# Identification of Hake Species (*Merluccius* Genus) Using Sequencing and PCR-RFLP Analysis of Mitochondrial DNA Control Region Sequences

Javier Quinteiro,<sup>\*,†</sup> Rodrigo Vidal,<sup>†</sup> Mónica Izquierdo,<sup>†</sup> Carmen G. Sotelo,<sup>‡</sup> María José Chapela,<sup>‡</sup> Ricardo I. Pérez-Martín,<sup>‡</sup> Harmut Rehbein,<sup>§</sup> Georgina L. Hold,<sup>∥</sup> Valerie J. Russell,<sup>∥</sup> Susan E. Pryde,<sup>∥</sup> Carla Rosa,<sup>⊥</sup> Ana T. Santos,<sup>⊥</sup> and Manuel Rey-Méndez<sup>†</sup>

Departamento de Bioquímica e Bioloxía Molecular, Facultade de Bioloxía, Universidade de Santiago de Compostela, 15706 Santiago de Compostela, Spain, Instituto de Investigaciones Marinas (CSIC), Eduardo Cabello 6, 36208 Vigo, Spain, Bundesforchungsanstalt für Fischerei, Institut für Biochemie und Technologie, Hamburg, Germany, Rowett Research Institute, Greenburn Road, Aberdeen, AB21 9SB Scotland, U.K., and Instituto de Investigação das Pescas e do Mar, Avenue Brasilia, 1400 Lisboa, Portugal

The use of DNA-based methodologies in identification of hake species belonging to the *Merluccius* genus was shown to be successful. A short fragment of the left hypervariable domain of the mitochondrial control region was amplified, sequenced, and digested from 11 hake species. The hake-specific PCR product, due to its limited size, was obtained in a variety of tissue samples with different levels of DNA concentration and degradation, including sterilized food products. On the basis of this phylogenetically informative 156-bp sequence were selected four restriction enzymes (*Apol, Ddel, Dra*III, and *Mbo*II) that allow the hake species discrimination. Species identification by phylogenetic analysis of sequences or by PCR–RFLP methodologies is useful in a variety of scenarios including authentication of thermally processed food, detection of food components, and species determination of individuals whose morphological characters are removed.

Keywords: Species identification; hake; Merluccius; PCR-RFLP; control region; seafood

## INTRODUCTION

Hake is a commercially and ecologically important group of 13 species of gadoids, classified into the genus Merluccius (Rafinisque, 1810), which constitutes a significant percentage of absolute fish consumption (1). External morphology in this group of species is highly conservative (2, 3), being laborious to identify individuals at species-level. Specific variances in their organoleptic features originate differences in consumer preferences and commercial values. Both issues call for strategies in ensuring product authenticity, adulteration, and improper description (4), particularly to prevent substitutions of most appreciated hake species, such as European hake (Merluccius merluccius; 5). Distribution areas included the Atlantic and Southern Oceans and the eastern Pacific margin (6), where the sympatric occurrence of two different species is common (3, 7, 8).

Morphological identification can be only carried out with whole fish after rigorous analysis. For this reason, new approaches must be used for species identification of incomplete fish with informative morphological characters removed, fresh or frozen portions of fish, pre-

 Table 1. List of Hake Species Used To Obtain Reference

 Sequences Including Geographic Origin and Number of

 Individuals Essayed (N)

· · ·			1	
species name	abbrev	common name	location	N
Merluccius albidus	MALB	offshore hake	NWA, USA	1
M. australis	MAUS	Austral hake	SEP, Chile	2
M. bilinearis	MBIL	silver hake	NWA, USA	3
M. capensis	MCAP	shallow-water cape hake	SEA, South Africa	2
M. paradoxus	MPAR	deep-water cape hake	SEA, South Africa	1
M. gayi	MGAY	Chilean hake	SEP, Chile	1
M. hubbsi	MHUB	Argentine hake	SWP, Argentina	1
M. merluccius	MMER	European hake	MED, NĔA, Spain	2
M. polli	MPOL	Benguela hake	SEA, Mauritania	2
M. productus	MPRO	Pacific hake	NEP, USA	2
M. senegalensis	MSEN	Senegalese hake	SEA, Senegal	2

cooked fillets, and ingredients in thermally and homogenized processed foods.

Molecular methodologies for fish species identification have been developed to circumvent limitations imposed by morphological characterization. Thus, specific isoelectric focusing (IEF) profiles of water-soluble proteins (9-11) and two-dimensional (2D) electrophoresis (5) allow the discrimination of raw fish. The main drawback of protein analysis is the loss of water solubility when samples to be analyzed came from thermally processed foods (11). Although, the use of enzymes or chemical cleavage of heat-denatured muscle proteins solve this problem, more resolution is needed to discriminate closely related species (12).

DNA-based techniques are the method of choice for species identification in thermally processed foods and

<sup>\*</sup> Corresponding author phone: 34-981-563100, ext 13317; fax: 34-981-596904; e-mail: bnjquint@usc.es.

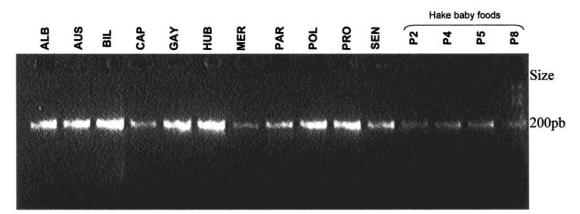
<sup>&</sup>lt;sup>†</sup> Universidade de Santiago de Compostela.

<sup>&</sup>lt;sup>‡</sup> Instituto de Investigaciones Marinas (CSIC).

<sup>&</sup>lt;sup>§</sup> Institut für Biochemie und Technologie.

<sup>&</sup>quot;Rowett Research Institute.

<sup>&</sup>lt;sup>1</sup> Instituto de Investigação das Pescas e do Mar.



**Figure 1.** Electrophoresis in agarose gel (2.5%) stained with ethidium bromide of PCR products obtained by amplification of a control region fragment with MERFPD1 and GADRPD1 primers. Amplifications were carried out in 11 species of *Merluccius* genus and four different baby food samples containing hake.

		*	20	*	40	*	60	*	
Maus30				тт	T			A :	78
Maus528								A	78
Mbi1032	.T1	r	G	A	T	TT	A		78
Mbill3	.T1	r		A	<b>T</b>	TT	A	c. :	78
Mbil03	.T7	r		A	T	TT		c. :	: 78
									. 78
Mpro21									: 78
Mgay11V	с	c.			T				: 78
								C. :	78
Mhub317									
Mcap	с	A		т				A	77
•								A	78
•									
Msen94									77
Mmer71									77
									. 77
									: 78
<b>2</b>									. 78
-								· · · · <del>·</del> · · · · · · · · · ·	
								• • • • • • • • • • • • • • • • • • •	. ,, : 78
P2C P4C								A	: 78
P4C P4A								· · · · A · · · · · · · · · · · · ·	. 78
P5A									. 78
P5A P5C								• • • • • • • • • • • • • • • • • •	. 78
								A	
POC								CTAGACTTGCtT	. /8
	LCCIA	S ALTAAA T	TATUTGIAT	SCUCIUA TTU	CITICCCCAAA	ACCCIANT.	TAAATCACCCG	JCIAGACIIGCUI	
	80	*	100	*	120	*	140	*	
Maus30	: 1	AG.	A			т			: 156
	: 1	AG.	A			т			: 156 : 155
Maus528 Mbi1032	:	AG. AG.	A			T TA		AG	: 155
Maus528 Mbil032 Mbil13	:	AG. AG. G.				TA TA			: 155
Maus528 Mbil032 Mbil13	:	AG. AG. G. 				TA TA		AG	: 155
Maus528 Mbil032 Mbil13 Mbil03	:	AG. AG. G. G. G.				TA A A			: 155 : 156 : 156
Maus528 Mbil032 Mbil13 Mbil03 Mprol8	:	AG. AG. G. G. G. G.				TA A A A			: 155 : 156 : 156 : 156
Maus528 Mbil032 Mbil13 Mbil03 Mpro18 Mpro21	:	AG. AG. G G G G G				TA A A A A		AG AG AG	: 155 : 156 : 156 : 156 : 156
Maus528 Mbil032 Mbil13 Mbil03 Mprol8 Mpro21	:	AG. AG. G. G. G. G. G.				TA A A A A		AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156
Maus528 Mbil032 Mbil13 Mbil03 Mpro18 Mpro21 Mgay11V	:	AG. AG. G. G. G. G. G. G. G.				TA AA AA AA AA AA		AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156
Maus528 Mbil032 Mbil13 Mbil03 Mpro18 Mpro21 Mgay11V Malb2	:	AG. AG. G. G. G. G. G. G. G. G.				T		.AG .AG .AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 156
Maus528 Mbil032 Mbil13 Mpro18 Mpro21 Mgay11V Malb2 Mhub317	:	A G. A G. 			G	T		.AG .AG .AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 156
Maus528 Mbil032 Mbil13 Mbil03 Mpr018 Mpr021 Mgayl1V Malb2 Mhub317 Mcap Mcap318	:	A G. A G. 			G. C.C	T		AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155
Maus528 Mbi1032 Mbi113 Mpr018 Mpr021 Mgay11V Malb2 Mhub317 Mcap318 Msen53	:	AG. AG. G. G. G. G. G. G. G. G. G.	A			TA. TA. A A A A TA. A A A		AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155 : 155 : 156
Maus528 Mbi1032 Mbi113 Mpr018 Mpr021 Mgay11V Malb2 Mhub317 Mcap318 Msen53	·	AG. AG. G. G. G. G. G. G. G. G. G. G. G. G.			 	TA. TA. A A A A TA. TA. A A A A A A A A A		AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155 : 156 : 155 : 156
Maus528 Mbil032 Mbil13 Mbil03 Mpro18 Mpro21 Mgay11V Malb2 Mhub317 Mcap Mcap318 Msen53 Msen94 Mmer71	:	AG. AG. G.				T		AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155 : 155 : 155 : 155
Maus528 Mbil032 Mbil13 Mbil03 Mpro18 Mpro21 Mgay11v Malb2 Mhub317 Mcap318 Msen53 Msen94 Mmer71 Mmer67	:	AG. AG. G.			 	T		AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155 : 155 : 155 : 155 : 155 : 155
Maus528 Mbil032 Mbil13 Mbil03 Mpro18 Mpro21 Mgay11V Malb2 Mhub317 Mcap318 Msen53 Msen94 Mmer71 Mmer67 Mpo192	:	AG. AG. G.				T		AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155 : 155 : 155 : 155 : 155 : 155 : 155 : 155 : 155
Maus528 Mbi1032 Mbi113 Mpr018 Mpr021 Mgay11V Malb2 Mhub317 Mcap318 Msen53 Msen94 Mmer71 Mmer67 Mp0192 Mp0105		AG. AG. G.	A			T	.c	AG AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155 : 155 : 155 : 155 : 155 : 155 : 156 : 156
Maus528 Mbi1032 Mbi113 Mpr018 Mpr021 Mgay11V Malb2 Mhub317 Mcap318 Msen53 Msen94 Mmer71 Mmer67 Mp0192 Mp0105 Mpar		A G. A G. 	A			TA. TA. A A A A TA. TA. A	.c. .c.	AG AG AG AG AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155 : 155 : 155 : 155 : 155 : 155 : 155 : 155 : 155
Maus528 Mbi1032 Mbi113 Mpr018 Mpr021 Mgay11V Malb2 Mhub317 Mcap318 Msen53 Msen94 Mmer71 Mmer67 Mp0192 Mp0105 Mpar P2C	·	A G. A G. 	A		 	TA. TA. A A A A TA. TA. A	.C	AG AG AG AG AG AG AG	: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155 : 155 : 155 : 155 : 155 : 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155 : 156 : 155 : 155 : 155 : 156 : 156 : 155 : 155 : 156 : 156 : 155 : 156 : 156 : 155 : 156 : 156 : 156 : 156 : 155 : 156 : 156
Maus528 Mbi1032 Mbi113 Mbi103 Mpro18 Mpro21 Mgay11V Malb2 Mhub317 Mcap Mcap318 Msen53 Msen94 Mmer71 Mmer67 Mpo192 Mpo105 Mpar P2C P4C	·	AG. AG. GG. G. G. G. G. G. G. G. G. G. G. G. 	A			T	.C		: 155 : 156 : 156 : 156 : 156 : 156 : 156 : 156 : 155 : 155 : 155 : 155 : 155 : 155 : 155 : 155 : 155 : 156 : 154 : 156 : 156
Maus528 Mbil032 Mbil13 Mbil03 Mpro18 Mpro21 Mgay11V Malb2 Mhub317 Mcap Mcap318 Msen53 Msen94 Mmer71 Mmer67 Mpo192 Mpo105 Mpar P2C P4C P4A	:	AG. AG. 				T	.c. .c. .c.	.AG .AG .AG .AG .AG .AG .AG .AG	$\begin{array}{c} : 155 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \end{array}$
Maus528 Mbil032 Mbil13 Mbil03 Mpro18 Mpro21 Mgayl1V Malb2 Mhub317 Mcap318 Msen53 Msen94 Mmer71 Mmer67 Mpo105 Mpar P2C P4C P4A P5A		A G. A G. 				T	.c. .c. .c.		$\begin{array}{c} : 155 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 156 \\$
Maus528 Mbi1032 Mbi113 Mpr018 Mpr021 Mgay11V Malb2 Mhub317 Mcap318 Msen53 Msen94 Mmer71 Mmer67 Mp0192 Mp0105 Mpar P2C P4C P4A P5A P5C		AG. AG. G.				TA. A A A A A T A A A A A A A A A A A A A A A A A T A.	.c. .c.	AG AG AG AG AG AG AG	$\begin{array}{c} : 155 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ \end{array}$
Maus528 Mbi1032 Mbi113 Mpr018 Mpr021 Mgay11V Malb2 Mhub317 Mcap318 Msen53 Msen94 Mmer71 Mmer67 Mp0192 Mp0105 Mpar P2C P4C P4A P5A P5C		AG. AG. 				TA. A. A. A. A. TA. 	.C	AG AG AG AG AG AG AG	$\begin{array}{c} : 155 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 155 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ : 156 \\ \end{array}$

**Figure 2.** Alignment of partial control region reference sequences from 19 individuals belonging to 11 different *Merluccius* species and homologous sequences obtained from processed food (P). Dots indicate identity with consensus sequence (bottom).

 Table 2. Restriction Patterns and Fragment Sizes Expected after Digestion of a Control Region Fragment Amplified with MER-F-PD1/GAD-R-PD1 Primers in *Merluccius* Species

species	ApoI r'AATT,y	<i>Dde</i> I C'TnA,G	<i>Dra</i> III CAC,nnn′GTG	MboII GAAGA(n)7,n'	pattern
M. albidus	164, 33	154, 43	157, 40	154, 43	AAAB
M. australis	164, 33	89, 65, 43		154, 43	AD0B
M. bilinearis		154, 43	157, 40	115, 43, 39	0AAD
M. capensis	164, 33	154, 43	157, 40	143, 43, 11	AAAA
M. gayi		154, 43	157, 40	104, 43, 39, 11	0AAC
M. hubbsi		154, 43		115, 43, 40	0A0D
M. merluccius	164, 32	101, 52, 43		153, 43	AB0B
M. paradoxus	164, 33			143, 43, 11	A00A
M. polli				154, 43	000B
M. productus	164, 33	154, 43		115, 43, 39	AA0D
M. senegalensis	164, 32	101,43,39,13	156, 40	153, 43	ACAB

samples preserved in nonoptimal conditions (11). Both thermal treatment and nuclease activity are responsible for the high levels of DNA degradation detected in processed food and poorly preserved samples, respectively. In these cases, the scarce and degraded DNA is available for analysis by means of Polymerase Chain Reaction (PCR). Therefore, short nucleotide sequences in amplified DNA and alternative methodologies to sequencing data, including restriction fragment length polymorphism (PCR-RFLP) and single-strand conformation polymorphism analysis (PCR-SSCP), provide information about the biological origin in a variety of foods. These methodologies have been used for species identification purposes in various commercial meat and fish samples (13, 14), snail meat (15), highly processed meat (16), whale and dolphin products (17), caviar (18, 19), and canned tuna (20-26).

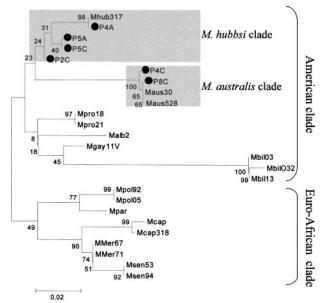
Mitochondrial DNA is a molecular marker widely used in phylogenetic analysis, population genetics, and species identification (27). The noncoding segment, namely, the control region, contains hypervariable sequences useful to analyze nucleotide variability in populations and close related species (28, 29). The left hypervariable domain of the control region has been used to elucidate phylogenetic relationships within the genus *Merluccius* (30, 31). These sequences provide the basis to search for informative short sequences, design of specific primers, and determination of methodologies for hake species identification in food products.

#### MATERIALS AND METHODS

**Samples.** A set of 19 individuals belonging to 11 species of the genus *Merluccius* (Rafinesque, 1810) was analyzed (Table 1). Two species with null or scarce commercial significance, *M. angustimanus* and *M. hernadezi* (species recently described), were unavailable for analysis. In addition, six 0.5-mL samples of baby foods containing hake were used to validate the protocol in a sterilized product.

**DNA Extraction.** DNA was extracted from 50 to 100 mg of frozen and ethanol-preserved muscle following standard protocols (*32*). In the case of DNA extraction from processed food samples, DNA was recovered from the aqueous phase by ultrafiltration using Microcon-100 (Amicon) (*23*). DNA was dissolved in a 20-50- $\mu$ L volume of sterilized water.

**Amplification.** Two primers for amplification of species belonging to the genus *Merluccius* were designed based on mitocondrial control region sequences (GenBank Accession Nos. AF112245–AF112255) (*30, 31*). Partial sequence was selected because of the limited length of successfully amplified sequences from degraded DNA, the low level of intra-specific nucleotide variability and the capacity to discriminate all hake species. The forward primer MERFPD1 (5'-TCAACCCAT-AATACWCATTCC-3') was specifically designed for *Merluccius* species, whereas the reverse primer GADRPD1 (5'-ATGGAC-



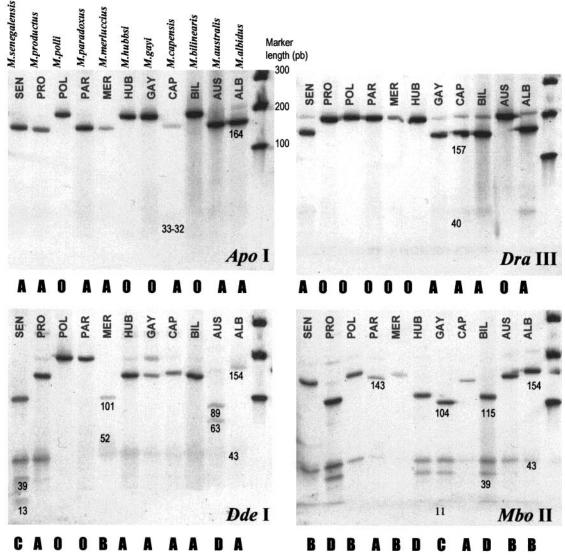
**Figure 3.** Phylogenetic tree elaborated from partial left domain of the control region sequences using Tamura-Nei distances and the of neighbor-joining method. Numbers in nodes indicate bootstrapping values (bootstrap replications = 1000). P sequences correspond to unknown processed food samples.

CTGAAGCTAGGCA-3') included a conserved sequence homologous to that other gadoids. Amplifications were performed in a reaction volume of 25  $\mu$ L of Promega Buffer × 1, 2.5 mM MgCl<sub>2</sub>, 200  $\mu$ M each dNTP, 0.1  $\mu$ M each primer, 0.625 unit/reaction *Taq* DNA polymerase (Promega), and 0.5–2  $\mu$ L of total DNA. The reaction was carried out in a Gene Amp PCR system 2400 (Perkin-Elmer), with an initial denaturing step of 3 min at 94 °C, followed by 30 cycles and each cycle with denaturation at 94 °C for 20 s, annealing at 50 °C for 30 s, extension at 72 °C for 50 s, and then a final extension at 72 °C for 5 min.

**Sequencing.** The PCR mixture was treated with the PCR Product Pre-sequencing Kit (Amersham-Pharmacia), and an aliquot of 2 mL of the PCR mixture was used for cycle sequencing with ABI PRISM dRhodamine Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems). Once the extension products were purified, electrophoresis was carried out in an ABI PRISM 377 DNA Sequencer (Applied Biosystems) using 6% polyacrylamide gels for fluorescent DNA sequencing.

Prior to sample loading, the pooled and dried reaction products were suspended in loading buffer (Applied Biosystems) containing five parts of deionized formamide to one part of 25 mM EDTA, pH 8.0, and 50 mg/mL Dextran Blue (Applied Biosystems) with the gel electrophoresed for 2.5 h at 50 °C and 32 W.

**Sequence Analysis.** Sequences were aligned using Clustal W (*33*). Phylogenetic methods were applied to permit identi-



**Figure 4.** Restriction patterns observed in 11 species of hake after digestion of the control region PCR product with restriction enzymes *Apol*, *Dra*III, *Dde*I, and *Mbo*II and separated by PAGE. Fragment size (bp) was indicated below observed silver stained bands.

fication of problem samples by clustering. The unknown sequences were included in a monophyletic clade with those closely related sequences belonging to the same species (13, 16, 23). Genetic distances were calculated using the Tamura-Nei model (34), and phylogenetic trees were constructed by the neighbor-joining method (35), rooted at midpoint, and evaluated by boostraping analysis with the MEGA program (36). To test the effects of using alternative phylogenetic reconstruction methods in clustering results, parsimony and maximum likelihood methods were also carried out with the PHYLIP package (37). Polymorphism information was obtained with DNAsp program (38). After the search for species-specific restriction sites, expected digestion profiles were calculated to construct combined restriction patterns informative for species discrimination.

**PCR**–**RFLP Analysis.** To perform the RFLP analysis, the PCR products were concentrated, and primers were removed using Microcon-50 (Amicon). Aliquots of 50–75 ng of PCR fragment were digested with *Apo*I, *Dde*I, *Dra*III, and *Mbo*II in a volume of 20  $\mu$ L, following manufacturer (New England Biolabs) standard conditions. Separation of DNA fragments was carried out in a GeneGel Excel 12.5 (T 12.5%, C 2%) (Amersham Pharmacia Biotech), loading 6.5  $\mu$ L of the digestion products on the gels. Anode buffer was 0.4% SDS and 0.45 M Tris acetate, pH 8.3; cathode buffer was 0.6% SDS and 0.08 M tricine. The electrophoresis was carried out on a GenePhor (Amersham Pharmacia Biotech) with a temperature of the

cooling plate of 15 °C and a voltage of 200 V. The run was stopped when the tracking dye reached the anode edge of the gel. DNA restriction fragments were visualized by a silver staining DNA kit (Amersham Pharmacia Biotech).

#### **RESULTS AND DISCUSSION**

The described set of primers MERFPD1/GADRPD1 successfully amplified a control region fragment using DNA extracts from raw and frozen as well as thermally processed food hake samples. This PCR product, located in the left hypervariable domain of the mitochondrial control region, was approximately 197 bp long (Figure 1).

Baby food is a homogenized and thermally processed food that include a sterilization step in their manufacture. Consequently, it is preferred to test the suitability of methodologies of species identification in comparison to other foods, containing hake, showing a weak DNA degradation level. The low concentration and highly severe degradation observed in DNA extracts from this processed food is comparable to those from other thermally processed food, such as canned tuna (*23*). The average size of this DNA is below 500 bp; therefore, a sufficient number of DNA molecules were isolated to be used as a suitable template, allowing the amplification of this 197-bp sequence. DNA extracted from frozen and ethanol-preserved tissues presents variability in concentration and degradation levels but constitutes a more efficient template, resulting in a higher yield of PCR product than DNA template from baby food (Figure 1).

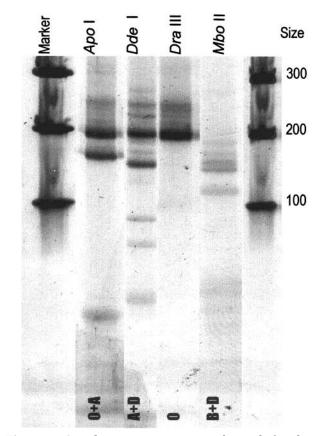
Baby food labels declare an approximate content of 15% hake tissues. The rest of animal (including related fish) and vegetable tissues contribute to the total DNA extracted from baby food. The specific PCR products obtained suggest the affinity of the primers sequence to the *Merluccius* template DNA even in low proportions, without interference of co-isolated DNA.

The partial sequences of the hypervariable left domain of the mitochondrial control region are representative of these domain features in hake, such as high proportion of AT and moderate variability (Figure 2) (30,31). In the 19 reference sequences of hakes were detected 37 polymorphic sites, including 32 parsimony informative sites, with a nucleotide diversity ( $\pi$ ) of 0.077 (SD = 0.006). Distance values ranged from 0.013 to 0.181 in inter-specific pairwise comparisons with a mean value of 0.086; whereas in intra-specific comparisons, the mean value was only 0.005, ranging from 0 to 0.021. Although the number of co-specific individuals analyzed was low, these values can be considered informative of the divergence level in intra- and inter-specific comparisons and are in congruence with results from genetic variation in hake population studies (31). Exceptionally, the closest related species, M. merluccius and M. senegalensis, showed inter-specific distance values comparable to intra-specific values obtained within M. hubbsi clade.

A phylogenetic reconstruction of the relationships within the genus *Merluccius*, based on this short control region sequence, reflects the same topology obtained with the complete left domain in the Euro-African clade, whereas in the American clade an alternative topology was recovered (Figure 3) (*30*, *31*). Differences in topology and branching support are due to the reduced number of variable positions caused by the limited size of the amplified fragment. However, the recovery of correct tree branching patterns is not essential for identification purposes. What is critical in this particular application of phylogenetic methods is the accurate definition of species in monophyletic clades and the estimation of the closest genetic distances.

Sequences obtained from baby food samples were introduced in the phylogenetic analysis together with reference sequences from each hake species analyzed. Each problem sequence was grouped with his closest relative reference sequence in the same monophyletic clade. The samples P2C, P4A, P5A, and P5C were located in the *M. hubbsi* clade, whereas P4C and P8C were grouped in *M. australis* clade. Two samples taken from the same recipient, P4A and P4C, give sequences belonging to different species, suggesting that a mixture of species was detected (Figure 3). The tree topologies obtained from parsimony analysis and maximum likelihood methods suggest the same clustering pattern.

On the basis of DNA sequence data, the selected set of diagnostic restriction enzymes provides a nonsequencing method for species identification. The expected results of PCR-RFLP analysis of the MERFPD1/ GADRPD1 fragment allow the hake species to be discriminated (Table 2). After digestion, congruent profiles were observed in all cases (Figure 4). Some short



**Figure 5.** Complex restriction patterns obtained after digestion of a PCR product amplified from a processed food containing a homogeneous mixture of tissues belonging to different hake species. Combined electrophoretic profiles suggest the presence of *M. hubbsi* and *M. australis* PCR products.

fragments generated, for instance after digestion of *M. gayi* PCR product with *Mbo*II (11 bp), are silver-stained with poor intensity. Nevertheless, the relative mobility of the longer fragments permits a correct resolution of the fragments profile. The composed restriction patterns, including the four enzymes, *ApoI, DdeI, Dra*III, and *Mbo*II, are species-specific and are conservative within each species (Figure 4). Searching and selection of restriction sites were restricted to partial sequences where intra-specific variability was not detected. Although hake control region variability is lower than in other teleosts (*39*), a higher number of individuals must be essayed to verify the absence of new haplotypes, especially those whose nucleotide variability is located within selected restriction sites.

In homogenized foods is the relatively frequent complete mixture of different species tissues contained in a recipient. Particularly, it is ordinary to find a mixture of closely related species belonging to a common commercial denomination (i.e., hake). Consequently, total DNA extracted contains a mixture of DNAs belonging to different species, and then PCR products are composed of diverse sequences. Sequencing methodologies applied to these complex PCR products result in technically hard to interpret electropherograms, with the exception of mixed PCR products where there is a particular overabundant sequence. In contrast, PCR-RFLP methodologies allow the analysis of these mixtures. After digestion, combined electrophoretic patterns can be resolved, providing evidence of the presence of different species (Figure 5). In a mixed sample P4, two types of results were obtained after sequencing: mixed and unreadable sequences and *M. hubbsi* sequences; whereas that in a single step, after PCR-RFLP, was possible to identify the mixture of *M. hubbsi* and *M. australis* restriction patterns (Figure 5). These alternative results suggest the limitations of DNA sequencing methods for identification purposes in a food containing an homogeneous species mixture.

A validation exercise was performed to evaluate the efficacy of the PCR-RFLP methodology. Problem samples of different species and baby foods were distributed within the laboratories participants, congruent results being achieved. Although the difficulty to obtain clear restriction patterns from eventual poor quality PCR products was observed.

The strength and limitations of DNA sequencing and PCR-RFLP methods for species identification purposes depend on the particular analysis to be carried out. A high number of tissues can be processed by PCR-RFLP analysis, although the eventual intra-specific variability in the restriction site selected must be take into account. The overall efficacy of sequencing methods is reduced in the case of mixed samples.

### LITERATURE CITED

- Yearbook of Fishery Statistics 1994; FAO: Rome, Italy, 1996; p 357.
- (2) Inada, T. Studies on the Merluccid fishes. *Bull. Far Seas Fish. Res. Lab.* **1981**, *18*, 1–172.
- (3) Cohen, D. M.; Inada, T.; Iwamoto, T.; Scialabba, N. FAO Species Catalogue. Gadiform fishes of the world (Order Gadiformes); FAO Fish Synopsis; FAO: Rome, Italy, 1990; p 442.
- (4) Hargin, K. D. Authenticity issues in meat and meat products. *Meat Sci.* **1996**, *43*, 277–289.
- (5) Piñeiro, C.; Barros-Velázquez, J.; Sotelo, C. G.; Pérez-Martín, R. I.; Gallardo, J. M. Two-dimensional electrophoretic study of the water-soluble protein fraction in white muscle of gadoid fish species. *J. Agric. Food Chem.* **1998**, *46*, 3991–3997.
- (6) Soliman, I. A. Variations of fishes of genus *Merluccius* in Atlantic Ocean and Mediterranean. *Sea Acta Ichthiol. Piscatoria* 1973, *3*, 29–64.
- (7) Howes, G. J. Biogeography of gadoid fishes. *J. Biogeogr.* 1991, *18*, 595–622.
- (8) López-Abellán, L. J.; Ariz-Tellería, J. Aspectos generales de la distribución y biología de las especies del género *Merluccius* Rafinesque, 1810, en aguas de Senegal y Gambia (16°00' N-12°25' N). *Bol. Inst. Esp. Oceanogr.* 1993, 9, 101-121.
- (9) Mackie, I. M. Identifying species of fish. Anal. Proc. 1990, 27, 89–92.
- (10) Rehbein, H. Electrophoretic techniques for species identification of fishery products. *Z. Lebensm.-Unters. Forsch.* **1990**, *191*, 1–10.
- (11) Sotelo, C.; Piñeiro, C.; Gallardo, J. M.; Pérez-Martin, R. I. Fish species identification in seafood products. *Trends Food Sci. Technol.* **1993**, *4*, 395–401.
- (12) Mackie, I. M.; Chalmers, M.; Reece, P.; Scobbie, A. E.; Ritchie, A. H. The application of electrophoretic techniques to the identification of species of canned tuna and bonito. In *Pelagic Fish: The resource and its explotation*, Burt, J. R., Hardy, R., Whittle, K. J., Eds.; Elsevier Science: Amsterdam, The Netherlands, 1993; p 511.
- (13) Bartlett, S. E.; Davidson, W. S. FINS (Forensically Informative Nucleotide Sequencing): A procedure for identifying the animal origin of biological specimens. *Biotechniques* **1992**, *12*, 408–411.
- (14) Meyer, R.; Candrian, U. PCR-based DNA analysis for thr identification and characterization of food components. *Lebensm.–Wiss. Technol.* **1996**, *29*, 1–9.

- (15) Borgo, R.; Souty-Grosset, C.; Bouchon, D.; Gomot, L. PCR-RFLP analysis of mitochondrial DNA for identification of snail meat species. *J. Food Sci.* **1996**, *61*, 1–4.
- (16) Forrest, A. R. R.; Carnegie, P. R. Identification of gourmet meat using FINS (Forensically Informative Nucleotide sequencing). *Biotechniques* 1994, 17, 24-26.
- (17) Baker, C. S.; Cipriano, F.; Palumbi, S. R. Molecular genetic identification of whale and dolphin products from commercial markets in Korea and Japan. *Mol. Ecol.* **1996**, *5*, 671–685.
- (18) De Salle, R.; Birstein, V. J. PCR identification of black caviar. *Nature* **1996**, *381*, 197–198.
- (19) Rehbein, H.; Mackie, I. M.; Pryde, S.; González-Sotelo, C.; Pérez-Martín R.; Quinteiro, J.; Rey-Méndez, M. Fish species identification in canned tuna by PCR–SSCP: Validation by a collaborative study and investigating of intra-species variability of the DNA-patterns. *Food Chem.* **1999**, *64*, 263–268.
- (20) Unseld, M.; Beyermann, B.; Brandt, P.; Hiesel, R. Identification of the species origin of highly processed meat products by mitochondrial DNA sequences. *PCR Methods Appl.* **1995**, *4*, 241–243.
- (21) Rehbein, H.; Mackie, I. M.; Pryde, S.; González-Sotelo, C.; Pérez-Martín R.; Quinteiro, J.; Rey-Méndez. M. Fish species identification in canned tuna by DNA Analysis (PCR-SSCP). *Inf. Fischwirtsch.* **1995**, *42*, 209–212.
- (22) Ram, J. L.; Ram, M. L.; Baidoun, F. F. Authentication of canned tuna and bonito by sequence and restriction site analysis of polymerase chain reaction products of mitochondrial DNA. J. Agric. Food Chem. 1996, 44, 2460–2467.
- (23) Quinteiro, J.; Sotelo, C. G.; Rehbein, H.; Pryde, S. E.; Medina, I.; Pérez-Martín R. I.; Rey-Méndez, M.; Mackie, I. M. The use of mtDNA direct PCR-sequencing and PCR-RFLP methodologies in species identification of canned tuna. J. Agric. Food Chem. **1998**, 46, 1662–1669.
- (24) Rehbein, H.; Mackie, I. M.; Pryde, S.; González-Sotelo, C.; Pérez-Martín R.; Quinteiro, J.; Rey-Méndez, M. Comparison of different methods to produce singlestranded DNA for identification of canned tuna by single-strand conformation polymorphism analysis. *Electrophoresis* **1998**, *19*, 1381–1384.
- (25) Mackie, I. M.; Pryde, S.; González-Sotelo, C.; Medina, I.; Pérez-Martín R.; Quinteiro, J.; Rey-Méndez, M.; Rehbein, H. Challenges in the identification of species of canned fish. *Trends Food Sci. Technol.* **1999**, *10*, 9–14.
- (26) Rehbein, H.; González-Sotelo, C.; Pérez-Matín, R.; Quinteiro, J.; Rey-Méndez, M., Pryde, S., Mackie, I. M.; Santos, A. T. Differentiation of sturgeon caviar by single strand conformation polymorphism (PCR–SSCP) analysis. Arch. Lebensmittelhyg. **1999**, *50*, 1–4.
- (27) Avise, J. C. *Molecular markers, natural history and evolution*; Chapman and Hall: New York, 1994; p 511.
- (28) Chow, S.; Okamoto, H.; Uozumi, Y.; Takeuchi, Y.; Takeyama, H. Genetic stock structure of the swordfish (*Xiphias gladius*) inferred by PCR–RFLP analysis of the mitochondrial DNA control region. *Mar. Biol.* **1997**, *127*, 359–367.
- (29) Faber, J. E.; Stepien, C. A. The utility of mitochondrial sequences for analyzing phylogenetic relationships among populations, species, and genera of the Percidae. In *Molecular systematics of fishes*; Kocher, T. D., Stepien, C. A., Eds.; Academic Press: London, 1997; pp 129– 143.
- (30) Quinteiro, J.; Vidal, R.; Rey-Méndez, M. Phylogeny and biogeographic history of hake (genus *Merluccius*) inferred from mitochondrial DNA control-region sequences. *Mar. Biol.* **2000**, *136*, 163–174.
- (31) Vidal, R. Filogeografía molecular del género Merluccius y estructura genética poblacional de la merluza chilena (*Merluccius gayi*) y de la merluza europea (*Merluccius merluccius*). Thesis, University of Santiago de Compostela, Santiago de Compostela, Spain, 1999.

- (32) Kocher, T. D.; Thomas, W. K.; Meyer, A.; Edwards, S. V.; Pääbo, S.; Villablanca, F. X.; Wilson, A. C. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 6196–6200.
- (33) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: improving the sensibility of progressive multiple sequence alignement through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994, 22, 4673–4680.
- (34) Tamura, K.; Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **1993**, *10*, 512–526.
- (35) Saitou, N.; Nei, M. The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406-425.
- (36) Kumar, S.; Tamura, K.; Nei, M. MEGA: *Molecular Evolutionary Genetics Analysis, version 1.01*; Pennsylvania State University: University Park, PA, 1993.

- (37) Felsenstein, J. PHYLIP (Phylogeny Inference Package) version 3.5c; Distributed by the author; Departament of Genetics, University of Washington: Seattle, WA, 1995.
- (38) Rozas, J.; Rozas, R. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **1999**, *15*, 174– 175.
- (39) Lee, W.; Conroy, J.; Howell, W. H.; Kocher, T. D. Structure and evolution of teleost mitochondrial control regions. *J. Mol. Evol.* **1995**, *41*, 54–66.

Received for review March 29, 2001. Revised manuscript received August 7, 2001. Accepted August 7, 2001. The authors thank the European Commission for the financial support for this study (FAIR CT97 3061).

JF010421F