Authentication of Canned Tuna and Bonito by Sequence and Restriction Site Analysis of Polymerase Chain Reaction Products of Mitochondrial DNA

Jeffrey L. Ram,* Michal L. Ram, and Fadi F. Baidoun

Departments of Physiology and Biochemistry, Wayne State University, Detroit, Michigan 48201

Methods of authenticating already canned fish were developed, using polymerase chain reaction (PCR) followed by sequencing and restriction site analysis. The canning process degrades DNA to fewer than 123 base pairs (bp) in length. Therefore, degenerate PCR primers were designed to amplify short (<123 bp) mitochondrial cytochrome *b* gene sequences known to differ at specific nucleotides among the species of interest. Sequences of canned tuna (*Thunnus albacares*, *Thuunus alalunga*, and *Katsuwonus pelamis*), bonito (*Euthynnus affinis*), and frigate mackerel (*Auxis thazard*) were reproducibly identified, and were used to determine which species or whether more than one species was present in individual cans. Restriction site analysis of two amplified regions of the cytochrome *b* gene demonstrated a faster and less expensive method than sequencing for distinguishing PCR products of different species. Thus, restriction site analysis of PCR products can be used in conjunction with sequencing to authenticate species in canned fish products.

Keywords: *Polymerase chain reaction; authentication; fish; tuna; bonito; Thunnus; Katsuwonus; Euthynnus; mitochondrial DNA; cytochrome b gene*

1. INTRODUCTION

Authentication of food products is important for labeling and assessment of value. The species in a canned product often determines what it can be called by the manufacturer, its price, and even whether the consumer is willing to buy it. Authentication is therefore necessary to avoid unfair competition and to assure consumers of accurate labeling.

For example, the European community has drawn a distinction between tuna and bonito, according to the list in Table 1. Although several species classified as bonito by the European community have common names of "tuna" and may be similar in taste and texture to authentic tuna, they may no longer be labeled as tuna. The bonito species are not as popular with consumers as tuna, and hence this distinction results in a much lower market price for bonito. Accurate authentication of these species is necessary to enforce this distinction.

This presents a problem in canned fish because the canning process obliterates many morphological distinctions and some biochemical markers used for species identification, such as allozymes. We therefore sought to develop a method of authenticating canned fish based on genomic differences. DNA sequences of the mitochondrial cytochrome *b* gene are known to differ by characteristic nucleotide substitutions among species of tuna and bonito (Block et al., 1993). The present paper demonstrates that, despite considerable DNA degradation present in canned fish products, polymerase chain reaction (PCR) amplification of selected regions of the mitochondrial cytochrome *b* gene can be used in conjunction with sequencing or restriction site analysis to provide a reliable means of identifying species in canned fish products.

* Author to whom correspondence should be addressed (Department of Physiology) [telephone (313) 577-1558; fax (313) 577-5494; e-mail jeffram@med.wayne.edu].

Table 1. Classification of Scombrid Species by the European Community (from the Annex to the *Off. J. Eur. Communities* 1992, June 17, L 163/4)

- **tuna**
  - species of the genus *Thunnus*
    - albacore or long-finned tuna (*T. alalunga*)
    - yellowfin tuna (*T.* (*neothunnus*) *albacares*)
    - bluefin tuna (*T. thynnus*)
    - bigeye tuna (*T.* (*parathunnus*) *obesus*)
    - other species of the genus *Thunnus*
  - skipjack or stripe-bellied tuna
    - [*Katsuwonus* (*Euthynnus*) *pelamis*]
- **bonito**
  - species of the genus *Sarda*
    - Atlantic bonito (*S. sarda*)
    - Pacific bonito (*S. chilliensis*)
    - oriental bonito (*S. orientalis*)
    - other species of the genus *Sarda*
  - species of the genus *Euthynnus* with the exception of *E.* (*K.*) *pelamis*
    - Atlantic little tuna (*E. alleteratus*)
    - eastern little tuna (bonito, *E. affinis*)
    - black skipjack (*E. lineatus*)
    - other species of the genus *Euthynnus*
  - species of the genus *Auxis*
    - frigate mackerel (*A. thazard*)
    - A. rochei

2. MATERIALS AND METHODS

2.1. Samples. Fish samples were obtained from canned tuna or fresh fish purchased at local supermarkets or various species of canned or fresh fish supplied by Unicord Public Co., Ltd. (Bangkok, Thailand). Supermarket canned tuna products were all Starkist brand and included Fancy Albacore Solid White tuna in water (lot SASGH X179C), Select Solid Light tuna in water (labeled as yellowfin tuna; S3SGB X146C), Chunk Light tuna in water (species not identified; CBSCC X174C), Chunk Light tuna in vegetable oil (species not identified; KB0ZF F161C), and Chunk White tuna in water (labeled pure albacore tuna; KADBD X151C). Canned fish products from Unicord were identified either as tuna in oil (lot U31DG W7J3G) or as bonito in oil (lot BTCO 3807). Canned fish commonly (and in all cans that we analyzed) consists of several (two to four) compact pieces, colloquially referred to as "chunks", the term that we will also use here as it is readily

S0021-8561(95)00822-3 CCC: $12.00

understood and refers to a distinct piece of fish of modest size (2−8 cm in each dimension), all parts of which are likely to have come from the same individual fish. Fresh yellowfin tuna (*Thunnus albacares*) and common mackerel (*Scomber scombrus*) were purchased at local supermarkets; fresh skipjack tuna (*Katsuwonus pelamis*) and eastern little tuna (bonito, *Euthynnus affinis*, two different samples) were shipped from Thailand by Unicord. Chunks of fresh fish were preserved in alcohol (ethanol for locally purchased fish; 2-propanol for fish shipped from Thailand) upon purchase or prior to shipping. Prior to use, canned tuna products were also preserved in ethanol.

2.2. PCR. DNA was extracted from 30−90 mg of each fish sample, assessed for quality and size distribution, and amplified by PCR with standard methods (Dougherty et al., 1996). Because considerable DNA degradation occurs as a result of the canning process (see Results and Discussion), PCR primers were designed to amplify only short species-distinctive sequences. Degenerate primers were designed to amplify bases 335−394 and 211−269 in the sequences of Block et al. (1993) (referred to in this paper as regions R335 and R211), in which sequence heterogeneity was sufficient to distinguish a large number of fish species. For R335, forward primer (TUNA334F) was 5′-TAGGGATCCTYCTHTCIGCAGTMCCMTAYGT and reverse primer (TUNA395R) was 5′-GGTCTCAGGAAGTGGAAKGCRAAGAAYCGG, where I = inosine, Y = T/C, M = A/C, H = A/C/T, R = G/A, and K = T/G, giving a PCR product 121 bases long. For R211, forward primer (TUNAFOR) was 5′-GGGAATTCCTMTACAAAGAAACMTGAAACA and reverse primer (TUNAREV) was 5′-DAGGGATCCTCAGAANGAYATYTGTCCTCA), giving a PCR product 119 bases long. Primers were manufactured by Operon Technologies, Inc. (Alameda, CA).

Only the R335 PCR products were used for sequence analysis using the techniques described below, whereas both R335 and R211 PCR products were used for restriction site studies. Sequences of R211 PCR products were obtained during the pilot stage of this project, using techniques that were less reliable and are therefore not described in detail. Brief mention of some of the results of these pilot studies and their implications for the robustness of these procedures will be given under Results and Discussion.

2.3. Electrophoresis, DNA Cleanup, Sequencing, and Sequence Identification of PCR Product. The presence of PCR product was confirmed by electrophoresis and ethidium bromide staining of 3 μL of each sample on a 2.5% agarose (FMC NuSieve 3:1) gel, with either a 123 base pair (bp) DNA ladder or DNA Ladder-Low (DNA molecular weight marker XI, Boehringer Mannheim, Indianapolis, IN) for standard. DNA was cleaned up using Qiagen Kit 28104, according to the manufacturer's instructions. The Qiagen kit gives excellent recovery of short (<125 bp) PCR products, which several alternative DNA cleanup procedures did not. The quality and amount of cleaned DNA was evaluated on agarose gels, using staining density relative to a known amount of DNA Ladder-Low as a quantity standard. DNA samples amplified from R335 (100 ng/10 μL) were sequenced in both directions by automated dye terminator sequencing at the University of Michigan DNA Sequencing Core (Ann Arbor, MI), using our custom primers TUNA334F and TUNA395R or an alternate reverse primer (TUNA396R) with the sequence 5′-GGTCTAGGAAGTGGAAKGCRAAGAATCG. Due to degeneracy of the primers, primer concentrations 8−32 pmol/20 μL were used instead of the standard ABI cycle sequencing condition of 1 pmol/20 μL.

The sequencer identified most bases unambiguously; however, the sequencer gave an "N" when two peaks in the chromatogram were superimposed and of approximately equal heights. In such cases, chromatograms were assessed visually to identify the two bases, and the "N" was replaced according to the following convention: R = G/A; Y = T/C; K = G/T; M = A/C; S = G/C; W = A/T. If two bases could not be clearly distinguished, the "N" was retained.

Resultant sequences from TUNA334F and the inverse complement of sequences obtained with TUNA395R or TUNA396R were compared to each other and to known tuna and bonito sequences in Genbank using the FASTDB Program of the Intelligenetics Suite (Mountain View, CA; Brutlag et al., 1990) and the BLASTN program of NCBI (Altschul et al., 1990). The programs aligned sequences, determined nearest matches, and in general graphically illustrated the quality of the sequencing over the entire sequence.

2.4. Restriction Site Analysis of PCR Products. For region R335, the enzymes *Rsa*I, *Kpn*I, and *Hin*fI were used to cut PCR products from fresh yellowfin, bonito, skipjack tuna, and common mackerel and from canned fish samples containing several of these species, as well as (based on sequence results to be described) albacore tuna and frigate mackerel. For region R211, the enzymes *Rsa*I and *Mbo*II were used to analyze samples from fresh yellowfin, bonito, and skipjack tuna and from canned samples of these species and frigate mackerel. Figure 1 shows the restriction sites in mitochondrial cytochrome *b* regions R335 and R211 that these enzymes would be expected to cut.

Restriction enzyme incubations were run in 20 μL or 50 μL volumes, consisting respectively of 2 or 5 μL of the appropriate 10× buffer (supplied by the manufacturers with their respective enzymes), 0.5−2 μL of restriction enzyme solution to give final concentrations (depending on the enzyme) of 6−24 units/assay, 3 or 6 μL of PCR product or purified PCR DNA solution (sufficient concentration to give a visible band), sometimes 2 or 5 μL of 1 mg/mL bovine serum albumin, and the remainder pure water. PCR solutions were compatible with the restriction enzyme buffer systems for *Rsa*I and *Mbo*II, for which purification of PCR product was unnecessary. For experiments with *Hin*fI, *Kpn*I, and *Rsa*I on R335 PCR products, amplified DNA was purified according to the Qiagen method before restriction analysis. Negative controls were incubated identically but with water in place of the restriction enzyme. Incubation was for 1 h at 37 °C, followed by agarose gel electrophoresis of 20 μL of the products.

3. RESULTS AND DISCUSSION

3.1. DNA Size Distribution and Yield of PCR Products. Figure 2A compares the size distributions of DNA isolated from fresh fish (yellowfin tuna, bonito, and common mackerel) and from canned tuna. A broad range of sizes up to thousands of base pairs was consistently present in the fresh fish samples, while the DNA extracted from canned tuna samples was consistently smaller than the 123 bp standard. A similar size range of DNA was present in canned samples of bonito and frigate mackerel (data not shown). These results indicate that considerable DNA degradation has occurred in canned fish products, compared to fresh fish. The implication of this observed DNA size distribution for the use of PCR in authenticating canned fish products is that the sequences to be targeted for PCR amplification should be no larger than 123 bases.

Results of amplifying DNA from both canned and fresh fish are illustrated in Figures 2B and 3. Fresh fish gave visible PCR products in all cases with primers for both region R335 (Figures 2B and 3B) and region R211 (Figure 3A). These data with fresh fish serve as positive controls demonstrating that PCR conditions were appropriate for obtaining amplified products if target sequences were present, as expected for DNA of fresh fish. Negative controls (no DNA added; lane 2 of Figure 2B) showed that PCR products were not due to contaminating DNA. Furthermore, DNA sequencing, described later, gave the expected species-distinctive sequences for all fresh fish samples. In addition to the positive results shown here with yellowfin tuna, skipjack tuna, bonito, and common mackerel, similar results have also been obtained using these primers with DNA from other fresh fish, including species as distant as gobies and salmon (data not shown).

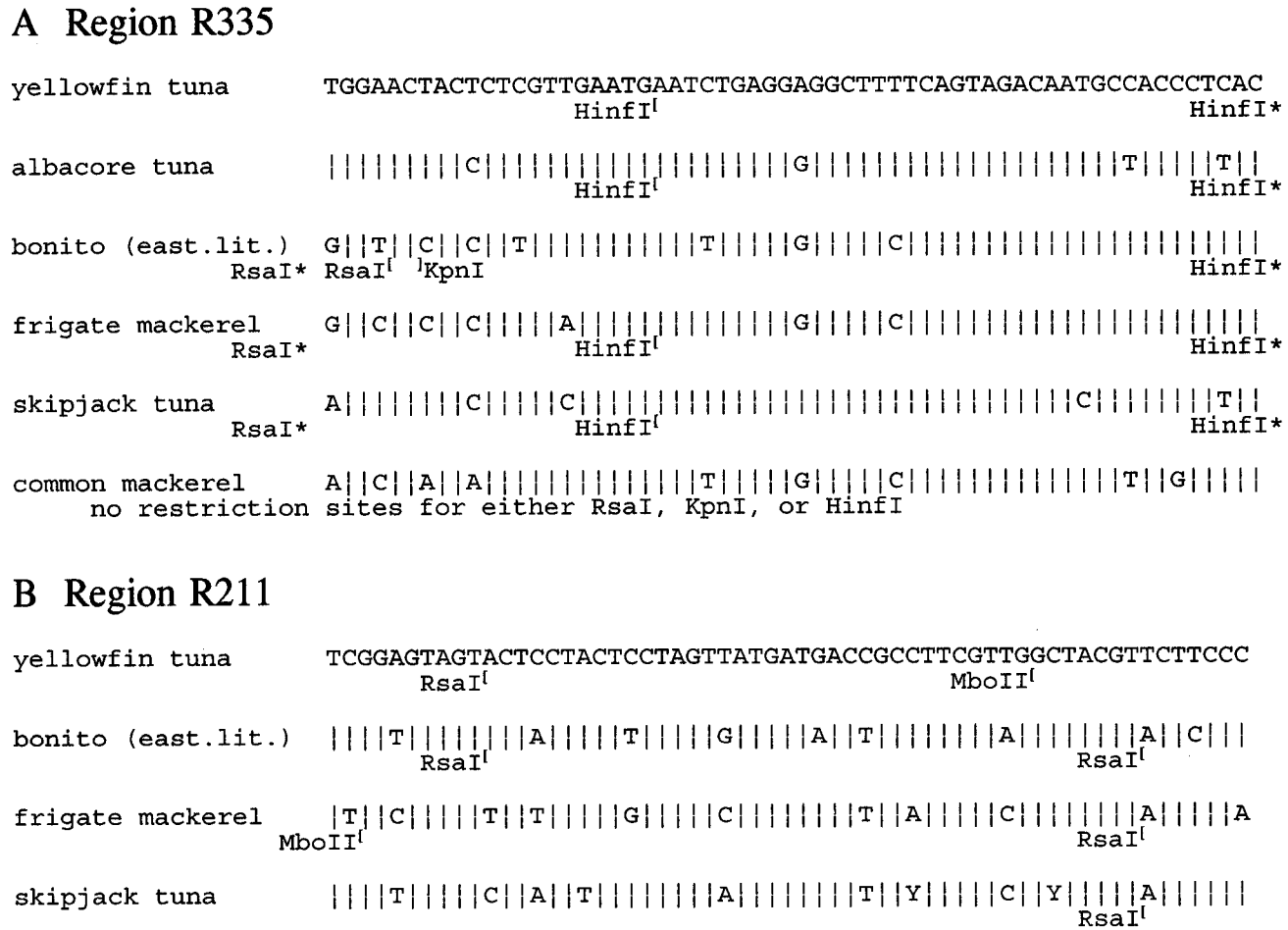

Figure 1. Restriction sites in regions R335 and R211 of the mitochondrial cytochrome *b* gene. Flanking primer regions (not shown) would add approximately 30 bases to each end. (A) *Rsa*I, *Kpn*I, and *Hin*fI sites in region R335. (B) *Rsa*I and *Mbo*II sites in region R211. Vertical lines represent nucleotides identical to the yellowfin tuna sequence. * indicates a restriction site in the primer region; however, since degenerate primers were used, the restriction site may not be present in the PCR product.

With DNA from canned fish, visible PCR products were obtained from 10 of the 11 cans of fish that were analyzed for this study (Table 2). The can that did not yield a PCR product (Figure 2B, lane 5) contained albacore packed in water, as labeled by the manufacturer. This can gave readily measurable amounts of DNA as large as 123 bp (Figure 2A, lane 2), indicating that the problem was not lack of DNA. Other possible causes for the lack of PCR product in this sample are discussed later, in section 3.4. Two other water-packed samples (yellowfin tuna, lot S3SGB X146C and chunk light tuna, lot CBSCC X174C; Figure 2B, lanes 6 and 7, respectively) gave a lower PCR yield than other samples; however, another water-packed sample gave a good PCR yield (albacore tuna, lot KADBD X151C, Figure 2B, lane 9). All oil-packed samples gave a good yield of PCR product (e.g., Figure 2B, lanes 8 and 12, and other samples not shown). Thus, despite the considerable DNA degradation present in all canned fish samples, PCR products could be obtained from a variety of species packed under several different canning conditions.

3.2. Sequence Data and Species Identification. *3.2.1. Fresh Fish.* Sequences of the R335 region were obtained from all samples for which PCR yielded measurable DNA. Sequences of fresh fish provided known sequences to serve as standards to which automatic sequencer chromatogram patterns could be compared. Sequences obtained from fresh yellowfin and skipjack tuna are shown in Figure 4, and those from fresh bonito are shown in Figure 5. In all cases the sequences of fresh fish samples identified the appropriate species, although, in all but one sample, at least one base disagreed with or was ambiguous compared to the reference sequence from Genbank (Block et al., 1993). Disagreement with Genbank could be due to sequencing inaccuracies; however, since all illustrated sequences are based on multiple sequencing runs (always in both directions and in some cases on multiple samples), a more likely explanation is intraspecific variation. Nevertheless, since in most cases the species of interest differ on several bases, identification of species can still be based on greatest similarity to the reference sequences, if not complete identity.

3.2.2. Canned Fish. Figure 4 illustrates sequence data used to distinguish canned yellowfin tuna (lot S3SGB X146C, specifically labeled yellowfin) from chunk light tuna (lot KB0ZF F161C, species not labeled). In Figure 4A, the canned yellowfin sample agreed with the Genbank reference sequence at all bases that were sequenced unambiguously in both directions. At four bases in which the sequence was ambiguous, the sequence was identical to the reference sequence for at least one direction of sequencing. Overall, the canned yellowfin sample was therefore decisively confirmed as yellowfin tuna. Similarly, in Figure 4B the fish sample from lot KB0ZF F161C, with somewhat poorer sequence data, was decisively identified as skipjack tuna. In this and several other samples (e.g. Figure 5, sequence 6), poor sequence data were obtained in the right half of

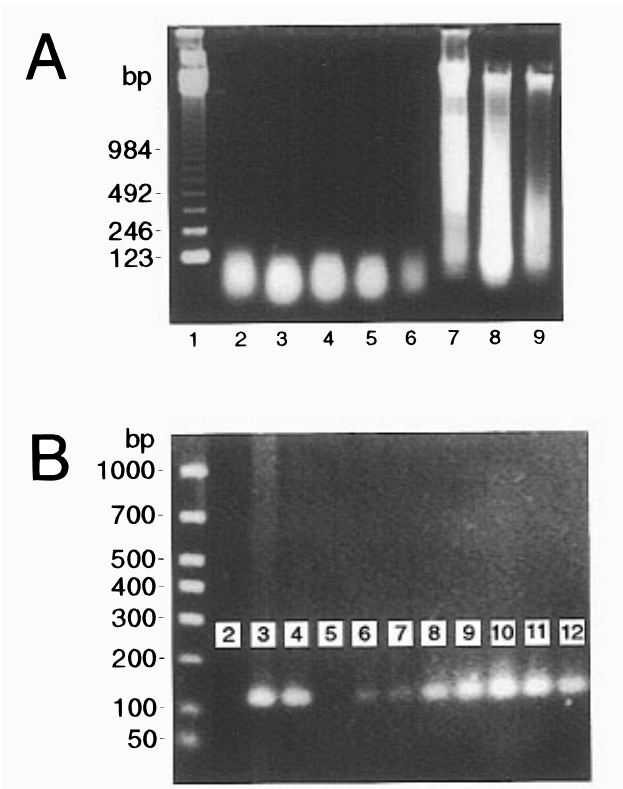

Figure 2. Electrophoretic analysis of purified DNA and PCR products. Samples were electrophoresed on 2.5% agarose gels and stained with ethidium bromide. (A) DNA purified from canned tuna and fresh fish. Lanes and samples were (1) 123 bp ladder; (2−6) canned tuna (2, lot SASGH X179C; 3, lot S3SGB X146C; 4, lot CBSCC X174C; 5, lot KB0ZF F161C; and 6, lot KADBD X151C); (7) fresh mackerel (*S. scombrus*); (8) fresh yellowfin tuna (*T. albacares*); and (9) fresh bonito (*E. affinis*). (B) PCR products obtained using primers TUNA334F and TUNA395R. Lanes and samples were (1) DNA Ladder-Low (molecular weights as marked); (2) no DNA negative control; (3) fresh yellowfin tuna (*T. albacares*); (4) fresh mackerel (*S. scombrus*); (5−9) canned tuna (5, lot SASGH X179C; 6, lot S3SGB X146C; 7, lot CBSCC X174C; 8, lot KB0ZF F161C; and 9, lot KADBD X151C); (10) fresh bonito (*E. affinis*); (11) fresh skipjack tuna (*K. pelamis*); and (12) canned bonito (lot BTCO 3807).

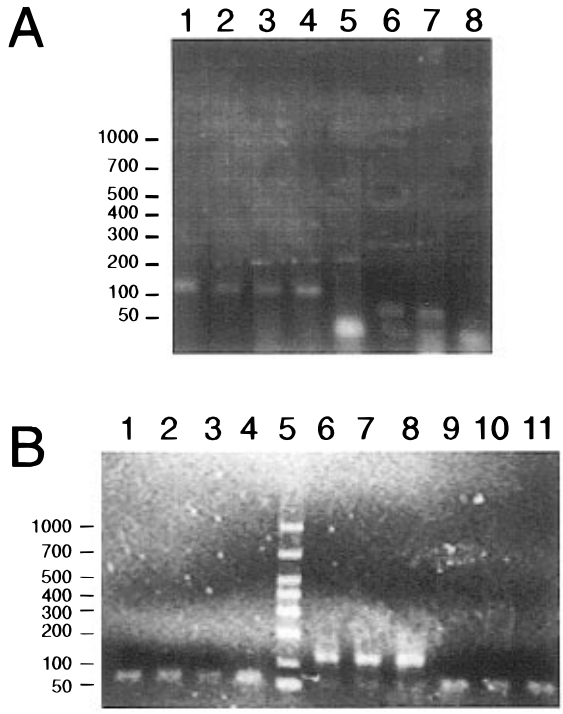

Figure 3. PCR products and results of restriction enzyme treatment, electrophoresed on 2.5% agarose gels and stained with ethidium bromide. (A) PCR product from cytochrome *b* region R211. Sample identities and restriction enzymes present for each lane are as follows: (1) fresh bonito 1, none; (2) fresh yellowfin tuna, none; (3) can 8 yellowfin tuna, none; (4) can 11 bonito, none; (5) fresh bonito 1, *Rsa*I; (6) fresh yellowfin tuna, *Rsa*I; (7) can 8 yellowfin tuna, *Rsa*I; and (8) can 11 bonito, *Rsa*I. Indicated sizes are based on markers run in another part of the gel. (B) PCR product from cytochrome *b* region R335. Lane 5 contains a DNA ladder of sizes shown to the left. All samples were treated under identical conditions with the enzyme *Hin*fI. Sample identities are as follows: (1) fresh skipjack tuna; (2) can 10 frigate mackerel; (3) can 7 skipjack tuna; (4) can 7 skipjack tuna; (6) can 9 bonito; (7) can 9 bonito; (8) can 9 bonito; (9) can 6 yellowfin tuna; (10) can 6 yellowfin tuna; (11) can 6 yellowfin tuna.

the sequence when using the reverse primers (usually poor from bases 371 to 393, but not as extensively in Figure 4B); however, reliable sequence was usually obtained in both directions from bases 338 to 370, providing enough species-distinctive bases to identify the species of origin. Furthermore, the right half of the sequence in the forward direction provided further corroborating data.

3.2.3. Reproducibility and Reliability. To assess reproducibility and reliability, multiple analyses were done on a number of samples from several cans. The set of sequences shown in Figure 5, from one of these cans, was chosen to illustrate some of the possible problems that may be encountered and is thus a *worst* case example, which nevertheless illustrates the reliability and usefulness of the methods. For this can, a total of 10 sequences were done, including 2 replicate sequences, sequences in both directions, sequences from two different samples from two chunks (defined as in section 2.1), and sequences from three different chunks.

3.2.3.a. Replicate Sequencing. With the forward primer (sequences 1 and 2), ambiguities were few and present for different bases, and in all cases, ambiguities in one sequence were consistent with the sequence in the other. For two other samples subjected to replicate sequencing with forward primer (not shown), sequences were fully consistent between both runs, except for one base (base 356 in a skipjack sample) that is invariant among these species and hence can be ignored with respect to canned fish identification. With the reverse primer (sequences 8 and 9), difficulty was encountered in the right half of the sequence. The base that disagreed (384) is invariant among the species of interest and hence can be ignored with respect to species identification. For five other samples subjected to replicate sequencing with reverse primers, there were no bases in the left half where the two sequences disagreed, thus providing ample reliable information for species identification. It was usually possible to tell if sequence data for the right half were poor from the chromatogram or by observing whether half or more of the bases in a span of 5−10 bases did not agree with any reference sequence or the forward sequence from the same sample. Replicate sequencing showed that, despite minor ambiguities in the forward direction and readily identifiable poor sequence data for a portion of the reverse complement, the PCR and sequencing procedures gave data that could be used reliably to identify the species of any given sample.

3.2.3.b. Multiple Samples from the Same Chunk. As illustrated by sequences 1 and 3 and sequences 6 and 7 of Figure 5, sequences from multiple samples of the

Table 2. Sequence Identity and Restriction Enzyme Fragments of PCR Products[a]

		size of restriction enzyme fragments (approx bp)				
		region R211		region R335		
source of PCR product[b]	species (based on sequencing)	*Rsa*I	*Mbo*II	*Rsa*I	*Kpn*I	*Hin*fI
can 1 (SAS)	no PCR product					
can 2 (S3S)	*T. albacares*				120	70
can 3 (CBS)	*K. pelamis*				120	70
can 4 (KB0)	*K. pelamis*				120	70
can 5 (KAD)	*T. alalunga*				120	70
can 6 (U31)	*T. albacares*				120	70
can 7 (U31)	*K. pelamis*	80, 40		120	120	70
can 8 (U31)	*T. albacares*	80, 40				
can 9 (BTC)	*E. affinis*				120, 80	120
can 10 (BTC)	*A. thazard*[c]	80, 40		120	120	70
can 11 (BTC)	*E. affinis*	40	120	120, 80		
fresh yellowfin	*T. albacares*	80, 40	80, 40	120	120	70
fresh bonito 1	*E. affinis*	40		80	80	120
fresh bonito 2	*E. affinis*			80	80, 40?	120
fresh skipjack	*K. pelamis*	80, 40		120	120	70
fresh mackerel	*S. scombrus*			120	120	120

[a] Blank spaces indicate samples/enzymes not tested. ? means a very faint band. [b] Codes in parentheses are the first three letters or numbers of the lot number. [c] Can 10 also contained *E. affinis*; however, only PCR products from *A. thazard* were subjected to restriction site analysis from can 10.

A

```
Thunnus albacares                TGGAACTACTCTCGTTGAATGAATCTGAGGAGGCTTTTCAGTAGACAATGCCACCCTCAC
                                 |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
canned yellowfin, forward        TGGAAYTACTCTCGTTGAATGAATCTGAGGAGGCTTTTCAGTAGACAATGCCACCCTCAC
                                 ||||||+|||||||||||||||||||||||||||||  |||||||||||||||||||||  |
canned yellowfin, reverse        TGGAACTACTCTCGTTGAATGAATCTGAGGAGGCTTKTCAGTAGACAATGCCACCCTGAG
                                 |||||  |||||||||||||||||||||||||||||  |||||||||||||||||||+|+
fresh yellowfin, forward         TGGAAYTACTCTCGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAATGCCACCCTCAC
                                 |||||+|||||||||||||||||||||||||||||*|||||||||||||||||||||||
fresh yellowfin, reverse         TGGAACTACTCTCGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAATGCCACCCTCAC
```

B

```
Katsuwonus pelamis               AGGAACTACCCTCGTCGAATGAATCTGAGGAGGCTTTTCAGTAGACAACGCCACCCTTAC
                                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
lot KB0ZF F161C, forward         AGGAACTACCCTCGTCGAATGAATCTGAGGAGGCTTTTCAGTAGACAATGCCACCCTTAC
                                 |||||||||||||||||||||||||||||||||||||||||||  |     |||||
lot KB0ZF F161C, reverse         AGGAACTACCCTCGTCGAATGAATCTGAGGAGGCTTTTCAGTAGACWAGAAYACCCTAGA
                                 ||||     |||||||||||||||||||||  ||||||||||||||  |++++|||||+++
fresh skipjack, forward          AGGANNNNCCCTCGTCGAATGAATCTGAGGGGGCTTTTCAGTAGACNACGCCACCCTTAC
                                 ||||++++|||||||||||||||||||||*|||||||||||||||+||||||||||||
fresh skipjack, reverse          AGGAACTACCCTCGTCGAATGAATCTGAGGGGGCTTTTCAGTAGACAACGCCACCCTTAN
```

Figure 4. Comparison of canned and fresh fish samples to reference sequences for region R335. Reference sequences for *T. albacares* (yellowfin tuna) and *K. pelamis* (skipjack tuna) are from Genbank. Sequences of PCR products are shown for canned yellowfin (lot S3SGB X146C), fresh yellowfin, unidentified tuna (lot KB0ZF F161C), and fresh skipjack tuna. Symbols: |, matches base on preceding line; +, does not match base on preceding line but does match reference sequence; *, matches preceding line but not reference sequence; blank, matches neither reference sequence nor preceding line. Abbreviations for ambiguities are given in section 2.3.

same chunk agreed with one another at least as well as replicate sequencing of an identical sample. This also held true for three other cans in which multiple samples from the same chunk were sequenced (data not shown).

3.2.3.c. Sequences from Different Chunks. Figure 5 illustrates the only can, of four in which two or more chunks were analyzed, in which the species identity differed from one chunk to another. In this can, one chunk had a consensus sequence most consistent with identification as bonito (*E. affinis*), allowing for possible intraspecies variation at base 359. The other two chunks had consensus sequences that agreed 100% with the reference Genbank sequence for frigate mackerel (*Auxis thazard*). We were not able to obtain fresh fish samples of frigate mackerel to verify the published sequence, but the agreement of the canned sample with the published sequence was perfect. Some of the ambiguities in the sequencing of chunk 1 may have been due to contamination with DNA from chunks 2 and 3. In three other cans in which multiple chunks were sequenced, the sequences agreed with one another as well as replicate sequences from the same sample.

3.2.4. Summary of Sequence Data. A summary of the fish species identified by sequencing in 10 of 11 cans and in 5 fresh fish is given in Table 2. In all cases in which a known sample was sequenced (i.e. fresh fish samples and explicitly labeled cans), sequence data were able to identify the species. Although in most cases the identity of the fish was clear from the forward sequence alone, the identities of all samples in Table 2 were based on sequences obtained in both directions since, as

```
#  Identification  Sequence

   chunk 1
     sample 1
1.     seq 1, f    GGGTACCACCCTCGTTGAATGAATCTGAGGGGGCTTCTCAGTAGACwATGCCACCCTCAC
2.     seq 2, f    GGGTACCACCCTCGTTGAATGAATCTGAGGGGGCTTCTCwGTAGACwATGCCwCCCTCwC
3.   sample 2, f   GGGTACnACCCTCGTTGAATGAATCTGAGGGGGCTTCTCAGTAGACAATGCCACyCTCAC
4. chunk 2, f      GGGCACnACCCTCGTAGAATGAATCTGAGGGrrCTTCTCAGTAGACAATGCCACTCTCAC
5. chunk 3, f      GGGCACnACCCTCGTAGAATGAATCTGAGGGarCTTCTCAGTAGACAATGCCACTCTCAC

   chunk 1
6.   sample 1, ic GGGyACCACCCTTGTwGAATGAATCTGAGGGGnCTT-----------------------

7.   sample 2, ic GGGyACCACCCTTGTwGAATGAATCTGAGGGGGCTTCTCAGTAGACAATGCCACCCTCA-
   chunk 2
8.   seq 1, ic    GGGCACCACCCTCGTAGAATGAATCTGAGGGGGCTTCTCAGTAGACAATaCCACTCTsA-
9.   seq 2, ic    GGGCACCACCCTCGTAGAATGAATCTGAGGGGGCTTCTCAGTAGACAATGCCACTCTCAC
10.chunk 3, ic    GGGCACCACCCTCGTAGAATGAATCTGAGGGGGCTTCTCAGTAGACAATGCCACTCTCAC

cons chunk 1      GGGTACCACCCTyGTTGAATGAATcTGAGGGGGCTTCTCAGTAGACAATGCCACCCTCAC
cons fresh bonito GGGTACCACCCTTGTTGAATGAATTTGAGGGGGCTTCTCAGTAGACAATGCCACCCTCAC
E. affinis (GB)   GGGTACCACCCTTGTTGAATGAATTTGAGGGGGCTTCTCAGTAGACAATGCCACCCTCAC
                      *           *        !                              *
cons ch 2 & 3     GGGCACCACCCTCGTAGAATGAATCTGAGGGGGCTTCTCAGTAGACAATGCCACTCTCAC
A. thazard (GB)   GGGCACCACCCTCGTAGAATGAATCTGAGGGGGCTTCTCAGTAGACAATGCCACTCTCAC
```

Figure 5. Sequence variability and reproducibility of samples from one can. Samples are from three different chunks of fish in a can from lot BTCO 3807. All samples were subjected to sequencing in both forward (f, using primer TUNA334F) and reverse [using either primer TUNA395R or TUNA396R; inverse complement (ic) shown] directions. Separate samples from chunk 1 are labeled samples 1 and 2. Replicate sequences of the same sample in the same direction are labeled seq 1 and 2. Abbreviations for ambiguities are given in section 2.3. In experimentally determined sequences (sequences 1−10), the nucleotide is shown in bold lower case if it was ambiguous or in disagreement with the consensus sequence for that chunk. Hyphens (-): sequence believed to be inaccurate, based on the chromatogram and/or >50% disagreement with possible reference sequences. In the consensus (cons) sequences, lower case bold is used if the base was ambiguous or in disagreement with the most likely Genbank (GB) reference sequence. *, bases differing between *E. affinis* and *A. thazard* which support the indicated species identification. !, a base in chunk 1 which disagrees with the Genbank reference sequence, possibly due to *E. affinis* intraspecies variation.

described above, ambiguous bases were sometimes encountered in sequencing one direction or another. Sequence data from mitochondrial DNA region R335 clearly enabled us to distinguish species classified as "tuna" from those classified as "bonito" according to the classification of the European community (Table 1).

Sequences from region R211 were obtained during the pilot stages of this project, prior to developing methods of purifying PCR products adequately to get high-quality sequence data. These sequences contained larger numbers of ambiguous or incorrect bases (data not shown). Nevertheless, by focusing mainly on specific species-distinctive bases, the correct species was identified for 9 of 10 samples sequenced. The fact that we were able to identify species correctly in all but one case even with relatively poor pilot sequence data demonstrated to us that identifying species using sequences obtained from PCR products is an extremely robust method that works well even when sequencing errors and ambiguities are present. We hasten to add, however, that identification is most convincing when PCR, DNA purification, and sequencing conditions have been optimized, as described under Materials and Methods, to obtain the most accurate sequence data.

3.3. Restriction Site Analysis. Overall results with restriction enzymes are summarized in Table 2. The uncut PCR products were approximately 120 bp in length. Some restriction enzymes cut PCR products from all species (e.g. *Rsa*I in region R211), but yielded different fragment patterns for different species. For other enzymes, one or more species were resistant to the enzyme, retaining an intact 120 bp PCR product, while other species were sensitive and yielded smaller fragments (e.g. *Hin*fI in region R335).

Typical examples are illustrated in Figure 3. In Figure 3A, *Rsa*I cut R211 of yellowfin tuna into two pieces, approximately 80 and 40 bp long, whereas, the *Rsa*I product from bonito was one rather broad band about 40 bp in size. As illustrated in Figure 1, yellowfin tuna has an *Rsa*I restriction site approximately 40 bp from one end (including primer regions), but bonito has two *Rsa*I restriction sites approximately 40 bases from each end. Thus, *Rsa*I produced the predicted result and clearly distinguished bonito from yellowfin tuna and, as summarized in Table 2, also from skipjack tuna and frigate mackerel.

In Figure 3B, *Hin*fI analysis of R335 PCR products gave a 70 bp fragment from yellowfin tuna, skipjack tuna, and frigate mackerel, but bonito was completely resistant to this enzyme. Although this was not quite the predicted fragmentation pattern, it did clearly distinguish bonito from the other species. The *Hin*fI sensitive site in the 3′-primer region of all of these species was apparently not present in the PCR products, possibly due to the degeneracy of the primer set. Furthermore, the 50 bp fragment that should also result from *Hin*fI in yellowfin and other species was not evident on any of the gels. It either ran anomalously with the 70 bp fragment or has been cut into even smaller fragments that were not seen. Regardless of the explanation of the missing band, it is clear that *Hin*fI distinguished bonito from several other species.

In region R211, the bonito PCR product was resistant to *Mbo*II, while yellowfin was cut into the predicted 80 and 40 bp products (data not shown). We attempted to cut the *Rsa*I product of this region in yellowfin with *Mbo*II, a procedure that would have been expected to distinguish yellowfin from skipjack tuna (see restriction

maps in Figure 1). However, for unexplained reasons, no change in the *Rsa*I products occurred. Modifications of the reaction conditions may enable this procedure to work.

*Rsa*I and *Kpn*I could also distinguish R335 PCR products of bonito from several other species (Table 2). However, in some cases (*Kpn*I for can 9 bonito; *Rsa*I for can 11 bonito), the reactions seemed not to go to completion, leaving some of the original 120 bp PCR product. The reason for these incomplete reactions is unknown, but it is not thought to be due to mixed species in the PCR product, since sequence data and other restriction enzymes did not yield such a mixed result.

None of the above-described restriction enzymes distinguishes frigate mackerel from yellowfin tuna and skipjack tuna. The presence of frigate mackerel in our samples (in can 10) was discovered too late in this study to test a restriction enzyme procedure that would distinguish frigate mackerel from yellowfin tuna and skipjack tuna. This would obviously be necessary to use restriction site analysis to determine if this species, classified by the European market with bonito, is present in canned fish samples. Restriction enzymes to accomplish this are, however, available: in region R335, the restriction enzymes *Bmy*I and *Sdu*I should cut only frigate mackerel; *Ban*I should cut bonito and frigate mackerel, but not any of the tunas. Future applications of these procedures should determine whether these enzymes yield the predicted fragment patterns for distinguishing frigate mackerel from tuna species.

3.4. Conclusions. PCR techniques followed by sequencing and/or restriction site analysis can be used reliably to differentiate species of previously canned fish. Despite the considerable DNA degradation that occurs in the canning process, PCR was able to amplify DNA from all oil-packed samples and most water-packed fish products tested. Two methods of analyzing the PCR products were demonstrated. The first method, sequencing the PCR products, was shown to be a robust and reliable method that gives precise information on the species present. The second method, analyzing for the presence of specific restriction sites, was shown to differentiate bonito from several other species. The restriction enzymes tested thus far were not appropriate to distinguish frigate mackerel, classified by the European market with bonito, from tuna species. However, we have suggested several restriction enzymes that should be able to identify frigate mackerel specifically.

The labeling on several store-bought cans was corroborated, and in a limited sample, skipjack tuna was identified in cans of Chunk Light tuna. However, analysis of canned fish samples from Unicord showed that some lots of canned fish may contain a mix of species, including more than one species in a single can, though, it should be noted, their contents were consistent with the European division between "tuna" and "bonito". The fact that these lots contained a mixed collection of fish clearly points up the usefulness of a monitoring method to authenticate the content of canned products.

The ability to get PCR products, and hence to do the analysis, may be affected by canning conditions. A visible PCR product was not obtained with one of the fish samples canned in water. Although the size range of DNA in this sample is as large as the DNA in other canned samples, it may be that the intense heat and pressure involved in the canning process would have degraded cytochrome *b* DNA beyond the detection abilities even of PCR. An alternative explanation of the lack of a PCR product is that PCR inhibitors may have been present. Additives or effects of the canning process might result in the production of PCR inhibitors. PCR inhibitors are common in degraded DNA (Akane et al., 1994; Goodyear et al., 1993). Possibly, as with degraded DNA from other sources, methods can be developed to get useful PCR products from these as well. However, the fact that we were not able to amplify DNA from all canned fish samples highlights the possibility that canning methods or additives that preclude the application of PCR could be intentionally used to prevent the application of these methods for species authentication.

Of the two methods demonstrated to be useful in identifying the species present, restriction site analysis is the more efficient and less expensive and thus more likely to be of use as an industrially viable detection method. For restriction site analysis, a large number of reactions can be incubated and then run on gels simultaneously, enabling a high throughput. The cost of the reactions varies according to the enzyme used, but an example is $0.50 per reaction using 12 units of *Hin*fI. The cost of buffers, gels, and labor would add only a few dollars more per sample, and analysis requires less expensive equipment than automated sequencing. Furthermore, the type of species present may be obvious by inspection. For example, if *A. thazard* were not a possibility for the data in Figure 3B, it is obvious by inspection which samples should be classified as "bonito".

In contrast, sequencing of individual PCR products presently has a commercial cost of about $35 per sequence, and for completely reliable conclusions concerning species, the PCR product may have to be sequenced in both directions, doubling the cost. An additional expense is purifying the PCR product prior to sequencing. Conclusions concerning species require detailed comparison of sequences and sometimes inspection of original chromatograms. For industrial application, sequencing could, however, be used for quality control or corroboration of restriction site analysis in cases of questionable labeling (e.g., if some samples in a lot labeled "tuna" gave restriction products with "bonito" fragmentation patterns).

Finally, it should be noted that we observed some intraspecies sequence variation not previously described in the scientific literature or Genbank. In the above results, we found substitution of C for T at base 371 in a fresh yellowfin sample and substitution of G for A at base 365 in a fresh skipjack tuna sample. Neither of these substitutions is at restriction sites analyzed in this study; however, it would be important to know whether any intraspecies variation is present at restriction sites as this would also affect conclusions regarding species identity. A more complete survey of intraspecific variation is needed for application of either sequencing or restriction site analysis for authentication of labeling. Nevertheless, this study demonstrates techniques that will work with canned fish samples and can be used with confidence once the degree of intraspecific variation is known.

ACKNOWLEDGMENT

We thank Wünsch International Ltd. for providing financial support for this research and for arranging for us to receive canned fish and fresh fish samples from their supplier, Unicord Public Co., Ltd.

LITERATURE CITED

Akane, A.; Matsubara, K.; Nakamura, H.; Takahashi, S.; Kimura, K. Purification of highly degraded DNA by gel filtration for PCR. *Biotechniques* **1994**, *16*, 235−238.

Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403−410.

Block, B. A.; Finnerty, J. R.; Stewart, A. F. R.; Kidd, J. Evolution of endothermy in fish: mapping physiological traits on a molecualr phylogeny. *Science* **1993**, *260*, 210−214.

Brutlag, D. L.; Dautricourt, J. P.; Maulik, S.; Relph, J. Improved sensitivity of biological sequence database searches. *Comp. Appl. Biosci.* **1990**, *6*, 237−245.

Dougherty, J. D.; Moore, W. S.; Ram, J. L. Mitochondrial DNA analysis of round goby (*Neogobius melanostomus*) and tubenose goby (*Proterorhinus marmoratus*) in the Great Lakes basin. *Can. J. Fish. Aquat. Sci.* **1996**, in press.

Goodyear, P. D.; MacLaughlin-Black, S.; Mason, I. J. A reliable method for the removal of co-purifying PCR inhibitors from ancient DNA. *Biotechniques* **1993**, *16*, 232−235.

Received for review December 19, 1995. Accepted May 17, 1996.[⊗]

JF950822T

[⊗] Abstract published in *Advance ACS Abstracts*, July 1, 1996.