dependable species identification as it includes diagnostic sites with the added advantage of the ability to amplify smaller, degraded DNA fragments. Clustering of an unidentified sample in the UPGMA tree was the quickest method of identification, but in cases where an individual did not clearly group with one species, discriminatory base positions were located in the unknown sequence and compared with the molecular key for identification.

Scombrid larvae from the Florida Straits in the western North Atlantic Ocean were used to test the efficacy of this marker. These individuals were sufficiently large (4.5–12 mm) to be identified morphologically to genus; however, some were damaged, making specific identification based on morphological characters difficult if not impossible. Fifty-two scombrid larvae were identified based on DNA sequence and, when possible, using morphological characters following Richards (2006) and
Table 2. Pairwise distances between western Atlantic scombrids based upon number of nucleotide differences for consensus haplotype COI sequences of each species. Differences in the longer fragment (945 bp) are shown in the lower left portion of the matrix and differences in the shorter fragment (260 bp) are shown in the upper right. The column labeled “Intra” contains the number of variable sites in the long fragment within each of the species in that row (MEGA version 3.0 Kumar, Tamura, Nei, 2004). Species abbreviations are given in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>ALBC</th>
<th>BLFT</th>
<th>BET</th>
<th>YFT</th>
<th>BLKF</th>
<th>AUXT</th>
<th>AUXR</th>
<th>EUTH</th>
<th>SARD</th>
<th>ASOL</th>
<th>CHMK</th>
<th>ATLM</th>
<th>KMCK</th>
<th>SPMK</th>
<th>CERO</th>
<th>SCBR</th>
<th>Intra</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALBC</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>29</td>
<td>25</td>
<td>23</td>
<td>17</td>
<td>23</td>
<td>22</td>
<td>43</td>
<td>36</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>36</td>
<td>7</td>
</tr>
<tr>
<td>BLFT</td>
<td>17</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>29</td>
<td>24</td>
<td>23</td>
<td>17</td>
<td>22</td>
<td>20</td>
<td>42</td>
<td>34</td>
<td>37</td>
<td>33</td>
<td>33</td>
<td>34</td>
<td>15</td>
</tr>
<tr>
<td>BET</td>
<td>14</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>27</td>
<td>23</td>
<td>23</td>
<td>16</td>
<td>23</td>
<td>17</td>
<td>44</td>
<td>36</td>
<td>36</td>
<td>32</td>
<td>32</td>
<td>34</td>
<td>15</td>
</tr>
<tr>
<td>YFT</td>
<td>14</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>29</td>
<td>24</td>
<td>23</td>
<td>17</td>
<td>23</td>
<td>18</td>
<td>43</td>
<td>37</td>
<td>35</td>
<td>33</td>
<td>33</td>
<td>34</td>
<td>17</td>
</tr>
<tr>
<td>BLKF</td>
<td>12</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>28</td>
<td>24</td>
<td>26</td>
<td>16</td>
<td>24</td>
<td>19</td>
<td>46</td>
<td>38</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>37</td>
<td>8</td>
</tr>
<tr>
<td>AUXT</td>
<td>98</td>
<td>97</td>
<td>95</td>
<td>95</td>
<td>6</td>
<td>26</td>
<td>18</td>
<td>22</td>
<td>22</td>
<td>48</td>
<td>44</td>
<td>28</td>
<td>31</td>
<td>31</td>
<td>33</td>
<td>33</td>
<td>8</td>
</tr>
<tr>
<td>AUXR</td>
<td>88</td>
<td>87</td>
<td>86</td>
<td>88</td>
<td>23</td>
<td>26</td>
<td>16</td>
<td>20</td>
<td>20</td>
<td>44</td>
<td>42</td>
<td>26</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>14</td>
</tr>
<tr>
<td>EUTH</td>
<td>89</td>
<td>89</td>
<td>86</td>
<td>85</td>
<td>90</td>
<td>89</td>
<td>92</td>
<td>18</td>
<td>19</td>
<td>23</td>
<td>39</td>
<td>40</td>
<td>28</td>
<td>27</td>
<td>26</td>
<td>28</td>
<td>8</td>
</tr>
<tr>
<td>SKJT</td>
<td>80</td>
<td>85</td>
<td>82</td>
<td>84</td>
<td>80</td>
<td>65</td>
<td>58</td>
<td>73</td>
<td>15</td>
<td>18</td>
<td>43</td>
<td>36</td>
<td>26</td>
<td>31</td>
<td>31</td>
<td>32</td>
<td>17</td>
</tr>
<tr>
<td>SARD</td>
<td>82</td>
<td>77</td>
<td>77</td>
<td>80</td>
<td>79</td>
<td>88</td>
<td>84</td>
<td>83</td>
<td>85</td>
<td>21</td>
<td>35</td>
<td>34</td>
<td>26</td>
<td>27</td>
<td>27</td>
<td>28</td>
<td>26</td>
</tr>
<tr>
<td>ASOL</td>
<td>100</td>
<td>99</td>
<td>95</td>
<td>97</td>
<td>96</td>
<td>107</td>
<td>100</td>
<td>108</td>
<td>86</td>
<td>38</td>
<td>35</td>
<td>27</td>
<td>28</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>16</td>
</tr>
<tr>
<td>CHMK</td>
<td>130</td>
<td>125</td>
<td>127</td>
<td>126</td>
<td>130</td>
<td>126</td>
<td>118</td>
<td>126</td>
<td>120</td>
<td>115</td>
<td>138</td>
<td>25</td>
<td>43</td>
<td>37</td>
<td>37</td>
<td>38</td>
<td>16</td>
</tr>
<tr>
<td>ATLM</td>
<td>119</td>
<td>120</td>
<td>123</td>
<td>122</td>
<td>126</td>
<td>126</td>
<td>126</td>
<td>124</td>
<td>127</td>
<td>117</td>
<td>134</td>
<td>82</td>
<td>45</td>
<td>33</td>
<td>33</td>
<td>34</td>
<td>15</td>
</tr>
<tr>
<td>KMCK</td>
<td>132</td>
<td>129</td>
<td>132</td>
<td>128</td>
<td>132</td>
<td>137</td>
<td>130</td>
<td>134</td>
<td>124</td>
<td>107</td>
<td>118</td>
<td>143</td>
<td>152</td>
<td>23</td>
<td>24</td>
<td>23</td>
<td>9</td>
</tr>
<tr>
<td>SPMK</td>
<td>115</td>
<td>114</td>
<td>113</td>
<td>110</td>
<td>118</td>
<td>131</td>
<td>127</td>
<td>123</td>
<td>124</td>
<td>105</td>
<td>118</td>
<td>133</td>
<td>133</td>
<td>133</td>
<td>107</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>CERO</td>
<td>118</td>
<td>118</td>
<td>117</td>
<td>114</td>
<td>122</td>
<td>132</td>
<td>126</td>
<td>124</td>
<td>126</td>
<td>107</td>
<td>119</td>
<td>135</td>
<td>134</td>
<td>108</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>SCBR</td>
<td>124</td>
<td>121</td>
<td>120</td>
<td>116</td>
<td>124</td>
<td>135</td>
<td>128</td>
<td>127</td>
<td>126</td>
<td>110</td>
<td>115</td>
<td>133</td>
<td>138</td>
<td>114</td>
<td>18</td>
<td>20</td>
<td>0</td>
</tr>
</tbody>
</table>
Nishikawa and Rimmer (1987). From these guides, useful morphological characters were: forebrain pigment and ventral pigment spot in *K. pelamis*, lower jaw pigmentation in *E. alletteratus*, and lateral tail pigmentation in *A. thazard*. Morphological identification to species level was possible for 18 larvae, and in each case, morphological assignment supported genetic identification. The remaining 34 unknown larvae were identified to species solely by noting their placement with known samples in a UPGMA tree of long fragment consensus sequences (Fig. 5).

To test the efficacy of the marker on degraded tissues, the short fragment was amplified from putative scombrids found in billfish stomach contents. The shorter fragment of COI was analyzed in the stomach contents as these tissues are generally deteriorated and therefore the DNA may also be degraded. When the sequences of these COI fragments were aligned to the known reference samples, nine samples clustered with *A. rochei* in the UPGMA tree (Fig. 6). Two samples from stomachs of billfishes caught in Hawaii did not cluster with any of the scombrid species. The search engine Basic Local Alignment Search Tool (BLAST) (National Center for Biotechnology Information) was used to find the closest match between these samples and species in GenBank (National Center for Biotechnology Information). One sample shared 84% identity with the COI gene of *Myripristis berndti* (blotcheye soldierfish) (Jordan and Evermann, 1903) and the other sample had 83% identity with the COI gene of *H. japonicus* (flintperch).
Discussion

Both the long and short COI fragments met the two requirements of a good molecular marker for scombrid identification: consistent interspecific differences and minimal intraspecific variation. This molecular key was developed to unambiguously identify all scombrids occurring in the western Atlantic Ocean, several of which have a circumtropical distribution. To evaluate the applicability of the marker outside the Atlantic, several Pacific conspecifics of some of the circumglobal species covered in this study were sequenced to determine if intraspecific differences exist between conspecifics from different ocean basins. Previous studies have shown evidence of inter-oceanic differences in bigeye tuna and albacore based on other gene regions (Chow and Ushiama, 1995; Alvarado Bremer et al., 1998; Chow et al., 2000). The diagnostic sites of the COI region still allowed for unambiguous identification of...
Figure 4. Molecular key of interspecific differences in the shorter COI fragment between consensus sequences. The sites that are useful in distinguishing very closely related species (i.e., *Thunnus albacares*/*Thunnus obesus*, *Thunnus atlanticus*/*Thunnus obesus*, and *Scomberomorus regalis*/*Scomberomorus maculatus*) have an asterisk and are bolded. Species abbreviations are given in Table 1.

Circumglobal species between ocean basins. This was not unexpected given the high level of conservation in COI (Brown, 1985). The success of this marker in other ocean basins suggests that its utility may easily be extended to other scombrid species that occur elsewhere.

While the COI marker has proven to be effective for species identification of scombroids it, like other molecular markers has limitations. Sequencing has drawbacks including limitations imposed by cost and time, but this technology is being improved...
upon continually, making it quite an attractive high resolution technique for species identification. An alternative method using PCR/RFLP analysis of an amplified gene region (Chow and Inoue, 1993; Daniel and Graves, 1994; McDowell and Graves, 2002; Chow et al., 2003) is practical, but it becomes more difficult with increasing number of species to find unique or unambiguous fragment patterns that will distinguish each species. Similarly, a multiplex assay increases speed and decreases cost of analysis, but requires the design of many species-specific primers, which would be challenging given the number of species in this study.

Another concern with using only a mitochondrial marker is the possibility that introgression may lead to the misidentification of samples. Mitochondrial introgres-
sion has been previously reported in scombrids. The mitochondrial genome of the albacore, *T. alalunga*, has introgressed onto the Pacific bluefin tuna *T. orientalis* genetic background within the Pacific at a high frequency (98%) (Chow and Kishino, 1995) and at a low frequency (5%; 6.8%) in the eastern Atlantic/Mediterranean in Atlantic bluefin tuna *T. thynnus* (Vinas et al., 2003; Carlsson et al., 2004). Additionally, mitochondrial introgression has also been reported in the genus *Scomberomorus*. Banford et al. (1999) posited that the *S. regalis* mitochondrial genome has introgressed into *S. maculatus*. In the present study, one *S. regalis* sample clustered with *S. maculatus*, suggesting either misidentification or introgression. Subsequent analysis of the nuclear ITS region demonstrated that introgression may be bidirectional, as the observed introgression is in the opposite direction as that seen in the study by Banford et al. (1999). Clearly there is a need for further analysis, including more *S. regalis* samples, to adequately resolve this issue. Until then, a nuclear marker should be employed in addition to the COI marker to verify species identity of any putative *S. regalis* or *S. maculatus* samples. The RFLP in ITS-1 found in this study, which generated different sized bands between these species, could be used for future studies as a quick diagnostic for introgression.

The demonstrated ability of this key to provide species identifications of scombrid larvae and scombrid remains in stomach contents indicates its potential for use
in population studies, forensic analyses, and early life history investigations. This marker has numerous applications, from verifying that samples are indeed the correct species in population studies employing analysis of nuclear microsatellite loci, to providing species level identification of fillets that are being sold illegally, which is critical for management enforcement (Lopez and Pardo, 2005). While DNA barcoding has been proven successful for species identification (Hebert et al., 2003; Ward et al., 2005), the present study demonstrates the potential of using a molecular marker to identify degraded tissue, a notoriously difficult task. In cases where morphological identification is not possible, a molecular key provides a reliable and relatively rapid means of unambiguously identifying scombrid species.

Acknowledgments

Many thanks to all those who generously provided tissue or DNA samples: D. Richardson, R. Cowen, B. Collette, D. Kerstetter, K. Neill, J. Walter, O. Sanjur, T. Orrell, A. Mahon, K. Carpenter, P. D’Antoni, and H. Banford. A. Pao assisted with stomach content analyses and E. Brasseur facilitated taking pictures of the larvae. Funding was provided through Cooperative Marine Education and Research (CMER) and NOAA Highly Migratory Species awards. Virginia Institute of Marine Science contribution no. 2823.

Literature Cited


Date Submitted: 10 July, 2006. Date Accepted: 14 February, 2007.