JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

ITS1-rDNA-Based Methodology To Identify World-Wide Hake Species of the Genus *Merluccius*

MONTSE PÉREZ,† JUAN M. VIEITES,‡ AND PABLO PRESA*,†

Department of Biochemistry, Genetics, and Immunology, Faculty of Biology, University of Vigo, 36310, Vigo, Spain, ANFACO, Asociación Nacional de Fabricantes de Conservas de Pescados y Mariscos, 36310, Vigo, Spain

Species-specific DNA-based tags are valuable tools for the management of both fisheries and commercial fish products. In this study, we have developed a two-step molecular tool to detect the presence of hake DNA (*Merluccius* spp.) and to identify the exact hake species present in an blind sample. The first test involves PCR amplification of an ITS1-rDNA fragment of 193 bp using nested primers that are interspecifically conserved in *Merluccius* spp. and Atlantic cod, *Gadus morhua*. The second test consists of the PCR amplification of a 602−659 bp DNA fragment spanning part of the ribosomal cluster 18S−ITS1−5.8S and digesting it with four restriction enzymes whose targets map at interspecifically nonconserved sites of the ITS1. Alternatively, the identification of hake species can be achieved by FINS or BLAST, using the nucleotide sequence of either the whole ITS1 sequence or its nested fragment of 193 bp. Because of their high reproducibility and ease of execution, these procedures allow for routine analysis and constitute high reliable tools for the rapid identification of 12 species of hake.

KEYWORDS: Hakes; *Merluccius* spp.; species identification; preliminary exclusion test; RFLPs; ITS1-rDNA marker; food science; traceability

INTRODUCTION

The genus *Merluccius* is composed of at least 14 species of hakes that occupy most temperate and tropical continental shelves except the Asian shores of the Pacific Ocean (*1*, *2*). Most hakes have been heavily fished along the last century (*3*), and their fisheries have attracted considerable interest from marine ecologists, fishery scientists, and commercial managers. In particular, there are key unresolved issues related to their conservation, exploitation, and traceability. However, the lack of a conspicuous identification key for each species seriously limits the assessment of both their commercial importance and the impact of fisheries on their sustainability. For instance, the existence of broad areas of bathymetric overlap between hakes (*1*) results in the simultaneous catching of two species in the same area, thus hampering independent management of each species' fishery. The industrial and legal sectors also face difficulties in regulating the commerce of products from mixed-species fisheries. The marked differences in price and marketability between hakes, together with declining fishing catches, increase the opportunities for fraudulent substitution with cheaper species of hakes or similar taxa (*4*).

The development of diagnostic tools for the unambiguous identification of hakes is a technology that would benefit both basic research studies on species biogeography and hybridization as well as applied fields such as fishery management, conservation genetics, fishery forensics, and commercial traceability. Even though experienced researchers can identify hakes using morphology (*1*), many samples such as those from museum collections, historical fishery surveys, or processed commercial fish products are devoid of morphological traits. Allozyme electrophoresis has been the most popular method used to describe genetic variation in hakes (*5*, *6*). However, the frequency dependence of the interspecific genetic distinctiveness, together with the low allelic variation of allozymes, do not allow us to unambiguously assign individuals to species for many practical purposes. Biochemical analyses based on species-specific sarcoplasmic proteins, using techniques such as isoelectrofocusing (IEF), two-dimensional electrophoresis (2DE), and SDS−PAGE, have been adapted to assist in the industrial identification of some hake species (*7*−*9*). However, most of those identification methods are based on tissue-dependent protein analyses, which depend heavily on the heat lability of proteins or on the maintenance of high levels of biological activity. Although species-specific heat-resistant muscle proteins could partially overcome such problems, closely related species usually share identical or apparently identical protein sequences, i.e., replacements between basic amino acids that do not bring about modifications in the net charge of the protein can appear as isoforms by electrophoresis.

Most of the disadvantages of protein-based identification methods can be circumvented using DNA-based techniques (*10*).

* To whom correspondence should be addressed. Telephone/Fax: +34-986-812567. E-mail: pressa@uvigo.es.
† University of Vigo.
‡ ANFACO.

10.1021/jf048012h CCC: $30.25
Published on Web 06/04/2005

Table 1. Names, Natural Ranges, and Sampling Locations of Hake Species from the Genus *Merluccius*

hake species	codes	common names	ranges	country[a]	sampling coordinates
Merluccius merluccius	ME	European hake	21–62° N	Spain	37° 35′ N/08° 50′ W
				Italy*	38° 03′ N/12° 56′ E
				United Kingdom*	55° 30′ N/04° 36′ E
Merluccius senegalensis	SE	Senegalese hake	10–33° N	Senegal*	15° 01′ N/18° 00′ W
				Namibia	18° 10′ N/16° 20′ W
				Mauritania*	21° 40′ N/17° 55′ W
Merluccius polli	PO	Benguela hake	20° N–19° S	Senegal*	15° 01′ N/18° 00′ W
				Mauritania	19° 37′ N/17° 06′ W
				Spain–Morocco*	27° 15′ N/14° 10′ W
Merluccius capensis	CA	shallow-water cape hake	0–34° S	Angola*	17° 10′ S/ 11° 21′ W
				Namibia	24° 10′ S/14° 23′ E
				South Africa*	25° 33′ S/15° 13′ E
Merluccius paradoxus	PA	deep-water cape hake	south 22° S	South Africa*	25° 33′ S/15° 13′ E
				South Africa*	34° 10′ S/17° 10′ E
Merluccius productus	PR	Pacific hake	25–51° N	Canada*	48° 08′ N/122° 20′ W
				Canada	49° 10′ N/123° 10′ W
				Canada *	50° 00′ N/125° 06′ W
Merluccius gayi	GA	Peruvian hake	3–10° S	Perú-Chile*	08° 50′ S/80° 00′ W
		Chilean hake	23–47° S	Chile*	24° 40′ S/70° 50′ W
				Perú–Chile	30° 00′ S/71° 55′ W
Merluccius australis	AU	Antarctic queen hake	40–57° S	Chile*	41° 20′ S/74° 35′ W
		New Zealand hake	south 40° S	Australia	43° 40′ S/169° 25′ E
		Austral hake	south 51° S	United Kingdom*	52° 40′ S/63° 35′ W
Merluccius hubbsi	HU	Patagonian hake	25–54° S	Argentina*	46° 30′ S/60° 45′ W
				Argentina	48° 30′ S/61° 30′ W
				United Kingdom*	53° 00′ S/61° 10′ W
Merluccius albidus	AL	offshore hake	20–35° N	United States*	35° 21′ N/70° 50′ W
				United States*	37° 21′ N/73° 33′ W
Merluccius hernandezi	HE	California hake	5–23° N	Mexico*	29° 50′ N/113° 20′ W
Merluccius bilinearis	BI	silver hake	36–47° N	United States*	39° 00′ N/73° 10′ W
				United States*	40° 40′ N/72° 00′ W
				United States	42° 30′ N/68° 33′ W

[a] The asterisks indicate the samples containing the individuals whose ITS1 were sequenced to calibrate the identification method. About 20 individuals were collected at each sampling location.

Ancient DNA samples or commercial products often contain DNA that can be amplified and compared to known "sample types", provided they had been standardized with reconstruction methods developed for phylogenetic analyses (*11*). Two previous DNA-based methods developed to identify hakes involved the use of four and seven restriction enzymes to digest a part of the mtDNA control region (*12*) and a cytochrome *b* fragment (*13*), respectively. However, the extremely large variation characterizing the left domain of the mtDNA control region (*14*) and the lack of species-specific restriction patterns of cytochrome *b* for several pairs of closely related hakes (*13*) seriously limit the utility of those methods for the full diagnostic identification within genus *Merluccius*.

The ITS1 spacer of the rDNA gene family accounts for hundreds of copies in fish genomes (*15*) and has the property to evolve concertedly within species and independently between species (*16*). Because concerted evolution results in sequence homogenization across most members of the rDNA family within reproductive units (*17*), the ITS1-rDNA spacer has been successfully applied in fish phylogeography (*18*) as well as in phylogenetic inference (*19*). This spacer is therefore one of the most promising DNA regions for species authentication and identification of closely related taxa (*20*). The goal of this work is to document the development of two ITS1-based diagnostic tools that allow detecting DNA from *Merluccius* and identifying 12 species of hakes from this genus, respectively.

MATERIALS AND METHODS

Sampling and Morphological Identification of *Merluccius* spp. To assess the impact of intraspecific genetic variation that could weaken the diagnostic signal, we sampled each of 12 hake species at distant sites of their oceanographic distributions (**Table 1**). This worldwide sampling was performed in cooperation with local fishermen, commercial factory ships, and research vessels across Euro-African (East-Atlantic) and American (West-Atlantic and Pacific) fisheries. Approximately 20 specimens per sample were frozen upon collection, and their GPS codes were recorded on board. Whole specimens were boiled to facilitate bone cleaning and were identified using species-specific morphological traits. Shape and length of clean structures such as otoliths, abdominal vertebrae (parapophysis), crane, and pectoral fins were inspected with optical microscopy to classify them according to criteria previously established for this genus (*1*). The ITS1-rDNA of two identified individuals per species was sequenced to calibrate the identification method. A total of 15 specimens per sample and 1−3 samples per species were used to assess the reproducibility of the identification method.

DNA Extraction and Purification. Genomic DNA was extracted by homogenization of 100 mg of gill tissue following an optimized DNA extraction method that combines the salting-out method (*21*) with the standard phenol/chloroform method (*22*) and is suitable for removing the mucopolysaccharides present in fish tissues.

Amplification and Sequencing of ITS1-rDNA. The ITS1-rDNA spacer was PCR-amplified from total DNA using a set of primers selected from coding regions of the 18S (*XelaITS1.1*: 5′-AAGTAAAAGTCGTAACAAGGTTTCCGTAGG-3′) and the 5.8S (*OnmyITS1.2*: 5′-CAAGCCGAGTGATCCACCGC-3′) genes of *Xenopus laevis* (*23*) and *Salmo gairdneri* (*24*), respectively. PCR amplifications of 50 μL containing 3 mM $MgCl_2$ were performed at 95 °C for 10 min and 35 cycles of 95 °C for 50 s, 55 °C for 40 s, and 72 °C for 2 min and 30 s, followed by a final step at 72 °C for 30 min. Amplicons of two individuals per species (**Table 1**) were purified from preparative gels (Marligen Biosciences, Valencia, Spain). Double-stranded DNA sequences were prepared in double with both the BigDye Terminator Cycle Sequencing Standard and the dGTP BigDye Terminator Ready Reaction Kit, using the primers *XelaITS1.1* and *OnmyITS1.2*. Sequenc-

ing reactions consisted of a denaturing cycle of 98 °C for 5 min, followed by 30 cycles of 96 °C for 30 s, 50 °C for 20 s, and 65 °C for 3 min and 30 s. Sequences were electrophoresed in an ABI Prism 310 DNA sequencer (Applied Biosystems, Madrid, Spain). All polymorphic sites were verified by re-amplifying all templates with two different T*aq* polymerases (Promega, Barcelona, and Eppendorf, Madrid, Spain) and resequencing them as described above. The consensus ITS1 sequence of each species was derived from single-strand chromatograms per template DNA using the CHROMAS software available online (http://www.technelysium.com.au/chromas.html).

Selection of Specific Primers for the Genus *Merluccius*. The 3′ end of the 18S gene and the 5′ end of the 5.8S gene were used to align the ITS1 sequences of 12 species of hakes using the SeqLab program from the GCG software package (*25*). This alignment allowed the identification of highly conserved regions between species, which were used to select several pairs of nested primers within the interspecifically conserved regions of the ITS1, using Oligo 4.05 (*26*). To check for the specificity of the primers to PCR-amplify exclusively in the genus *Merluccius*, all of them were also tested in closely related taxa (tailed hakes, *Macruronus novaezelandiae* and *Macruronus magellanicus*; Atlantic cod, *Gadus morhua*), as well as in distantly related taxa (salmonids, *Salmo salar*, *Salmo trutta*, and *Oncorhynchus mykiss*; flatfishes, *Scophthalmus maximus*, *Scophthalmus rhombus*, and *Platichthys flexus*; and mollusks, *Octopus vulgaris* and *Mytilus galloprovincialis*). All PCR products were visualized in 3% agarose gels to verify the length of the amplification products.

Selection of Diagnostic ITS1 Targets and Establishment of Species-Specific Restriction Patterns. The restriction maps of the 602−659 bp fragments spanning the ITS1 spacer were developed with WEBCUTTER 2.0 (*27*) and the enzymatic database REBASE (*28*) (New England Biolabs, Barcelona, Spain). The comparison of ITS1 restriction maps from each species allowed the selection of a minimum number of restriction enzymes necessary to achieve a complete discrimination of species. The consistency of the restriction patterns predicted was assessed by digesting 15 individuals per sampling site (45 individuals per species, including *Gadus morhua*) with four restriction enzymes. Independent digestions of 1 μg of the amplicon per each enzyme were allowed to proceed for 5 h, and the products were electrophoresed in 3% agarose gels at 70 V for 1 h. The restriction patterns of each species were established from the gels by (i) their comparison with a molecular weight marker, (ii) side by side comparisons of the patterns obtained for each pair of species, and (iii) verification of the exact size of fragments as determined from the ITS1 sequence of each species.

Phylogenetic Assessment of the Diagnostic Power. Because all conspecific samples showed the same restriction pattern for all of the enzymes, a rooted tree was constructed with one individual per species to verify the correct phylogenetic discrimination of species. The polymorphism parsimony method of DOLLOP from PHYLIP 3.6 (*29*) was used to find the most parsimonious trees built from a data matrix of presence−absence of restriction fragments (*30*). To search for the best tree, the analyses were performed by randomizing the input order of species through 100 iterations. The parsimonious trees recovered were summarized in a consensus tree using CONSENSE from PHYLIP 3.6 (*29*). The consensus nodal values were considered as a measure of the resolution power achieved at identifying species.

Species Identification using FINS and BLAST. The use of FINS (*31*) to identify hake species consisted of three steps. First, the DNA of test samples were extracted and purified following the procedures described above. Second, the nested fragment of 193 bp from the ITS1-rDNA spacer of 10−45 individuals per species were PCR-amplified and sequenced using capilar electrophoresis. Third, a phylogenetic reconstruction of the ITS1 (either from the whole ITS1 sequence or from the 193 bp nested fragment) was performed using sample types of the 12 species and the ITS1 sequences from all samples. This analysis allowed the confirmation of the assignment accuracy of samples to species upon their ascribing into each species' cluster.

A second alternative to the PCR−RFLPs identification method developed was the calculation of the expectation value of random sequence identity using the BLAST package (*32*). For this calculation, the ITS1 nucleotide sequences (from either the whole ITS1 sequence or the 193 bp sequence) of tested samples were compared to ITS1-rDNA sequences of all hakes made available in GenBank (accession numbers in **Figure 1**).

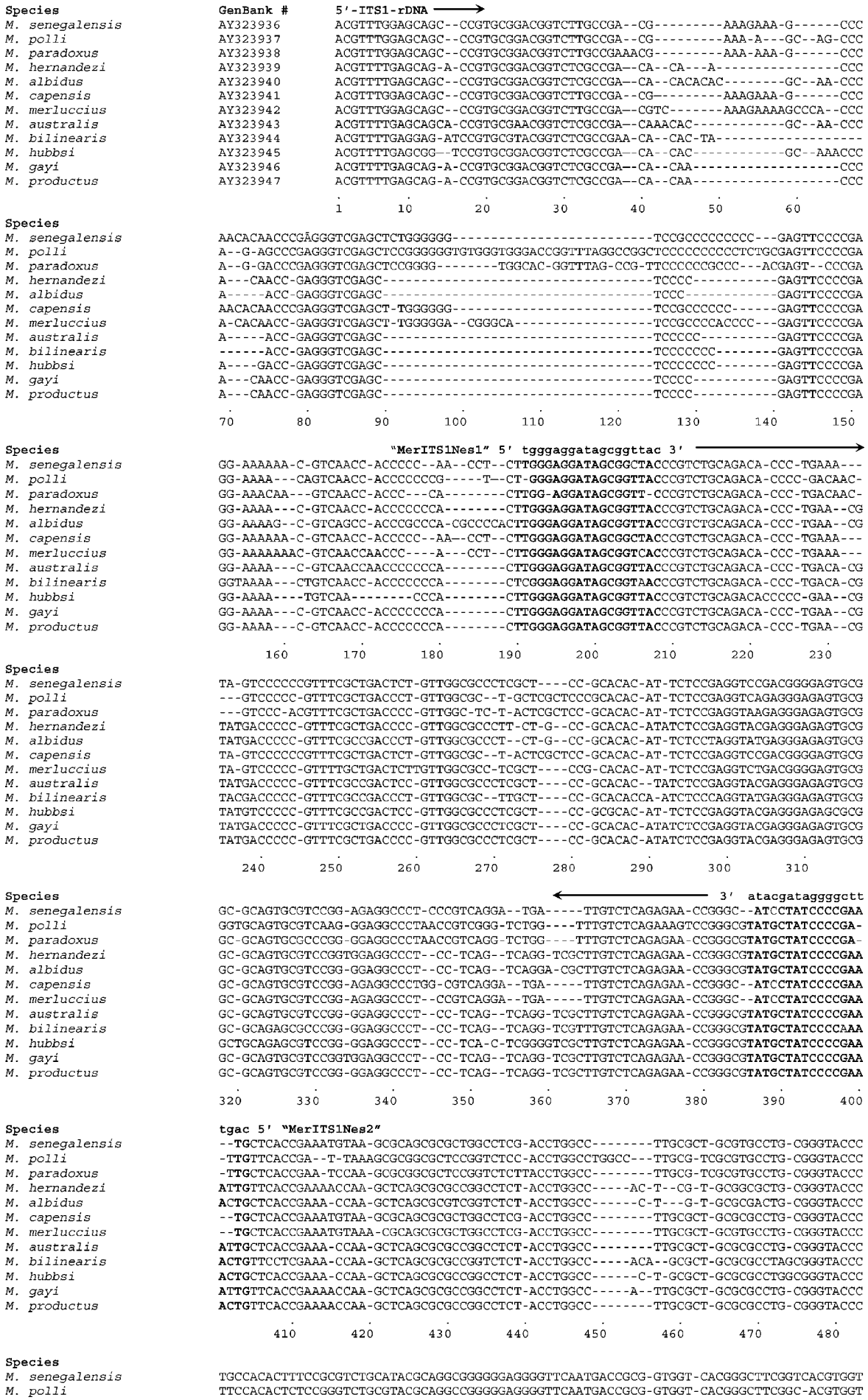
RESULTS

The double sequencing of each template as well as the use of two additional *Taq* DNA polymerases to amplify the ITS1 region, followed by its further resequencing, served at verifying the accuracy of the ITS1 sequences recovered. Identical ITS1 sequences were obtained for all individuals of the same species across independent amplifications and sequencing methods (**Figure 1**). The alignment yielded a nucleotide matrix of 692 bp that comprised the ITS1 sequence plus 53 bp from the 3′ end of the 18S gene and 20 bp from the 5′ end of the 5.8S gene. Most gaps introduced in the alignment to properly pair the ITS1 sequences from the 12 hake species were due to repetitive regions of simple nucleotide stretches (**Figure 1**).

The low number of ambiguities in the ITS1 alignment provided large conserved regions between species that were used to design three pairs of conserved primers. Two primer pairs were not specific of *Merluccius* and amplified the same length product in *Gadus morhua*, *Macruronus novaezelandiae*, *Macruronus magellanicus*, *Salmo salar*, *Salmo trutta*, and *Oncorhynchus mykiss*. Amplification was not observed in flatfish species or in mollusks; therefore, the latter were used as negative controls (data not shown). The third primer pair termed *MerITS1Nes1* and *MerITS1Nes2* (**Figure 1**) rendered a 193 bp PCR product of satisfactory quantity and quality in *Merluccius* spp. and in *G. morhua* (data not shown).

The nucleotide differences between the ITS1 sequences of the species provided the basis for selecting diagnostic enzymes from the ITS1 restriction maps of each species. The application of four restriction enzymes to aliquots of ITS1 amplicons allowed the full discrimination of all hake species and the Atlantic cod, with all composite haplotypes being species-specific (**Table 2**). The four enzymes recognized and cut the targets GT!AC/CA!TG (*Afa* I), GGCGC!C/C!CGCGG (N*ar* I), TGG!CCA/ACC!GGT (M*lu* NI), and CCTC$(N)_7$!/GGAG$(N)_6$! (M*nl* I). The restriction patterns generated after the digestion with A*fa* I allowed the identification of *M. polli*, *M. hubbsi*, *M. bilinearis* (**Table 2** and **Figure 2**), and *G. morhua* (**Table 2** and **Figure 3**). The remaining species were grouped in two clusters by the similarity of their restriction patterns, i.e., *M. merluccius*, *M. senegalensis*, *M. capensis*, *M. paradoxus*, and *M. albidus*, in one group, and *M. productus*, *M. gayi*, *M. australis*, and *M. hernandezi*, in a second group. After the identification of the first four species with A*fa* I, the enzyme N*ar* I allowed the direct discrimination of *M. merluccius* and *M. senegalensis* and also of *M. capensis*, *M. paradoxus*, *M. albidus*, and *M. australis* when combined with the pattern of the enzyme A*fa* I (**Table 2** and **Figure 2**). The enzyme M*lu* NI allowed the distinction of *M. productus* from *M. gayi* and *M. hernandezi* (**Table 2** and **Figure 2**). The enzyme M*nl* I was used to distinguish *M. gayi* from *M. hernandezi* by a 164 bp band specific to the latter species (**Table 2** and **Figure 4**). The composite pattern of the four enzymes gave a full diagnosis between the 12 species. It should be noted that the agarose gels used did not allow the clear resolution of the smallest restriction fragments (<30 bp), otherwise unnecessary for species identification, and that digestions of the ITS1 sequence in some species did not go to completion, leaving some of the original ITS1 fragment uncut (e.g., **Figure 2**).

The presence−absence matrix of restriction fragments from the digestion of the ITS1 amplicon of 10−45 individuals per


Species
GenBank #
5′-ITS1-rDNA
M. senegalensis AY323936
M. polli AY323937
M. paradoxus AY323938
M. hernandezi AY323939
M. albidus AY323940
M. capensis AY323941
M. merluccius AY323942
M. australis AY323943
M. bilinearis AY323944
M. hubbsi AY323945
M. gayi AY323946
M. productus AY323947
1 10 20 30 40 50 60
70 80 90 100 110 120 130 140 150
"MerITS1Nes1" 5′ tgggaggatagcggttac 3′
160 170 180 190 200 210 220 230
240 250 260 270 280 290 300 310
3′ atacgataggggctt
tgac 5′ "MerITS1Nes2"
410 420 430 440 450 460 470 480

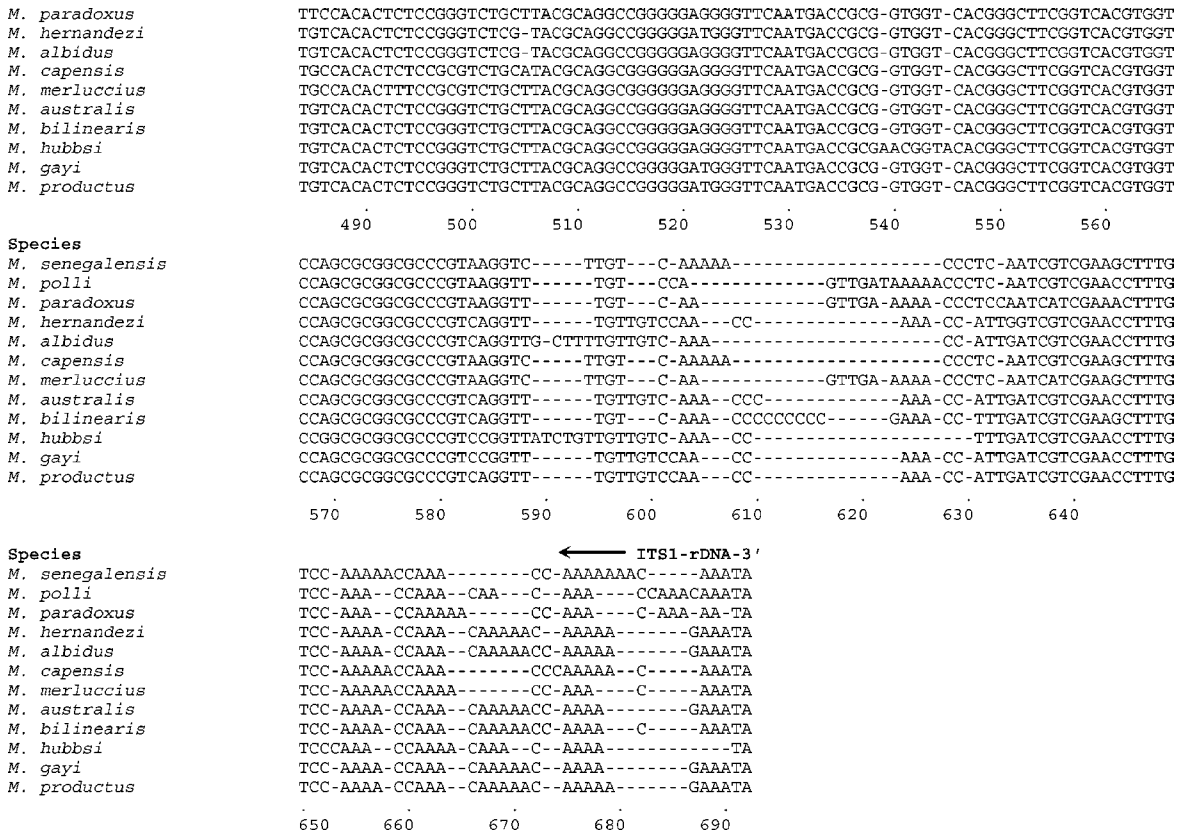


Figure 1. Alignment of ITS1-rDNA sequences from 12 hake species. The nested PCR primers *MerITS1Nes1* and *MerITS1Nes2* used to detect the presence of DNA from *Merluccius* spp. and *G. morhua* (GenBank accession number AY323948) are shown in bold above their annealing positions.

Table 2. Restriction Fragments (in bp) Obtained after Digestion of PCR–ITS1 Products from *Merluccius* spp. and *G. morhua* with the Enzymes A*fa* I, N*ar* I, M*lu* NI, and M*nl* I[a]

		A*fa* I		N*ar* I		*Mlu* NI		M*nl* I		
hake species	ITS1 length	fragment sizes	type	fragment sizes	type	fragment sizes	type	fragment sizes	type	pattern
M. merluccius	629	189, 440	A	92, 265, 272	A	629	A	5, 8, 22, 26, 36, 47, 61, 67, 74, 79, 84, 120	A	A**A**AA
M. senegalensis	618	186, 432	A	89, 255, 274	B	618	A	5, 8, 22, 26, 36, 47, 61, 67, 74, 77, 80, 115	A	A**B**AA
M. polli	659	27, 166, 466	B	97, 562	C	659	A	7, 16, 30, 31, 55, 65, 84, 91, 112, 168	B	B**C**AB
M. capensis	614	185, 429	A	88, 526	C	614	A	8, 26, 36, 38, 63, 74, 78, 85, 91, 115	A	A**CAA**
M. paradoxus	630	190, 440	A	93, 537	D	630	A	7, 27, 29, 55, 63, 83, 89, 113, 164	B	A**D**AB
M. productus	603	156, 193, 254	C	96, 226, 281	E	603	A	5, 7, 23, 25, 29, 30, 69, 86, 106, 223	C	CEAC
M. gayi	604	158, 192, 254	C	95, 226, 283	E	216, 388	B	5, 7, 24, 30, 30, 69, 106, 110, 223	C	CE**BC**
M. australis	610	155, 194, 261	C	97, 234, 279	F	610	A	5, 7, 22, 26, 29, 31, 70, 74, 85, 111, 150	D	C**F**AD
M. hubbsi	602	65, 125, 158, 254	D	96, 227, 284	E	602	A	5, 7, 22, 23, 28, 30, 70, 75, 86, 110, 146	D	**D**EAD
M. albidus	610	193, 417	A	97, 239, 274	F	610	A	5, 24, 28, 29, 38, 59, 112, 150, 155	E	AFAE
M. hernandezi	602	156, 192, 254	C	96, 226, 280	E	214, 388	B	5, 7, 30, 69, 106, 164, 221	F	CEBF
M. bilinearis	607	76, 123, 194, 214	E	97, 510	H	219, 388	B	5, 28, 29, 32, 40, 59, 64, 98, 102, 150	D	**E**HBD
G. morhua	605	72, 173, 360	F	88, 517	I	605	A	35, 45, 98, 113, 156, 158	G	**F**IAG

[a] The fragment sizes shown are exact as determined by sequencing. Bold letters indicate the restriction pattern of the species identified with that enzyme.

species with four restriction enzymes is given in **Table 3**. All conspecific samples presented identical restriction patterns. The consensus tree that summarizes the pool of 110 parsimonious trees recovered showed a full nodal resolution (100%) for each species' cluster (**Figure 5**).

The phylogenetic reconstruction using either the ITS1 amplicon or the ITS1 nested fragment of 193 bp of 10−45 individuals per species (FINS) showed an unambiguous grouping of all test samples in their expected cluster defined by samples types. The same full species assignment (data not shown) was obtained using the BLAST engine to match test samples to GenBank entries (**Figure 1**) of *Merluccius* spp.

DISCUSSION

The *a priori* morphological identification of the sample types used to calibrate any molecular identification method is indispensable for the creation of a reliable key (*33*). Therefore, all specimens were assigned to 1 of 12 well-recognized hake species. The exceptions were *M. hernandezi* and *M. albidus*, from which no entire specimens were available. Tissue samples from these two species were identified following four criteria, (i) the records of the trawling surveys during which they were captured, (ii) the latitude and longitude describing the exact oceanographic catching point, (iii) previous allozyme data

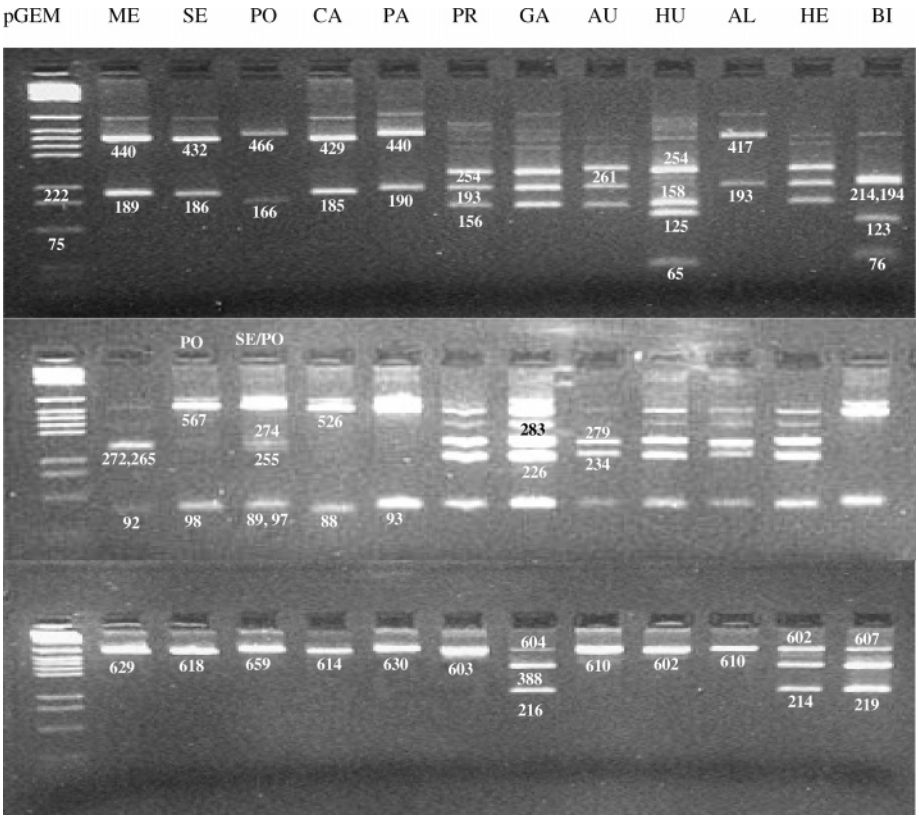


Figure 2. Agarose gel showing restriction fragments (in bp) from digestion of the ITS1–PCR product of hakes with A*fa* I (top panel), *N*ar I (middle panel; here, the third lane corresponds to *M. polli* and the fourth to a *M. polli*/*M. senegalensis* hybrid individual included only in this panel), and M*lu* NI (bottom panel). The first lane corresponds to the molecular marker pGEM (Promega) and the remaining lanes are identified using the codes of the species described in **Table 1**. The exact fragment sizes are given in **Table 2**.

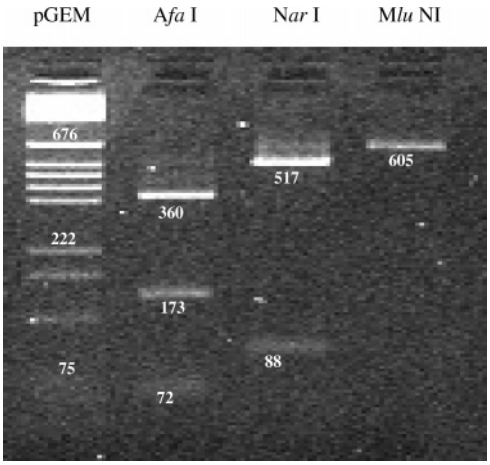


Figure 3. Agarose gel showing the restriction fragments (in bp) from digestion of the ITS1–PCR product of *G. morhua* with A*fa* I, N*ar* I, and M*lu* NI. The first lane corresponds to the molecular marker pGEM (Promega). Exact fragment sizes are given in **Table 2**.

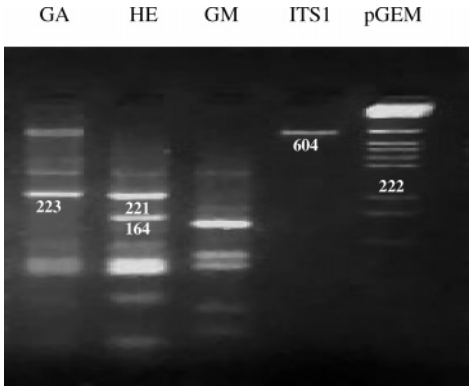


Figure 4. Agarose gel showing the 164-bp fragment that distinguishes *M. gayi* (GA) from *M. hernandezi* (HE), after digestion of their ITS1–PCR products with M*nl* I. GM, *G. morhua*; ITS1, undigested ITS1 sequence of *M. gayi*; pGEM, molecular marker (Promega).

obtained on the same individuals of *M. albidus* (M. Roldán, personal communication), and (iv) the morphology of the scales of *M. hernandezi* that serves as unambiguously differentiating this species from the neighboring hake *M. angustimanus* (C. P. Mathews, personal communication). Although the 10 species of commercial relevance are included in this study, there are two additional species from this genus which were not obtainable for this study, i.e., *M. angustimanus* or Panama hake, distributed from Baja California to Colombia (*34*), and the recently discovered *M. patagonicus* (*2*). Previous studies have attempted to differentiate 11 hake species using a PCR−RFLP protocol that used four and seven enzymes to digest the left domain of the mitochondrial DNA control region (*12*) and a cytochrome *b* gene fragment (*13*), respectively. While both methods perform well at identifying some species, they are devoid of several indispensable properties, i.e., (1) a full genetic discrimination between the nine species of hake analyzed (see Table 2 of ref *13*), (2) an *a priori* morphological identification of species and origin of the sample types analyzed, (3) an assessment of the usual intraspecific variation of mtDNA markers (*14*), which might lead to sample misidentifications, (4) a sufficient number of sample types from the distribution range of the species, and (5) a final validation step on the large number of individuals per species.

Table 3. Matrix of Presence and Absence of Restriction Fragments (in bp) Obtained with Four Enzymes on the ITS1-rDNA Amplicon of Hakes and *Gadus Morhua*

		Afa I													*Nar* I								*Mlu* NI				*Mnl* I					
species	N[a]	27	65	72–76	123–125	155–166	173	185–194	214	254–261	360	417	429–440	466	88–97	226–239	255	265	272–284	510–517	526	537	562	214–219	388	602–659	25–32	36–45	47–55	59–70	74–80	83–91	98–120	146–168	221–223
M. merluccius	45	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0
M. senegalensis	34	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	1	1	1	1	1	0	1	0	0
M. polli	33	1	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	1	0	1	1	1	1	1	0
M. capensis	45	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	1	1	0	1	1	1	1	0	0
M. paradoxus	30	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	1	0	1	1	0	1	1	1	1	1	0
M. productus	45	0	0	0	0	1	0	1	0	1	0	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	1	0	1	1	1	1	0	1
M. gayi	45	0	0	0	0	1	0	1	0	1	0	0	0	1	1	1	0	0	1	0	0	0	0	1	0	0	1	0	1	0	0	1	0	1
M. australis	45	0	0	0	0	1	0	1	0	1	0	0	0	1	1	1	0	0	1	0	0	0	0	1	0	1	1	0	1	1	1	1	1	0
M. hubbsi	45	0	1	0	1	1	0	0	0	1	0	0	0	1	1	1	0	0	1	0	0	0	0	1	0	1	1	0	1	1	1	1	1	0
M. albidus	20	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1	0	0	1	0	0	0	0	1	0	1	1	0	1	0	0	1	1	0
M. hernandezi	10	0	0	0	0	1	0	1	0	1	0	0	0	0	1	1	0	0	1	0	0	0	0	1	1	0	1	0	0	1	0	0	1	1
M. bilinearis	45	0	0	1	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	1	0	1	0	0	1	0
G. morhua	25	0	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	1	1	1	0	0	0	0	1	0

[a] *N* is the sample size used to validate the identification test (1–3 samples per species). Length fragments not straightforwardly distinguished from each other in agarose gels are grouped into the same length class (e.g., 72–76 bp).

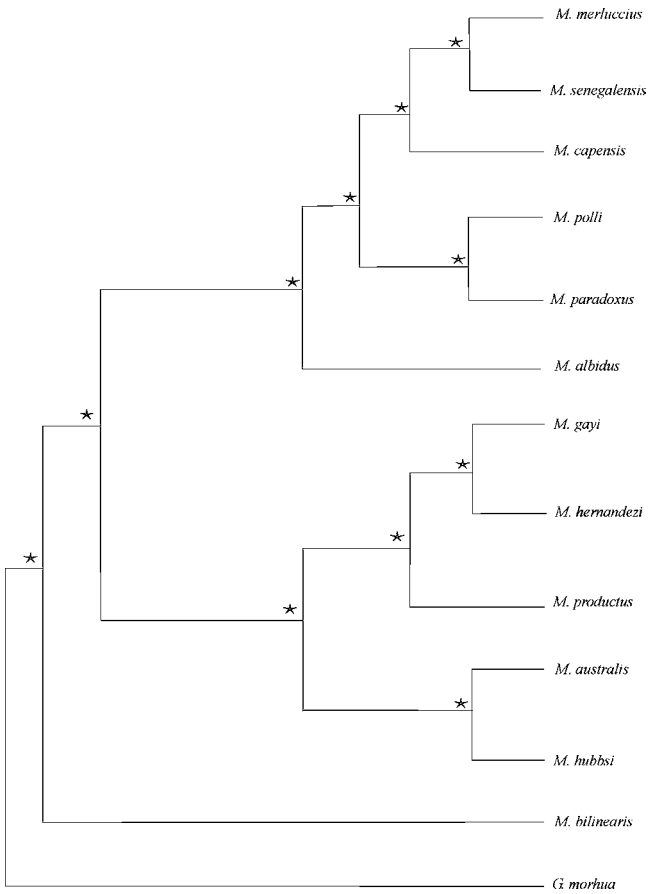


Figure 5. Consensus diagram of 110 equally parsimonious trees recovered with the polymorphism parsimony method. Because all conspecific samples showed the same restriction pattern, only one representative individual from each species was used in this reconstruction. (*) Branching topology was recovered in 100% of trees.

In this study, we have designed a preliminary test to shortcut the development of tedious and expensive identification protocols when there is no hake in the sample. It consists of PCR amplification of a ITS1-rDNA fragment using ITS1 nested primers (*MemeITS1Nes1* and *MemeITS1Nes2*) that are interspecifically conserved in hakes. This exclusion test is very robust because the amplified fragment spans only 193 bp, a DNA length usually recovered in most cases, including ancient DNA samples (*35*). While Atlantic cod was not excluded from the hake-only DNA test, no PCR amplification was observed from Gadidae species, such as grenadiers, or from distant species such as salmonids, flatfishes, and mollusks.

The second step of the identification protocol consisted of three equally performant approaches, (i) species-specific PCR−RFLPs on the whole ITS1 amplicon, (ii) nucleotide sequence comparison using BLAST against the GenBank nucleotide database, and (iii) molecular phylodiagnosis using FINS. The universal primers selected at the conserved ends of the ITS1-flanking 18S (*XelaITS1.1*) and 5.8S genes (*OnmyITS1.2*) produced a PCR fragment from each species that spanned 53 bp of the 3′ end of the 18S ribosomal gene, a 602−659 bp fragment of the specific ITS1 sequence, and 20 bp of the 5′ end of the 5.8S gene. Noteworthy, this primer pair amplifies the ITS1 regions of all eukaryotic taxa so far screened in our laboratory. The application of an array of four restriction enzymes with restriction sites tied to interspecifically nonconserved regions of the ITS1 permitted the unequivocal identification of 12 hake species, including the never-before genetically

studied *M. hernandezi* (*36*) and the Atlantic cod *G. morhua*. The restriction patterns were differentiated in agarose gels and matched those predicted with the restriction maps of the ITS1 sequences. Because a single enzyme identifies three species of hake and the Atlantic cod, two enzymes distinguish 9 species, three enzymes identify 11 species, and the four enzymes together discriminate the 12 species of hake and the Atlantic cod; it may not be necessary to apply all four enzymes to identify unknown hake samples. This PCR−RFLP method would also serve to identify hakes in seafood products where more than one species is present because of a partial substitution (M. Pérez, unpublished data). For this purpose, one should apply the three first restriction enzymes to detect the expected composite patterns between species for each enzyme.

Two important properties of the ITS1 region facilitated designing the species-specific restriction patterns. First, nucleotide divergence between species in portions of the ITS1 region allowed us to select sequence-specific restriction targets. Second, the presence of a consensus ITS1 sequence within species (*18*) allowed for reproducibility among the species-specific restriction patterns. These two properties make ITS1 one of the most valuable regions for development of nuclear DNA markers to easily identify species by PCR−RFLP (*20*), FINS (*31*), and BLAST (*32*).

The intraspecific polymorphism of the ITS1 region is very low in vertebrates (*18*) and is apparent in some restriction patterns of hakes, e.g., N*ar* I (**Figure 2**). These weak fragments are not due to cross-species contamination, as concluded from replicated experiments, and are most likely due to a low ITS1-rDNA intragenic variability that does not weaken the reproducibility of the diagnostic patterns, which have been worked out on the consensus ITS1 sequence present in all conspecific individuals.

The confidence intervals of genetic distances used to identify commercial fishes usually overlap within and between species, rendering this methodology useless for an exact identification of species. Also the bootstrap method used to calculate a "quantitative estimate of the degree of confidence" of the species' assignment (*37*) usually gives nodal values smaller than 100%, which do not provide the level of confidence required for identification purposes. In this study, we have applied a polymorphism parsimony method of restriction fragment data (*29*) to assess the diagnostic power of our methodology. The correct grouping of all individuals from the same species and the maximal nodal scores of the tree demonstrated the full identification power of this method. The phylogenetic tree recovered with the ITS1 nucleotide sequence allows for the correct assignment of unknown samples to 1 of the 12 hake species (FINS). Complementary to the FINS method, the BLAST comparison of the ITS1 nucleotide sequence (the 193 bp ITS1 nested fragment as well as the full ITS1 sequence) from an unknown hake sample also permitted its quick assignment to species provided that the ITS1 sequences of the 12 hakes have been made available at the GenBank nucleotide database.

The present identification key has interesting applications in various tasks such as the conservation and management of hake species, the study of their ecological interactions, and the commercial tracking of hake products. For instance, fishing effort in mixed fisheries of hakes could be conveniently measured for each hake by analyzing the species composition of the catches. This method can also be used to discourage fraudulent fishing (*38*) in protected areas and to detect illegal trade of hake products. From an industrial perspective, the methodology developed could be of great help at identifying commercial products of hake at different stages of the food chain. After the exact hake species in mixed and diversified products was determined, the fish industry could use this key as a marketing tool to guarantee the quality and authenticity of hake-based products for sale (*39*). From a legal perspective, there is need for analytical approaches to enforce labeling regulations and to authenticate imports and exports of hakes. In particular, fishery forensics, in cases of food alarms or fishing conflicts, can benefit from a reliable genetic diagnose.

The diagnostic method described herein is one of the most reliable tools so far developed for the identification of hakes from the genus *Merluccius*, because of its ease of execution and high reproducibility. In addition, the total time required to achieve a reliable diagnosis is approximately a working day for a large number of samples, making this methodology suitable for routine analysis. These properties have also been tested by independent researches of the seafood quality control reference laboratory of CECOPESCA (Centro Técnico Nacional de Conservación de Productos de la Pesca). Several case studies have shown that most industrial processes applied to hake meat, such as heating, cooking, and food additives, neither degraded the DNA nor inhibited the PCR reaction to such an extent as to prevent the amplification of the two diagnostic fragments of the ITS1-rDNA (*20*, *40*).

ABBREVIATIONS USED

EDTA, ethylenediaminetetraacetic acid; bp, base pair(s); FINS, forensically informative nucleotide sequencing; GPS, global positioning system; ITS1, internal transcribed spacer 1; rDNA, ribosomal DNA; SDS, sodium dodecyl sulfate; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TBE, Tris-borate-EDTA.

ACKNOWLEDGMENT

The authors are indebted to all of the fishing organizations that assisted in collection of hakes in their native habitats. Particularly, we acknowledge the sampling effort developed by C. Ruíz, J. Irisarri, B. Carrillo, I. Crespán, M. José, J. Ellis, L. Sanpedro, F. Sánchez, F. Velasco, B. Patiño, M. Sainza, F. Badalamenti, R. Gustafson, R. K. Smedbol, P. Chase, M. Saunders, M. Showell, E. Rubidge, P. Tamkee, G. de Joya, M. Salgado, O. Pedrín, J. Guardado, and M. I. Roldán. We especially thank the expert taxonomic and sampling assistance granted by C. P. Mathews. The identification methodology based on rDNA-RFLPs described in this paper is a property of Vigo University, protected under the Spanish patent P200300996; use of this method is permitted for research purposes only.

LITERATURE CITED

(1) Inada, T. Studies on merluciid fish. *Bull. Far Seas Fish. Res. Lab. Shimizu, Japan.* **1981**, *18*, 1−172.

(2) Lloris, D.; Patallanas, J. Description of a new species of hake: *Merluccius patagonicus* sp. Nov. (Gadiformes: Merluciidae) from the waters of Argentina. *Sci. Mar.* **2003**, *67*, 323−326

(3) Jones, B. W. World resources of hakes of the genus *Merluccius*. In *Sea Fisheries Research*; Harden-Jones, F. R., Ed.; John Wiley and Sons: New York, 1974; pp 139−166.

(4) Mackie, I. M. Fish species identification—A retrospective view. In *Authenticity of Species in Meat and Seafood Products*; Pérez-Martín, R. I., Sotelo, C. G., Eds.; CSIC−Chapela Press: Vigo, Spain, 2003; pp 9−26.

(5) Grant, W. S.; Leslie, R. W. Inter-ocean dispersal is an important mechanism in the zoogeography of hakes (Pisces: *Merluccius* spp.). *J. Biogeogr.* **2001**, *28*, 699−721.

(6) Jones, B. W.; Mackie, I. M. An application of electrophoretic analyses of muscle myogenes to taxonomic studies in the genus *Merluccius*. *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol.* **1970**, *32*, 267−273.
(7) Mackie, I. M.; Jones, B. W. The use of electrophoresis of the water-soluble (sarcoplasmic) proteins of fish muscle to differentiate the closely related species of hake (*Merluccius* spp.). *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol.* **1978**, *59*, 95−98.
(8) Etienne, M.; Jérôme, M.; Fleurence, J.; Rehbein, H.; Kündiger, R.; Mendes, R.; Costa, H.; Pérez-Martín, R.; Piñeiro-González, C.; Craig, A.; Mackie, I.; Malmheden, I.; Ferm, M.; Martínez, I.; Jessen, F.; Smelt, A.; Luten, J. Identification of fish species after cooking by SDS−PAGE and urea IEF: A collaborative study. *J. Agr. Food Chem.* **2000**, *48*, 2653−2658.
(9) Piñeiro, C.; Vázquez, J.; Marina, A. I.; Barros-Velázquez, J.; Gallardo, J. M. Characterization and partial sequencing of species-specific polypeptides from commercial hake species by mass spectrometry following two-dimensional electrophoresis. *Electrophoresis* **2001**, *22*, 1545−1552.
(10) Becker, I. I.; Grant, W. S.; Kirby, R.; Robb, F. T. Evolutionary divergence between sympatric species of Southern African hakes, *Merluccius capensis* and *M. paradoxus* II: Restriction enzyme analysis of mitochondrial DNA. *Heredity* **1988**, *61*, 21−30.
(11) Baker, C. S.; Palumbi, S. R. Population structure, molecular systematics, and forensic identification of whales and dolphins. In *Conservation Genetics: Case Histories from Nature*; Avise, J. C., Hamrick, J. L., Eds., Chapman and Hall: New York, 1996; pp 10−49.
(12) Quinteiro, J.; Vidal, R.; Izquierdo, M.; Sotelo, C. G.; Chapela, M. J.; Pérez-Martín, R. I.; Rehbein, H.; Hold, G. L.; Russel, V. J.; Pryde, S. E.; Rosa, C.; Samtos, A. T.; Rey-Méndez, M. Identification of hake species (*Merluccius* genus) using sequencing and PCR−RFLP analysis of mitochondrial DNA control region sequences. *J. Agr. Food Chem.* **2001**, *49*, 5108−5114.
(13) Hold, G. L.; Russell, V. J.; Pryde, S. E.; Rehbein, H.; Quinteiro, J.; Vidal, R.; Rey-Méndez, M.; Sotelo, C. G.; Pérez-Martín, R. I.; Samtos, A. T.; Rosa, C. Development of a DNA-based method aimed at identifying the fish species present in food products. *J. Agr. Food Chem.* **2001**, *49*, 1175−1179.
(14) Brown, J. R.; Beckenbach, A. R.; Smith, M. J. Intraspecific DNA sequence variation of the mitochondrial control region of white sturgeon (*Acipenser transmontanus*). *Mol. Biol. Evol.* **1993**, *10*, 326−341.
(15) Long, E. O.; Dawid, I. B. Repeated genes in eukaryotes. *Annu. Rev. Biochem.* **1980**, *49*, 727−764.
(16) Cortadas, J.; Pavon, M. C. The organization of ribosomal genes in vertebrates. *EMBO J.* **1982**, *1*, 1075−1080.
(17) Castro, J.; Pardo, B. G.; Sánchez, L.; Martínez, P. rDNA RFLPs as genetic markers for resource management in brown trout. *J. Fish Biol.* **1999**, *55*, 221−225.
(18) Presa, P.; Pardo, B.; Martínez, P.; Bernatchez, L. Phylogeography of brown trout (*Salmo trutta*) using the ribosomal internal transcribed spacers (ITS). *Mol. Biol. Evol.* **2002**, *19*, 2161−2175.
(19) van Oppen, M. J. H.; Willis, B. L.; van Vught, H. W. J. A.; Miller, D. J. Examination of species boundaries in the *Acropora cervicornis* (Sclaractinia, Cnidaria) using nuclear DNA sequence analyses. *Mol. Ecol.* **2000**, *9*, 1363−1374.
(20) Pérez, M.; Cabado, A. G.; Vieites, J. M.; Presa P. Experimental assessment of a new rDNA-based method for the identification of *Merluccius capensis* and *Merluccius paradoxus* in commercial products. *J. Aquatic Food Prod. Technol.* **2004**, *13*, 49−57.
(21) Miller, S. A.; Dykes, D. D.; Polesky, H. F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* **1988**, *16*, 1215.
(22) Sambrook, J.; Fritsch, E. F.; Maniatis, T. Appendix E: Commonly used techniques in molecular cloning. In *Molecular Cloning: A Laboratory Manual*, 2nd ed.; Ford, N, Nolan, C., Ferguson, M., Eds.; Cold Spring Harbor Laboratory Press: New York, 1989; pp E3−E4.
(23) Salim, M.; Maden, B. E. Nucleotide sequence of *Xenopus laevis* 18S ribosomal RNA inferred from gene sequence. *Nature* **1981**, *291*, 205−208.
(24) Nazar, R. N.; Roy, K. L. Nucleotide sequence of rainbow trout (*Salmo gairdneri*) ribosomal 5.8S ribonucleic acid. *J. Biol. Chem.* **1978**, *253*, 395−399.
(25) Edelman, I.; Olsen, S.; Devereux, J. *Program Manual for the Wisconsin Package*, version 1.0; Genetics Computer Group (GCG): Madison, WI, 1996.
(26) Rychlick, W.; Rhoads, R. E. A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing, and *in vitro* amplification of DNA. *Nucleic Acids Res.* **1989**, *17*, 8543−8551.
(27) Heiman, M. *Webcutter 2.0*.; http://www.firstmarket/cutter/cut2.html; 1997.
(28) Roberts, R.; Macelis, D. REBASE—Restriction enzymes and methylases. *Nucleic Acids Res.* **2001**, *29*, 268−269.
(29) Felsenstein, J. *PHYLIP (Phylogeny Inference Package)*, version 3.5; Department of Genetics, University of Washington, Seattle, WA, 1993.
(30) Felsenstein, J. Alternative methods of phylogenetic inference and their interrelationship. *Syst. Zool.* **1979**, *28*, 49−62.
(31) Barlett, 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.
(32) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389−3402.
(33) Maddison, W. P. Molecular approaches and the growth of phylogenetic biology. In *Molecular Zoology, Advances, Strategies, and Protocols*; Ferraris, J. D., Palumbi, S. R., Eds.; Wiley-Liss: New York, 1996; pp 47−63.
(34) Garman, S. Results of an exploration of the west coasts of Mexico, Central and South America, and off the Galapagos Island. *Mem. Mus. Comput. Zool., Harvard College* **1899**, *26*, 183−185.
(35) Pääbo, S. Ancient DNA. *Sci. Am.* **1993**, *269*, 86−92.
(36) Mathews, C. P. Meristic studies of the Gulf of California species of *Merluccius*, with a description of a new species. *J. Nat. Hist.* **1985**, *19*, 697−718.
(37) Terol, J.; Mascarell, R.; Fernández-Pedrosa, V.; Pérez-Alonso, M. Statistical validation of the identification of tuna species: Bootstrap analysis of mitochondrial DNA sequences. *J. Agr. Food Chem.* **2002**, *50*, 963−969.
(38) Baker, C. S.; Palumbi, S. R. Which whales are hunted? A molecular genetic approach to monitoring whaling. *Science* **1994**, *265*, 1538−1539.
(39) Vázquez, X. The effect of overfishing on species substitution in seafood products: A general approach. In *Authenticity of Species in Meat and Seafood Products*; Pérez-Martín, R. I., Sotelo, C. G., Eds.; CSIC−Chapela Press: Vigo, Spain, 2003; pp 27−38.
(40) Pérez, M. Calibrado y aplicación de marcadores moleculares para el diagnóstico genético de origen y autenticidad de stocks, productos y subproductos de merluza, *Merluccius* spp. Ph.D. thesis, University of Vigo, Spain, 2003, http://biblio.cesga.es/search*.spi.

Received for review November 26, 2004. Revised manuscript received May 2, 2005. Accepted May 6, 2005. This research was partially supported with grants from Xunta de Galicia (PGIDT00MAR02E) to J.V. and P.P. and from the Spanish Ministerio de Ciencia y Tecnología to P.P. (INIA RZ00/020) and M.P. (MIT99/236122076).

JF048012H