JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Fish Species Identification Using PCR−RFLP Analysis and Lab-on-a-Chip Capillary Electrophoresis: Application to Detect White Fish Species in Food Products and an Interlaboratory Study

JOHN J. DOOLEY, HELEN D. SAGE, MARIE-ANNE L. CLARKE, HELEN M. BROWN, AND STEPHEN D. GARRETT*

Biochemistry Section, Department of Chemistry and Biochemistry, Campden & Chorleywood Food Research Association, Chipping Campden, Gloucestershire, GL55 6LD, UK

Identification of 10 white fish species associated with U.K. food products was achieved using PCR−RFLP of the mitochondrial cytochrome *b* gene. Use of lab-on-a-chip capillary electrophoresis for endpoint analysis enabled accurate sizing of DNA fragments and identification of fish species at a level of 5% (w/w) in a fish admixture. One restriction enzyme, *Dde*I, allowed discrimination of eight species. When combined with *Nla*III and *Hae*III, specific profiles for all 10 species were generated. The method was applied to a range of products and subjected to an interlaboratory study carried out by five U.K. food control laboratories. One hundred percent correct identification of single species samples and six of nine admixture samples was achieved by all laboratories. The results indicated that fish species identification could be carried out using a database of PCR−RFLP profiles without the need for reference materials.

KEYWORDS: Capillary electrophoresis; food authenticity; species identification; PCR−RFLP; DNA fingerprinting; lab-on-a-chip; cytochrome *b*; fish species

INTRODUCTION

The diversity of fresh, frozen, and fish-based products available to the consumer has increased significantly in recent years. Products can range from premium-grade fish steaks to low-cost fish fingers. As fish are caught, processed, and distributed by a global network of operators, there is a need to ensure the authenticity and the origin of fish used in the products. This is especially true in the European Union (EU) where stringent fish catch quotas have been introduced in an attempt to limit the decline of native fish stocks. There is, therefore, a need to have reliable and simple species identification methods to support enforcement and compliance with labeling legislation (EC Council Regulation No. 104/2000 and EC Commission Regulation No. 2065/2001).

Methods of fish species identification based on morphological characteristics are suited to whole or lightly processed fish; however, the identification of fish species becomes more difficult once it has been processed. The use of unique species-specific protein profiles has been reported for fish identification (*1*, *2*); however, these methods are less reliable when applied to processed food products as the proteins become denatured. Furthermore, they require the analysis of species reference materials along with the samples. In terms of simplicity and speed, antibody-based methods would be most appropriate. However, only a limited number of immunoassays have been developed, and none are available for wide-scale commercial use (*3*, *4*). DNA-based methods offer an alternative approach to species identification as DNA remains detectable in all but the most heavily processed samples.

Direct sequencing is the most definitive method of identification; however, it cannot easily be applied to samples suspected or known to contain more than one species. Alternative techniques, using the polymerase chain reaction (PCR), have been applied to variable regions of DNA such as the cytochrome *b* or 5S rDNA genes. Although specific PCR assays have been used successfully to differentiate sole (*Solea solea*) and Greenland halibut (*Reinhardtius hippoglossoides*) species (*5*), in general more all-inclusive methods are used to identify a wider range of species. These methods include single strand conformation polymorphism (SSCP) (*6*), PCR−RFLP (restriction fragment length polymorphism) (*7*−*12*), and random amplified polymorphic DNA (RAPD) fingerprinting (*13*−*15*).

A PCR−RFLP technique, which involves digesting an amplified 464 bp region of the cytochrome *b* gene with restriction enzymes to generate DNA profiles, has been used for the identification of a variety of species, including salmon, hake, sardine, eel, and flatfish (*9*, *11*, *16*; G. Hold, personal communication).

Although useful for identification purposes, the PCR−RFLP technique relies on the use of gel electrophoresis and staining for endpoint detection, methods that are potentially hazardous

* Corresponding author. Tel.: +44 (0)1386 842175; fax: +44 (0)1386 842100; e-mail: s.garrett@campden.co.uk.

10.1021/jf047917s CCC: $30.25
Published on Web 03/31/2005

and time-consuming and can sometimes produce variable results. An improvement to this technique, in which the gel electrophoresis step was replaced by a microfluidic, lab-on-a-chip-based capillary electrophoretic system, has recently been reported (*16*). The improved method uses the Agilent 2100 Bioanalyzer to separate and detect DNA fragments by the application of capillary electrophoresis (CE) technology. The generation of species-specific PCR−RFLP profiles on the 2100 Bioanalyzer enabled accurate sizing and quantification of individual DNA fragments, which gave the system a significant advantage over gel-based approaches in terms of ease of use, speed, and safety.

This paper reports the development of PCR−RFLP profiles on the 2100 Bioanalyzer that can be used to identify a range of white fish species without the need for concurrent analysis of reference materials. The objective of this study was to extend the method developed by Dooley et al. (*16*) on salmon and trout to include white fish species widely used in the production of fish products consumed in the United Kingdom and to trial the method in an interlaboratory study.

MATERIALS AND METHODS

Authentic Fish Samples. Authentic samples (frozen, preserved in ethanol, or dried dorsal fin clips) from five individuals of 10 white fish species were obtained from various sources. Sample sizes of frozen or preserved fish were between 3 and 5 g. All samples were stored at −20 °C upon receipt. Details of samples and source are shown in **Table 1a**. Samples (200 mg) for DNA extraction were removed from each frozen fish sample using a fresh, sterile scalpel. The work area was cleaned with 80% ethanol between each sample from a single species and with hypochlorite solution (available chlorine = 5000 ppm) and 80% ethanol between each species. Fresh gloves were worn for each sample.

Additional Fish Samples. Fish fillets were obtained from local fish mongers and authenticated using PCR−RFLP profiles. Samples were stored, for a maximum of 2 days, at 4 °C until processed.

Fish products, as detailed in **Table 1b**, were obtained from local retailers and stored at −20 °C until required. Products were thawed overnight at 4 °C prior to processing. Where products (e.g., fish fingers, fishcakes, etc.) had coatings (e.g., breadcrumbs), these were removed as far as possible before the fish part was blended to produce a homogeneous sample. DNA was extracted from portions (2 g) of this homogeneous sample. Products comprising fish in a sauce were blended whole to ensure homogeneity prior to DNA extraction.

Freeze-Dried Fish Admixtures. Samples of fish fillet from single fish species were cut into cubes (∼2 cm^3) and freeze-dried to produce dry fish powders. Portions of the powdered fish species were finely ground and mixed together using a pestle and mortar, cooled in liquid nitrogen, to produce admixtures. Freeze-dried admixtures containing 2, 5, or 10% (w/w) of one species in another were produced. Details of admixtures are shown in **Table 1c**. All freeze-dried samples were stored dried at −20 °C until required. Portions (200 mg) of the freeze-dried admixtures were used for DNA extraction.

DNA Extraction. DNA was extracted using the CTAB−Proteinase K method described by Dooley et al. (*16*) with scaling for sample size. Briefly samples were lysed, in a CTAB buffer, with proteinase K. Proteins were removed using precipitation and chloroform extraction steps. DNA was finally recovered by precipitation using 2-propanol and purified using Promega's Wizard Purification Resin as per the manufacturer's protocol. Final DNA concentrations (ng/μL) were determined using a GeneQuant pro DNA calculator (Pharmacia).

DNA was extracted from freeze-dried material using the Tepnel Biokits DNA Extraction kit (Tepnel BioSystems Ltd., One Newtech Square, Deeside Industrial Park, Deeside, Flintshire, CH5 2NT, UK) as per the manufacturer's instructions. The protocol for DNA extraction from Raw Meat Samples was used with the following modifications; the volume of tissue extraction solution was doubled to 800 μL and the volume of RNA solution was doubled to 40 μL to compensate for the use of dried material. DNA was quantified using the GeneQuant pro DNA calculator.

Table 1. Details of (a) Authentic Fish Species, (b) Products Containing White Fish, (c) Freeze-Dried Materials, and (d) DNA Admixtures Used in This Study

(a) Details of Authentic Fish Species Used in This Study

common name	Latin name	source
Atlantic cod	*Gadus morhua*	IFR, Norwich
Pacific cod	*Gadus macrocephalus*	IFR, Norwich
coley (saithe)	*Pollachius virens*	IFR, Norwich
haddock	*Melanogrammus aeglefinus*	IFR, Norwich
European hake	*Merluccius merluccius*	IFR, Norwich
South African hake	*Merluccius paradoxus*	IFR, Norwich
European Plaice	*Pleuronectes platessa*	IFR, Norwich
whiting	*Merlangus merlangus*	IFR, Norwich
Alaskan (walleye) pollock	*Theragra chalcogramma*	AFSC, Alaska
hoki	*Macruronus novaezelandiae*	NIWA, New Zealand

(b) Details of Products Containing White Fish Used in This Study

sample name	declared species
haddock portion in oven crisp crumb	haddock (54%)
100% hoki fish fingers	hoki fillet (64%)
cod fish cakes	minced cod (45%)
salmon cakes in crunch crumb	salmon (37%), white fish (37% pollock, haddock, cod)
fish fingers	minced white fish (62%)

(c) Details of Freeze-Dried Materials Used in This Study

each of the 10 white fish species as shown in **Table 1a**

2–10% Atlantic cod[a]	90–98% haddock[a]
2–10% haddock[a]	90–98% Atlantic cod[a]
10% whiting	90% coley
10% Pacific cod	90% coley
5% hoki	95% Atlantic cod

(d) Details of DNA Admixtures Used in this Study

species 1	species 2
5% coley	95% Atlantic cod
5% hoki	95% Atlantic cod
5% whiting	95% haddock
5% SA hake	95% haddock
5% coley	95% Pacific cod
5% hoki	95% haddock
5% SA hake	95% Atlantic cod
5% whiting	95% Pacific cod
5% hoki	95% coley
5% whiting	95% Atlantic cod
5% A. pollock	95% Atlantic cod
5% hoki	95% Pacific cod
5% coley[b]	95% Pacific cod
5% whiting[b]	95% haddock
5% SA hake[b]	95% Atlantic cod
10% coley[b]	90% Pacific cod

[a] Admixtures prepared on a w/w basis. [b] Admixture contains 59% (v/v) total fish DNA and 41% (v/v) soya DNA.

DNA Admixtures. DNA extracts (10 ng/μL) from several authentic samples were pooled and used to prepare DNA admixtures. Admixtures of 5% (v/v) one fish species in a second fish species were prepared from pooled DNA stocks of each species. Appropriate volumes of DNA were mixed thoroughly in 1.5 mL Eppendorf tubes before being aliquoted into smaller volumes and stored at −20 °C until required. A second set of DNA admixtures containing two fish species and soya DNA were also prepared as described above. These admixtures contained a total of 59% (v/v) fish DNA and 41% (v/v) soya DNA. The proportions of the two fish species relative to each other were 5 or 10% (v/v) one fish in the second. Details of all DNA admixtures used are shown in **Table 1d**.

DNA Amplification. PCR products (464 bp of the cytochrome *b* gene) were produced by amplification of DNA extracts (50 ng) in 25

or 20 μL reactions containing 1× Amplitaq Gold PCR buffer (Applied Biosystems), 300 nM each of primers L14735 (5′-AAA AAC CAC CGT TGT TAT TCA ACT A-3′) and H15149 (5′-GCI CCT CAR AAT GAY ATT TGT CCT CA-3′) (*8*), 200 nM dNTPs, 5 mM $MgCl_2$, and 0.05 U/μL of Amplitaq Gold (Applied Biosystems). Amplification profiles (94 °C for 5 min [denaturation]; 40 cycles of 94 °C for 40 s, 50° C for 80 s, 72 °C for 80 s [amplification]; 72 °C for 7 min [final extension]) were applied using an Applied Biosystem's GeneAmp PCR machine. Unpurified PCR products (1 μL) were applied to the bioanalyzer to confirm amplification.

Restriction Digestion. Restriction enzymes were obtained from New England Biolabs and used as per the manufacturer's instructions. Unpurified PCR product (2.5 μL) was digested overnight with 2−5 units of enzyme in a total volume of 5 μL. Reactions were terminated by incubation at 65 °C for 10 min.

PCR−RFLP Profiling. Reagents and DNA500 LabChips were prepared following manufacturer's instructions. Batches (∼500 μL) of gel matrix (used to fill LabChip capillaries) were prepared as required, or at four weekly intervals. All reagents were stored at 4 °C when not in use and allowed to reach room temperature for 1 h before use. Digested PCR products (5 μL) were mixed with 5 μL of 20 mM EDTA, to achieve a final concentration of 10 mM EDTA, prior to loading on to LabChips. Aliquots (1 μL) of the reaction mix were loaded on to the LabChip, as per the manufacturer's instructions, and analyzed on the 2100 Bioanalyzer.

DNA Sequencing. Sequencing PCRs were performed in 25 μL reaction volumes. All DNA sequencing was performed by Lark Technologies (Saffron Walden, UK) using the BigDye Terminator protocol (Applied Biosystems). PCR products were cleaned using a Qiagen PCR purification kit before being sequenced in both directions. Sequence data was provided to CCFRA as a text file and graphical sequence output file.

Sequence Analysis. Full DNA sequences were generated from contigs using the SeqMan module of the LaserGene software suite (Ver 5.05; DNASTAR Inc, Madison, USA). Discrepancies were corrected manually using information in the graphical sequence file. All sequence alignment was performed using the MegAlign module of the LaserGene package.

Predicted RFLP patterns were produced using the AnnHyb software (version 4-17; http://annhyb.free.fr) and all commercially available enzymes in REBASE (file gcgenz.304; http://rebase.neb.com/).

Interlaboratory Study. The PCR−RFLP method was assessed in an interlaboratory study involving five independent laboratories. The participants, including those involved with U.K. government food control and enforcement, had no previous experience of using the PCR−RFLP approach for fish identification, and only one had previous experience of using the 2100 Bioanalyzer. Four of the participants shared two 2100 Bioanalyzers, and the fifth used their own 2100 Bioanalyzer. Participants attended a 1 day training event at CCFRA and were then provided with freeze-dried samples (19 unknowns and an Atlantic cod positive), a protocol, all reagents needed to perform the analysis, a DNA extraction kit, a sheet for recording results, and the expected PCR−RFLP profile patterns and fragment sizes for each species with each enzyme. Participants were asked to report the PCR−RFLP fragment sizes obtained from each sample with each enzyme along with the species that they identified as being present in each sample. Participants were also provided with a CD-R so that they could return raw data files to CCFRA.

For the purposes of the trial and because most of the participants did not have access to DNA quantification instruments, DNA was extracted using the Tepnel Biokits DNA Extraction kit and then diluted 1 in 100 prior to use in PCR reactions. This dilution had been determined to be optimal based on work at CCFRA.

RESULTS AND DISCUSSION

Sequencing and Theoretical Analysis of the cytb Gene. Sequence data for the 464 bp amplicon from the cytochrome *b* gene was generated from all five individuals of all white fish species. Sequence information is available from GenBank as accession numbers AY946297−AY946346. Owing to the degeneracy of primer H15149 exact matching of this end of the PCR products was not possible. This resulted in a small variation (±2 bp) in the length of PCR products. Sequences from the five individuals of each species were aligned to produce species-specific consensus sequences. These alignments highlighted variations (point mutations) between the individuals of each species, which is as expected in large populations.

The consensus sequences from the 10 species were aligned to confirm that sufficient sequence variation between the species was present for the PCR−RFLP method to be applied. Aligned consensus sequences for the 10 species are shown in **Figure 1**. Interestingly, this alignment revealed the presence of an additional 5 bp in the hoki (*Macruronus novaezelandiae*) sequence, which was not present in any other species. From an evolutionary point of view, this suggests an early divergence of the hoki species from the other fish species used in this study. This longer (469 bp) PCR product has been consistently amplified in all hoki samples subsequently analyzed, which indicates that this insert is not an artifact.

Following analysis of consensus sequences with AnnHyb, the predicted PCR−RFLP patterns for all fish species, using restriction enzymes with a recognition sites of >4 bp, were visualized graphically using an Excel spreadsheet. Enzymes that produced the greatest pattern variability across the species were selected for further analysis. From over 120 enzymes examined, *Dde*I was predicted to show the greatest variability across the 10 species. This enzyme was predicted to produce nine distinct profiles from the 10 fish species, which would allow the differentiation of all species except Alaskan pollock (*Theragra chalcogramma*) and Pacific cod (*Gadus macrocephalus*). Two additional enzymes, *Hae*III and *Nla*III, were selected on the basis that they were predicted to differentiate between Alaskan pollock and Pacific cod.

PCR−RFLP Profiling of White Fish Species. PCR−RFLP patterns for the white fish species were produced experimentally by amplifying DNA extracts from the authentic fish and then digesting the PCR products with the three enzymes, *Dde*I, *Hae*III, and *Nla*III. Profiles were resolved using a DNA500 LabChip on the 2100 Bioanalyzer. DNA fragments comprising each profile are automatically detected and sized by the 2100 Bioanalyzer software if their fluorescence exceeds the default detection threshold. Additional fragments can be detected by adjusting the detection threshold or by manually selecting individual peaks of interest. PCR−RFLP profiles for all 10 species using the three enzymes and default detection settings on the 2100 Bioanalyzer are shown in **Figure 2**, which is a computer-generated gel image of the data. This clearly shows that, as predicted, most species can be differentiated using *Dde*I alone (**Figure 2A**). **Figure 2B,C** also shows that profiles generated by the enzymes *Hae*III and *Nla*III complement profiles generated by enzyme *Dde*I when identifying the different fish species.

Initially analysis was performed using 25 μL PCR reactions. PCR products were purified using the QIAquick PCR purification kit (Qiagen, Crawley, UK) before being cleaved with restriction enzymes in a total restriction digest reaction volume of 10 μL. However, as only 1 μL of sample is required for analysis on the 2100 Bioanalyzer, the potential of reducing the reaction volumes was investigated. Restriction digest reactions were optimized in a total volume of 5 μL, which was the smallest volume that produced consistent results, and PCR volumes were reduced to 20 μL from 25 μL. Removal of the PCR cleanup step was also found to have no effect on the final fragment numbers or sizes observed. Additionally analysis of


A A A A A C C A C C G T T G T T A T T C A A C T A C A A G A A C C T - - - - - T A A T G G C C A G C C T T C G G A A A A Majority
10 20 30 40 50
1 . . . . . T . . . . . T A G C A A C . . . . . T T . A . . A . . . . Hoki
1 . . . . . - - - - - . . . . . A. Pollock
1 . . . . . - - - - - . . . . . A. Cod
1 . . . . . - - - - - . . . . . P. Cod
1 . . . . . - - - - - . . . . . Coley
1 . . . . . C T - - - - . . . . . C . . A . . . . E. Hake
1 . . . . . C C - - - - . . . . . C . . A . . . . SA. Hake
1 . . . . . T . - - - - . . . . . Haddock
1 . . . . . C - - - - - . . . . . A . . . C . . T . . . T Plaice
1 . . . . . - - - - - . . . . . Whiting
C C C A T C C A A T C C T A A A A A T T G C T A A T G A C G C A T T A G T T G A T C T C C C C G C C C C C T C C A A T A Majority
60 70 80 90 100 110
61 . T . . . . . T C . G . . T . . G . . C . . A . . C . . . . . T C . T . . G . . C T . A . . A . . A . . T . . . . . C . Hoki
56 . . . . . A G . . . . . A. Pollock
56 . . . . . A G . . . . . A. Cod
56 . . . . . A G . . . . . P. Cod
56 . . . . . C . . . . . A . . . . . T . . . . Coley
57 . T . . . . . C C . T . . . . . G . . . . . T . . . . . A . . C . . T . . T . . . . . T . . C C E. Hake
57 . T . . C . . C C . T . . . . . G . . . . . T . . . . . A . . C . . T . . T . . . . . C . SA. Hake
56 . T . . . . . T . . . C . T . . . T . . . . . T . . . A . T . . . . . Haddock
56 . . . . C . . C C . T . . T . . . . . C . . A . . C . . . T . . . T . . . . . C . . . . . A . . . . . T . . C . Plaice
56 . . . . . T . . . . . T . . . . . Whiting
T C T C A G T A T G A T G A A A T T T T G G C T C T C T T C T A G G C C T T T G C T T A A T T A C T C A A C T T C T A A Majority
120 130 140 150 160 170
121 . T . . C . . T A T . . . . . C . . A . . . . . A . . . C . T . . C . . A . . G A . C T . . . Hoki
116 . T . . . . . C . . . . . A. Pollock
116 . . . . . A. Cod
116 . T . . . . . P. Cod
116 . . . . . C . . . . . C . . . T . . . . . Coley
117 . . . . . A C . . T . . . . . C . . C . . G . . . . . C . . . . . C . . . . . G C C G . C . . . A . C T . . . E. Hake
117 . . . . . A C G . T . . . . . C . . . . . G . . . . . T . . . . . C . . . . . G C C G . C . . . A . C T . . . SA. Hake
116 . T . . . . . T . . . . . T . . . . . C . . . . . T C . . . . . A . G . Haddock
116 . . . . T . . C . . . . . C . . . . . G . . C . . C T . . . . A . . C . . T . . . G . G . . . . . A . . G C T . Plaice
116 . . . . . T . . . . . Whiting
C A G G A C T A T T T C T A G C C A T A C A C T A T A C C T C A G A C A T C G A G A C A G C C T T C T C A T C C G T A G Majority
180 190 200 210 220 230
181 . . . . C . . C . . C . . . . . T . . . . . T . . . . . A . . . . . T . . . . . T . . T . Hoki
176 . . . . . T . . . . . A . . G . . . . . A. Pollock
176 . . . . . A. Cod
176 . . . . . T . . . . . C . . . . . P. Cod
176 . . . . . C . . T . . . . . T . . . . . T . . G . . T . . G . Coley
177 . . . . G . . . . . G . . . . . T . . . . . G . . A . . G . . . . . T . . . T . . . . . C . E. Hake
177 . . . . . T . . . . A . . . . . C . . . G . . A . . G . . . . . T . . . T . . . . . C . SA. Hake
176 . . . . T . . . . . T . . . . . G . . . . . T . . . . . T . . A . . . . . T . . . . . Haddock
176 . C . . C T . . . . C T . . . . . G . . . . . A . . T . . T . . T . C T . . T . . . . . A . C . . . . . T . Plaice
176 . . . . . T . . . . . G . . . . . C . . A . . . . . T . . . . . Whiting

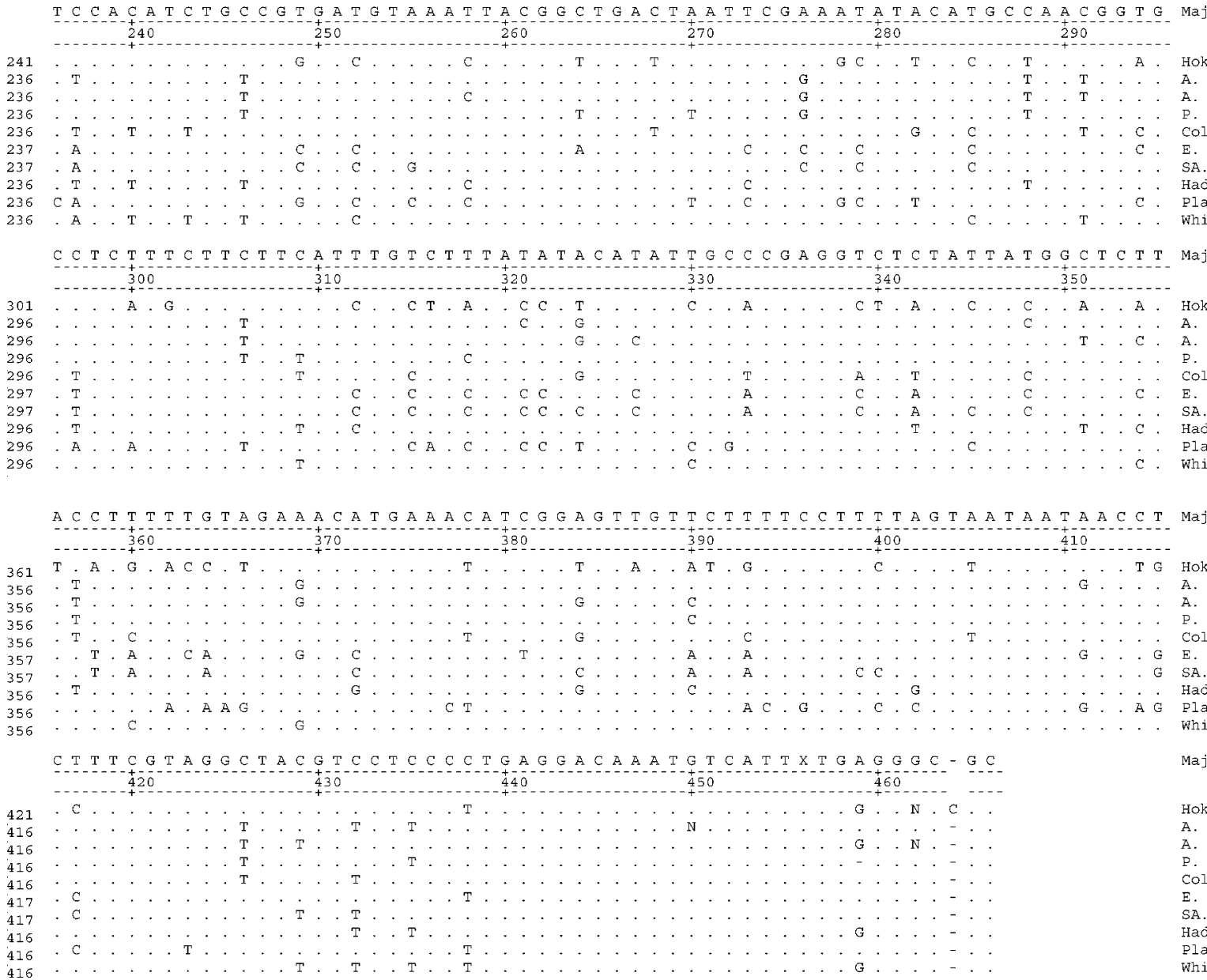

Figure 1. Alignment of cytochrome *b* gene sequence data generated from authentic fish samples used in this study. Alignment is of species-specific consensus sequence data generated from five individuals of each species. Sequence data from all samples was generated during this study and is available from GenBank (Accession numbers AY946297–AY946346). The majority base at each position of the sequence is shown at the top. Bases that vary from the majority are indicated by use of the altered base at each point of the sequence.

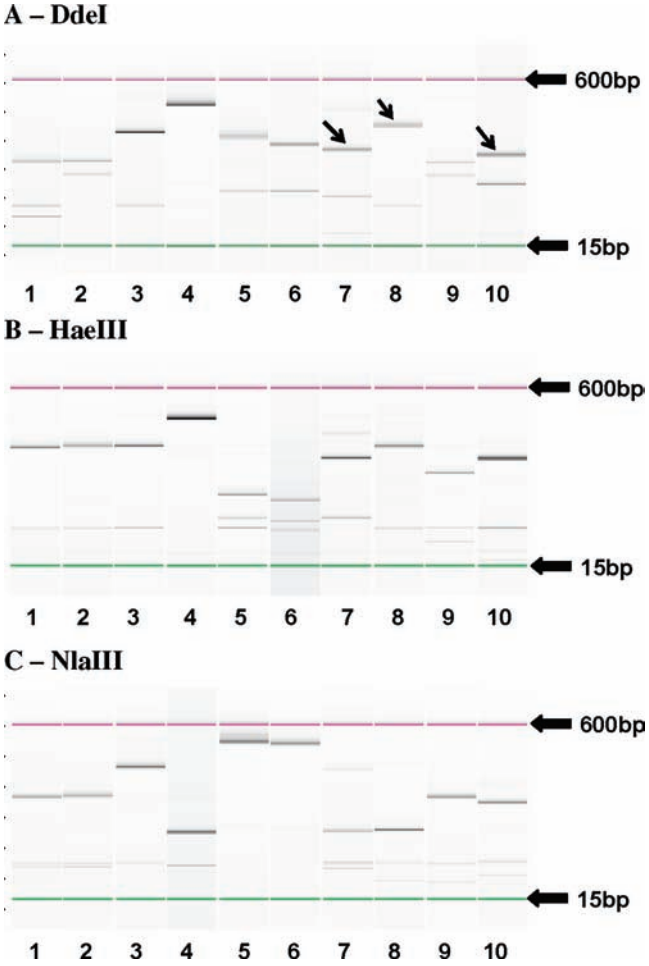

Figure 2. PCR–RFLP profiles from the 10 white fish species used in this study. Profiles were generated using enzymes *Dde*I (A), *Hae*III (B), or *Nla*III (C). Lanes 1–10 contained A. cod (1), P. cod (2), coley (3), haddock (4), E. hake (5), SA. hake (6), plaice (7), whiting (8), A. pollock (9), and hoki (10). Images were computer generated by the 2100 Bioanalyzer. Doublets generated by *Dde*I (A) are arrowed for clarity. Size markers (15 and 600 bp) run in each lane are shown.

unpurified PCR products resulted in more intense fragment peaks in the PCR−RFLP profiles, which made the detection of some fragments easier. However, a consequence of removing the PCR cleanup step was that it was necessary to add an equal volume of 20 mM EDTA to the PCR product to both reduce the high salt concentration in the *Dde*I reaction buffer and maintain 10 mM EDTA required for enzyme inactivation. Without the EDTA buffer, the stability and migration rate of DNA fragments on the LabChip were affected. These adjustments did not affect the final PCR−RFLP profiles; however, they did result in a faster analysis time compared to the conventional gel-based method described by Russell et al. (*8*). Subsequent results are based on the use of this faster protocol.

Expected and observed PCR−RFLP fragment sizes are shown in **Table 2**. In the majority of cases, the experimental PCR−RFLP patterns matched those predicted using AnnHyb software; however, some differences (which are detailed below) were seen. Profiles were defined as matching if the predicted and observed fragment sizes were within 5% of each other. A variation of 5% was chosen as the 2100 Bioanalyzer is reported to size fragments within ±5 bp for fragments of 25−100 bp and ±5% for fragments of 100−500 bp. The largest observed variation in fragment size in all profiles generated was 3.7%. The main differences between predicted and observed profiles were noticed when enzymes *Dde*I and *Nla*III were used to generate profiles as described below.

Plaice, whiting, and hoki samples, when digested with enzyme *Dde*I, produced a DNA fragment doublet, i.e., two fragments with a size difference of about 10 bp (**Figure 2A**, lanes 7, 8, and 10). Occasionally, the two peaks were not fully resolved in whiting samples; however, in these samples, a shoulder, corresponding to the smaller 341 bp fragment, was always seen on the electropherogram peak corresponding to the 347 bp fragment. This doublet was reported previously following analysis of salmon and trout (*16*) and is due to the introduction (into some PCR products) of a *Dde*I restriction site by the primer H15149.

Although some natural genetic variation within each fish population was revealed in the sequence data (data not shown), this did not manifest itself as differences in the PCR−RFLP profiles between the individuals of a single species, except for hoki. Analysis of the five hoki individuals with *Dde*I produced three distinctly different PCR−RFLP profiles as shown in **Table 2**. The most common profile (32, 176, 252, and 261 bp) was observed in three individuals. The other two individuals each produced different profiles comprising (a) 206, 254, and 261bp or (b) 35 and 447 bp, respectively. Sequence data from the five hoki individuals indicated that these profiles were due to the loss of *Dde*I restriction sites in these individuals. These restriction site losses had been masked during the original profile predictions due to the use of the hoki consensus sequence. No variation between the five individuals of hoki was observed when PCR−RFLP profiles were generated using the other two enzymes, *Hae*III and *Nla*III. The identification of hoki can be achieved using either *Hae*III or *Nla*III (**Table 2**), and as such the occurrence of multiple *Dde*I profiles does not detract from identifying this species in a sample. Although not observed in the samples analyzed in this study, it is possible that a novel PCR−RFLP profile could be generated from an individual of one of the 10 species, due to natural variation. However, profiles generated with the other two enzymes should ensure that such an individual is correctly identified.

Haddock and whiting PCR−RFLP profiles generated with enzyme *Nla*III produced some unexpected results. These species were predicted to produce near-identical profiles, containing three DNA fragments of about 91, 180. and 193 bp (haddock) or 89, 179, and 196 bp (whiting) with this enzyme. Experimental results indicated that haddock contained two fragments of 94 and 183 bp and whiting contained two fragments of 100 and 186 bp. The smaller observed fragments correspond to the predicted 91 bp fragment; however, the larger fragment observed in both species is part way between the sizes of the two predicted larger fragments. Sequence data from all 10 individuals of these two species was checked for predicted digestion patterns. This confirmed that the predicted three fragment PCR−RFLP profiles were correct. It is believed that the two larger predicted *Nla*III fragments from each species comigrate as a single band, as observed experimentally. Work to verify this comigration is underway; however, in the absence of comigration it is difficult to explain where 180 bp of DNA could be lost.

During analysis of samples with all three enzymes, some smaller fragments of between 25 and 42 bp were inconsistently detected. This was probably because they were too close to the sizing limits for the DNA500 LabChip (25−500 bp) or did not fluoresce sufficiently to be detected. Poor fluorescence is likely to be due to the incorporation of only small amounts of dye by these small, low concentration fragments. It should be noted

Table 2. Predicted and Observed PCR–RFLP Fragment Sizes Generated with the Three Enzymes *Dde*I, *Hae*III, and *Nla*III from the 10 Authentic White Fish Species Used in This Study

	*Dde*I fragments		*Hae*III fragment		*Nla*III fragments	
white fish species	predicted	observed[a]	predicted	observed[a]	predicted	observed[a]
A. cod	(25),[b] 87, 118, 234	84, 115, 234	41, 109, 316	37, 102, 321	88, 92, 284	89, 100, 280
P. cod	(24), 205, 234	198, 235	41, 109, 316	37, 102, 320	88, 92, 284	89, 100, 279
coley	(25), 118, 321	117, 328	41, 109, 316	37, 103, 323	89, 375	101, 372
haddock	(25), 439	433	41, 423	37, 429	91, 180, 193	94, 183
E. hake	158, 306	155, 307, 314	42, 109, 124, 189	38, 101, 127, 185	uncut	473 [uncut]
SA hake	25, 158, 282	155, 285	42, 109, 124, 189	38, 101, 128, 185	uncut	477 [uncut]
E. plaice	42, 145, 270, 277	[32],[c] 138, 266, 273	41, 131, 292	37, 129, 286	88, 89, 91, 196	86, 100, 187
whiting	117, 339, 347	115, 344, 355	41, 71, 109, 242	37, 101, 326	89, 179, 196	100, 186
A. Pollock	(24), 205, 234	198, 232	41, 109, 316	38, 67, 101, 243	38, 50, 88, 287	38, 47, 98, 278
hoki[d]	32, 176, 252, 261	[35], 175, 255, 262	30, 46, 109, 285	[23], 44, 101, 288	45, 58, 90, 263	44, 67, 105, 264

[a] Observed sizes are the mean sizes observed from analysis performed on at least three different occasions. [b] Numbers in parentheses, ( ), are predicted fragments which are not always detected by the 2100 Bioanalyzer. [c] Numbers in square brackets, [ ], are small fragments detected intermittently by the 2100 Bioanalyzer. [d] Point mutations in the authentic hoki samples produced three PCR–RFLP profiles. Alternative PCR–RFLP profiles contain fragment sizes of (a) 206, 254, and 261 or (b) 35 and 447 bp.

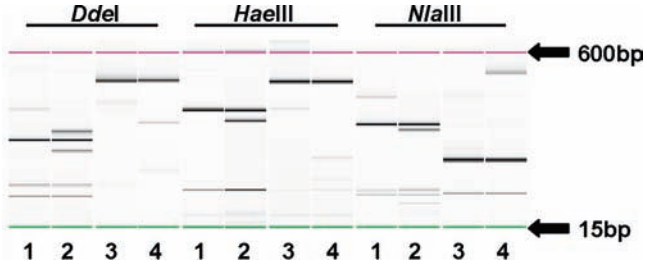

Figure 3. PCR–RFLP profiles generated from white fish DNA admixtures. Profiles were generated using enzymes *Dde*I, *Hae*III, or *Nla*III and DNA amplified from the first four DNA admixtures (1–4) shown in **Table 1d**. Images are computer generated by the 2100 Bioanalyzer. Size markers (15 and 600 bp) run in each lane are shown.

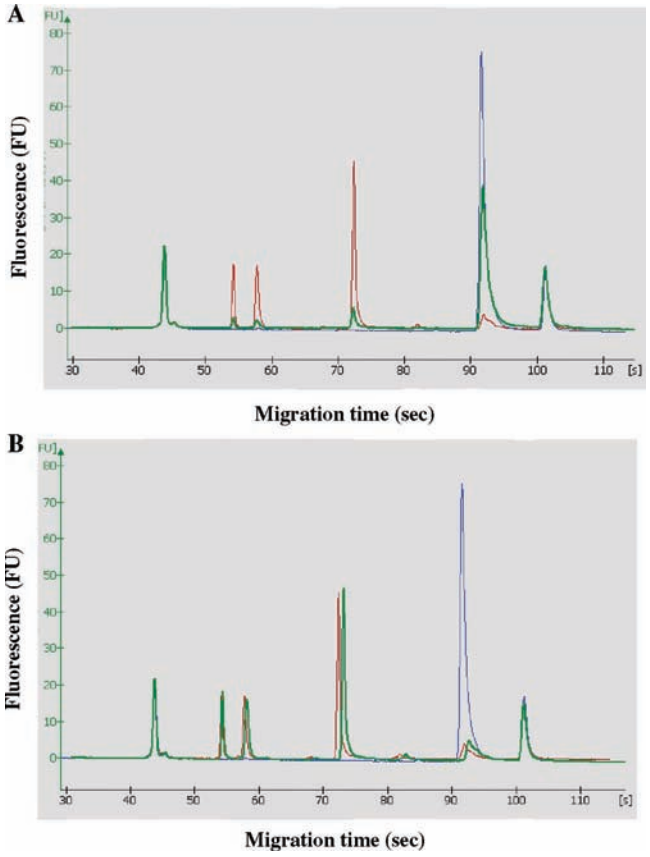

Figure 4. PCR–RFLP profiles generated from DNA admixtures of Atlantic cod and haddock. Profiles were generated from DNA admixtures of 5% Atlantic cod in haddock or 5% haddock in cod using enzyme *Dde*I. Electropherograms show the presence of species defining peaks for cod (red line) or haddock (blue line). The species mixture 5% cod in haddock (A) or 5% haddock in cod (B) are superimposed (green line).

that not detecting these fragments did not affect the ability to identify individual species.

PCR−RFLP Profiling of Admixtures. The application of this method for detecting fish species in samples containing more than one species was initially investigated using DNA admixtures. Admixtures containing two species were analyzed to demonstrate that a species present at low level (5%) could be detected in a background of another fish species. Admixtures containing more than two species were not investigated due to the complexity of superimposed profiles making interpretation difficult.

An example of the profiles obtained following analysis with all three enzymes and DNA admixtures 1−4 (see **Table 1d**) is shown in **Figure 3**. The PCR−RFLP profiles for both species, in each admixture, are detected; however, the profile for the species present at 5% is not as intense as the profile for the species present at 95%. This is as expected given the 20-fold difference in the proportions of both fish; however, this does not detract from the ability to detect the fish species present at the lower amount. Using this approach, both species in all DNA admixtures shown in **Table 1d** were detected, even in those admixtures containing soya DNA.

Additional analysis using haddock and Atlantic cod DNA admixtures at levels of 1, 2, 5, or 10% was performed to determine if a lower detection limit was possible. In some replicates, it was possible to detect as little as 1% cod in haddock or 2% haddock in cod; however, reliable detection of either species was achieved at a 5% limit. **Figure 4** shows results from analysis of admixtures containing 5% haddock in cod or 5% cod in haddock with enzyme *Dde*I. The peak height of fragments produced by the species present at the 5% level in the admixture are low compared to the peak heights for fragments produced by the species present at 95%. However, they can still be detected, indicating that this method is suitable for the detection of these species when present in an admixture at 5%.

To confirm that this approach could be used for the detection of fish species in products, the analysis was performed on admixtures prepared from freeze-dried samples of fish fillets using mixtures shown in **Table 1c**. These samples were prepared for use in an interlaboratory study of this method to facilitate

Table 3. Summary of Results Obtained Following Analysis of Freeze-Dried Fish Samples by Participants in Ring-Trial

sample ID code	species present in sample	percentage of laboratories correctly identifying main species[a]	percentage of laboratories correctly identifying minor species[a]	percentage of laboratories fully identifying sample correctly[a]
1	whiting (10%) coley (90%)	100	100[b]	100
2	haddock	100	n/a	100
3	haddock (10%) A. cod (90%)	100	100	100
4	hoki	100	n/a	100
5[c]	SA hake	100	n/a	100
6[c]	Pacific cod (10%) Coley (90%)	100	25	25
7[c]	A. cod (10%) haddock (90%)	100	100	100
8[c]	E. plaice	100	n/a	100
9[c]	whiting	100	n/a	100
10	haddock (5%) A. cod (95%)	100	100	100
11	A. cod (2%) haddock (98%)	100	100	100
12	A. cod (5%) haddock (95%)	100	100[d]	100
13	Alaskan pollock	100	n/a	100
14	E. hake	100	n/a	100
15	haddock (2%) A. cod (98%)	100	80[e]	80
16	hoki (5%) A. cod (95%)	100	100	100
17	Pacific cod	100	n/a	100
18	coley	100	n/a	100
19	Alaskan pollock	100	n/a	100
COD[f]	Atlantic cod	100	n/a	100

[a] A total of five laboratories participated in the trial, and all five provided results contributing to percentages shown here. [b] One laboratory reported whiting as a possible species due to presence of 185 bp fragments in *Nla*III digests. [c] Results for these five samples are based on data from four laboratories only. [d] One laboratory reported the presence of hoki rather than Atlantic cod; however, analysis of the data files indicated only that Atlantic cod and haddock were present in this sample. Values have been adjusted to reflect this correction. [e] Haddock was not reported by one laboratory. [f] This was a positive control sample used to confirm progress at each stage of the analysis. Although not reported by two laboratories, a check of the PCR–RFLP profiles produced with this sample indicated that Atlantic cod had been correctly identified.

distribution to participants. Results of this analysis (data not shown) were similar to those obtained with DNA admixtures and indicated that it was possible to consistently detect all species when present in the admixtures at a level of 5%. Additionally cod or haddock were detected in some replicate analyses of admixtures containing 2% of these species.

Interlaboratory Study. Results from the interlaboratory study are summarized in **Table 3**. It was clear that one laboratory had problems analyzing samples 5−9, as evidenced by the presence of extraneous PCR−RFLP fragments in the final profiles for these samples. An examination of the raw data submitted to CCFRA suggested that the problems initially occurred during PCR amplification as up to eight extra PCR products of between 129 and 709 bp were observed in some of these samples. Additionally, this laboratory appeared to have had a problem with the *Hae*III profiles, due to the loss of the upper marker. This is likely to be due to the incomplete mixing of EDTA with the samples prior to loading them onto the 2100 Bioanalyzer. This meant fragment sizes were not calculated correctly, and so it was not possible to identify the species present in these samples. Because of these problems, results for samples 5−9 for this laboratory have been omitted from the analysis of results described below.

The results (**Table 3**) show that all 11 samples comprising a single species were correctly identified by all laboratories (100%), confirming that species can be identified using this approach without the need to run authentic samples with the unknowns. This is a distinct advantage of this method compared to the conventional protein identification method routinely used in many laboratories.

Admixtures of cod and haddock at, below, and above 5% (of one species in the other) were provided to participants to verify that 5% was a realistic limit of detection (LOD) for this method. It was expected that all samples containing 5 or 10% of one species in the other would be correctly identified but that some of the species at 2% in the admixtures may not be identified, as these are below the LOD.

Analysis of cod and haddock admixtures indicated that all laboratories correctly identified the major species in all admixtures. When haddock comprised the minor species in the admixtures, it was correctly identified by all laboratories in all samples except sample 15 where one laboratory failed to identify the presence of 2% haddock. Similarly, when Atlantic cod comprised the minor component of the admixture it was correctly identified in all samples except sample 12 where one laboratory misclassified it as hoki. An examination of the data files from this laboratory revealed profiles indicating the presence of Atlantic cod but not hoki in this sample. The data files also showed that the amount of amplified DNA in this sample was low, which resulted in a reduced fragment intensity on the PCR−RFLP profiles. This meant peaks indicating the presence of Atlantic cod were below the detection threshold. A similar low intensity profile was observed in sample 15 where the 2% haddock had not been identified. It was concluded that in these cases if analysts had more experience of analyzing

Table 4. Results Obtained Following Analysis of Commercial Products Using the PCR–RFLP Method

sample[a]	enzyme	predicted fragment sizes (bp)	observed fragment sizes (bp)	potential species identified
haddock portions in oven crisp crumb (haddock)	*Dde*I	433	441	haddock
	*Hae*III	37, 429	36, 419	haddock
	*Nla*III	94, 183	92, 179	haddock
100% hoki fish fingers (hoki)	*Dde*I	35, 175, (255), 262	26, 170, 250, 258	hoki
	*Hae*III	23, 44, 101, 288	22, 42, 101, 286	hoki
	*Nla*III	44, 67, 105, 264	42, 67, 103, 264	hoki
cod fish cakes (cod)	*Dde*I	84, 115, 234	86, 115, 238	Atlantic cod
	*Hae*III	37, 102, 321	36, 101, 318	Atlantic cod
	*Nla*III	89, 100, 280	88, 100, 277	Atlantic cod
salmon cakes in crunch crumb (salmon & white fish [pollock, haddock, cod])	*Dde*I	110, 312, 321 & 198, 232	110, 195, 229, 310, 319	Atlantic salmon, Alaskan pollock
	*Hae*III	35, 98, 318 & 38, 67, 101, 243	35,[b] 67, 99,[b] 241, 319	Atlantic salmon, Alaskan pollock
	*Nla*III	438 & 278, 98, 47, 38	276, 435	Atlantic salmon, Alaskan pollock
fish fingers (white fish)	*Dde*I	n/a	194,[b] 229[b]	Pacific cod, Alaskan pollock
	*Hae*III	n/a	37,[b] 66, 100,[b] 242, 324	Pacific cod, Alaskan pollock, Atlantic cod, coley, whiting
	*Nla*III	n/a	38, 47, 98, 275	Alaskan pollock

[a] Species in brackets were those declared on product. [b] Fragments derived from both species.

profiles they would be better able to determine the presence of species present at low levels.

For three other admixtures (samples 1, 6, and 16), sample 1 (10% whiting in 90% coley) and sample 16 (5% hoki and 95% Atlantic cod) were correctly identified by all five laboratories (100%); however, only one laboratory identified the presence of Pacific cod in sample 6 (10% Pacific cod in 90% coley). From an examination of the data files, it was possible to identify small peaks (especially the 235 and 198 bp fragments of *Dde*I) that indicate the presence of Pacific cod in these samples. Again, the height of these peaks was below the default detection threshold, so interpretation of these data was dependent on the user.

The results from this trial were encouraging especially considering the lack of experience of the participants in the use and interpretation of data from the 2100 Bioanalyzer and in the use of PCR−RFLP analysis for species identification. This confirms that this method is easy-to-perform and robust and should be readily transferable with a minimum of training to analytical laboratories wishing to perform this type of analysis. This method can be used as a general screen to identify a wide range of fish species, which provides advantages over direct PCR assays when applied to products with unknown species content. However, if a questionable result was obtained with this method the analyst would still be able to perform a secondary check using a more specific technique.

Analysis of Commercial Products. The method was also applied to fish-based products purchased locally. A prediction of possible species in these types of products was made by matching fragment sizes obtained to those predicted for the 10 species. A summary of the results is shown in **Table 4**, which contains expected and observed PCR−RFLP profiles for five different fish products detailed in **Table 1b**. The expected fragments shown are the average fragment sizes previously obtained from authentic materials rather than the predicted fragments sizes obtained from sequence data. For those products where a particular species was declared, it was relatively easy to check for the presence of the predicted fragment sizes corresponding to the declared species and hence confirm its presence in the sample. Additional species were also readily identified by the presence of extra DNA fragments in the three enzyme profiles. Some products, however, were described as containing "white fish", and in this case it was not possible to check the predicted fragment sizes to confirm the presence of a particular species.

The results in **Table 4** show that none of the five products tested contained undeclared species. The first three products, which contained a single species of either haddock, hoki, or cod, were readily identified from their respective PCR−RFLP profiles generated with all three enzymes. Analysis of the salmon fish cakes clearly indicated the presence of Atlantic salmon; however, smaller peaks corresponding to Alaskan pollock were also observed, suggesting the presence of this species. Alaskan pollock was declared on the package as one of three possible white fish species forming 37% of the product. There was no evidence of the other two species (haddock or cod) in this sample. The fish finger product was declared to contain white fish only with no named species. Examination of PCR−RFLP profiles suggested that this product contained Alaskan pollock as PCR−RFLP profiles consistent with this species were obtained from all three enzymes. Although other potential species were identified by enzymes *Dde*I and *Hae*III, only Alaskan pollock was identified with enzyme *Nla*III (**Table 4**).

Conclusion. These results indicate that this method is well suited to confirm the presence of a declared species in fish-based products and that it is also possible to determine some of the species present in a product described as containing white fish. It should be noted, however, that the determination of species in a product where no specific species has been declared is limited to those species for which authenticated profiles are available. To improve the reliability of this approach (i.e., to identify undeclared fish species), it would be useful to expand the number of fish species for which PCR−RFLP profiles are available. Despite the need for predetermined profiles, for species identification, this is still a very useful method for rapidly confirming the presence of a fish species in a fish product and for recognizing the presence of additional species that should not be present. An additional use of this method is for sample comparison. We have used this approach to analyze several samples to confirm that they are identical or to confirm the presence of one sample in a second sample. In this case, there is no requirement for a species-specific profile from an authentic sample.

This method is relatively robust and accurate to a limit of around 5%, but prior treatment of fish samples, which is beyond the control of the analyst, may affect this limit. The effects of processing, especially canning, can affect the detection of fish species in samples due to DNA degradation. PCR targets of greater than 150 bp are especially vulnerable; however, we have had some success applying this method, with its large 464 bp

amplicon, to some commercial brands of canned salmon (data not shown). The method is, however, offered as a qualitative test for screening raw, frozen, and lightly processed samples for the presence or absence of fish species. As such, this method is extremely useful as it has the potential to detect a wide range of fish species following the development of suitable PCR−RFLP profiles or can be applied as a rapid means of comparing samples. This, along with the relatively cheap cost of the instrument, should make the assays developed suitable for wide use in quality control and enforcement laboratories.

ACKNOWLEDGMENT

We thank Agilent Technologies for the loan of 2100 Bioanalyzers used during the interlaboratory study and to the participants from the Laboratory of the Government Chemist, Tepnel Biosciences, Worcestershire Scientific Services, Central Science Laboratory and Greater Manchester Scientific Services. We also thank Dr. Mike Canino (AFSC, Alaska), Dr. Vicky Webb (NIWA, New Zealand) and Dr. Mark Winterbone (IFR, UK) for supplying authentic fish material for this work.

LITERATURE CITED

(1) Colombo, M. M.; Colombo, F.; Biondi, P. A.; Malandra, R.; Renon, P. Substitution of Fish Species Detected by Thin-Layer Isoelectric Focusing and a Computer-Assisted Method for the Evaluation of Gels. *J. Chromatogr. A.* **2000**, *880*, 303−309.

(2) Tepedino, V.; Berrini, A.; Borromeo, V.; Gaggioli, D.; Cantoni, C.; Manzoni, P.; Secchi, C. Identification of Commercial Fish Species Belonging to the Orders Pleuronectiformes and Gadiformes: Library of Isoelectric Focusing Patterns. *J. AOAC Int.* **2001**, *84*, 1600−1607.

(3) Cespedes, A.; Garcia, T.; Carrera, E.; Gonzalez, I.; Fernandez, A.; Asensio, L.; Hernandez, P. E.; Martin, R. Indirect Enzyme-Linked Immunosorbent Assay for the Identification of Sole (*Solea solea*), European Plaice (*Pleuronectes platessa*), Flounder (*Platichthys flesus*) and Greenland Halibut (*Reinhardtius hippoglossoides*). *J. Food Protein* **1999**, *62*, 1178−1182.

(4) Asensio, L.; Gonzalez, I.; Rodriguez, M. A.; Mayoral, B.; Lopez-Calleja, I.; Hernandez, P. E.; Garcia, T.; Martin, R. Identification of Grouper (*Epinephelus guaza*), Wreck Fish (*Polyprion americanus*), and Nile Perch (*Lates niloticus*) Fillets by Polyclonal Antibody-Based Enzyme-Linked Immunosorbent Assay. *J. Agric. Food Chem.* **2003**, *51*, 1169−1172.

(5) Cespedes, A.; Garcia, T.; Carrera, E.; Gonzalez, I.; Fernandez, A.; Hernandez, P. E.; Martin, R. Identification of Sole (*Solea solea*) and Greenland Halibut (*Reinhardtius hippoglossoides*) by PCR Amplification of the 5S rDNA Gene. *J. Agric. Food Chem.* **1999**, *47*, 1046−1050.

(6) Cespedes, A.; Garcia, T.; Carrera, E.; Gonzalez, I.; Fernandez, A.; Hernandez, P. E.; Martin, R. Application of Polymerase Chain Reaction-Single Strand Conformational Polymorphism (PCR−SSCP) to Identification of Flatfish Species. *J. AOAC Int.* **1999**, *82*, 903−907.

(7) Cespedes, A.; Garcia, T.; Carrera, E.; Gonzalez, I.; Sanz, B.; Hernandez, P. E.; Martin, R. Identification of Flatfish Species Using Polymerase Chain Reaction (PCR) Amplification and Restriction Analysis of the Cytochrome b Gene. *J. Food Sci.* **1998**, *63*, 206−209.

(8) Russell, V. J.; Hold, G. L.; Pryde, S. E.; Rehbein, H.; Quinteiro, J.; Rey-Mendez, M.; Sotelo, C. G.; Perez-Martin, R. I., Santos, A. T.; Rosa, C. Use of Restriction Fragment Length Polymorphism to Distinguish Between Salmon Species. *J. Agric. Food Chem.* **2000**, *48*, 2184−2188.

(9) Hold, G. L.; Russell, V. J.; Pryde, S. E.; Rehbein, H.; Quinteiro, J.; Rey-Mendez, M.; Sotelo, C. G.; Perez-Martin, R. I.; Santos, A. T.; Rosa C. Validation of a PCR−RFLP Based Method for the Identification of Salmon Species in Food Products. *Eur. Food Res. Technol.* **2001**, *212*, 385−389.

(10) Quinteiro, J.; Vidal, R.; Izquierdo, M.; Sotelo, C. G.; Chapela, M. J.; Perez-Martin, R. I.; Rehbein, H.; Hold, G. L.; Russell, V. J.; Pryde, S. E.; Rosa, C.; Santos, A. T.; Rey-Mendez, M. Identification of Hake Species (*Merluccius* Genus) using Sequencing and PCR−RFLP Analysis of Mitochondrial DNA Control Region Sequences. *J. Agric. Food Chem.* **2001**, 49, 5108−5114.

(11) Sotelo, C. G.; Calo-Mata, P.; Chapela, M. J.; Perez-Martin, R. I.; Rehbein, H.; Hold, G. L.; Russell, V. J.; Pryde, S.; Quinteiro, J.; Izquierdo, M.; Rey-Mendez, M.; Rosa, C.; Santos, A. T. Identification of Flatfish (*Pleuronectiforme*) Species Using DNA-Based Techniques. *J. Agric. Food Chem.* **2001**, *49*, 4562−4569.

(12) Sanjuan, A.; Comesana, A. S. Molecular Identification of Nine Commercial Flatfish Species by Polymerase Chain Reaction-Restriction Fragment Length Polymorphism Analysis of a Segment of the Cytochrome b Region. *J. Food Protein* **2002**, *65*, 1016−1023.

(13) Partis, L.; Wells, R. J. Identification of Fish Species Using Random Amplified Polymorphic DNA (RAPD). *Mol. Cell. Probes* **1996**, *10*, 435−441.

(14) Bielawski, J. P.; Pumo, D. E. Randomly Amplified Polymorphic DNA (RAPD) Analysis of Atlantic Coast Striped Bass. *Heredity* **1997**, *78*, 32−40.

(15) Asensio, L.; Gonzalez, I.; Fernandez, A.; Rodriguez, M. A.; Lobo, E.; Hernandez, P. E.; Garcia, T.; Martin, R. Application of Random Amplified Polymorphic DNA (RAPD) Analysis for Identification of Grouper (*Epinephelus guaza*), Wreck Fish (*Polyprion americanus*) and Nile Perch (*Lates niloticus*) Fillets. *J. Food Protein* **2002**, *65*, 432−435.

(16) Dooley, J. J.; Sage H. D.; Brown, H. M.; Garrett, S. D. Improved Fish Species Identification by Use of Lab-on-a-Chip Technology. *Food Control.* **2005**, *16*, 601−607.

Received for review December 10, 2004. Revised manuscript received March 2, 2005. Accepted March 3, 2005. The financial support of the Food Standards Agency, for which this work forms part of project Q01069, is gratefully acknowledged.

JF047917S