

Validation of a tRNA-*Glu*-cytochrome *b* Key for the Molecular Identification of 12 Hake Species (*Merluccius* spp.) and Atlantic Cod (*Gadus morhua*) Using PCR-RFLPs, FINS, and BLAST

MONTSE PÉREZ* AND PABLO PRESA

Department of Biochemistry, Genetics and Immunology, Faculty of Marine Sciences—ECIMAT, University of Vigo, 36310, Vigo, Spain

The goal of this study was to develop a diagnostic key for hake meat to solve the limitations of previous identification methodologies, mainly related to the high degradation of the DNA recovered from processed foods. We describe the development of two molecular tools based on polymerase chain reaction (PCR) and PCR-restriction fragment length polymorphisms of the cytochrome *b* gene, respectively, to identify DNA from 12 hake species in commercial products. The first assay is an exclusion test consisting of the PCR amplification of a 122 bp fragment using nested primers interspecifically conserved in *Merluccius* spp. and in *Gadus morhua*. This 122 bp amplicon, being the shortest one so far designed for hake DNA, is a useful traceability tool for highly degraded samples because its sequence contains enough interspecific diagnostic variation to identify 10 hake species and cod and has been successfully amplified from most commercial products so far tested. The second identification key follows a positive outcome of the exclusion test and consists of the PCR amplification of a 464–465 bp fragment and its digestion with three restriction enzymes whose targets map at interspecifically nonconserved sites of the cytochrome *b*. The key presented here has passed through a rigorous methodological calibration including its testing for genus specificity, its validation on a large number of authenticated sample types from each species range, and its implementation with a maximum likelihood method for the assignment of unknown samples. Together, these two procedures constitute the most complete molecular key so far developed for *Merluccius* spp., which is optimal for routine identification of hakes in large commercial samples at a reasonable cost–time ratio.

KEYWORDS: *Merluccius* spp.; hakes; molecular traceability; identification key; PCR-RFLPs; cytochrome *b*; FINS; BLAST

INTRODUCTION

The genus *Merluccius* comprises at least 14 species of hakes, which occupy most temperate and tropical continental shelves except the Asian shores of the Pacific Ocean (1–3). Most hakes occupy large areas of bathymetric overlap (1), resulting in the simultaneous catching of two species in the same fishery. From a management perspective, such multispecies catches hamper the independent assessment of each species fishery in terms of biomass prospective, genetic structuring, seasonal distributions, or cyclic migration routes followed by hake stocks. Also, the industrial and administrative sectors face difficulties in regulating the commerce of products from mixed fisheries (4). For instance, the marked differences in price and marketability between species, together with the overfishing of the most popular hakes,

increase the opportunities for their fraudulent commercial substitution either by other hakes or by similar taxa (5).

The identification of hakes is affordable using morphological keys on whole specimens such as typified measurements of body parts (1, 2). However, those morphological keys are neither fully reliable nor always available from processed commercial products. Several biochemical tools have been applied to classify hakes, such as allozyme electrophoresis (4). However, this technology does not always provide a clear-cut assignment of individuals to species due to the partial diagnosis affordable from the generally low-variable and highly conserved protein-coding loci. Also, several techniques based in species-specific sarcoplasmic proteins, such as isoelectrofocusing (IEF), two-dimensional electrophoresis (2-DE), and sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), have been optimized to improve the commercial identification of some hakes (6–8). Nevertheless, those protein-based identification methods are dependent on the proteomic expression of a

* To whom correspondence should be addressed. Fax: +34 986 81 25 67. E-mail: mon@uvigo.es.

Table 1. Name and Latitudinal Range of the 12 Hakes Identified and the Eight Outgroup Taxa Used as External Controls^a

hake species	codes	common name	range sampled	Ns	Ni
<i>Merluccius merluccius</i>	ME	European hake	37° N–55° N	3	45
<i>Merluccius senegalensis</i>	SE	Senegalese hake	15° N–21° N	3	45
<i>Merluccius polli</i>	PO	Benguela hake	15° N–27° N	3	45
<i>Merluccius capensis</i>	CA	shallow-water cape hake	17° S–25° S	3	45
<i>Merluccius paradoxus</i>	PA	deep-water cape hake	25° S–34° S	3	45
<i>Merluccius productus</i>	PR	Pacific hake	48° N–50° N	3	45
<i>Merluccius gayi</i>	GA	Peruvian hake	08° S–30° S	3	45
<i>Merluccius australis</i>	AU	Chilean hake Antarctic queen hake New Zealand hake Austral hake	41° S–52° S	3	45
<i>Merluccius hubbsi</i>	HU	Patagonian hake	46° S–53° S	3	45
<i>Merluccius albidus</i>	AL	offshore hake	35° N–37° N	1	11
<i>Merluccius angustimanus</i>	HE	Panama hake	29° N	1	10
<i>Merluccius bilinearis</i>	BI	silver hake	39° N–42° N	3	30
<i>Gadus morhua</i>	GM	Atlantic cod	commercial origin	1	25
<i>Macruronus magellanicus</i>	MM	Patagonian grenadier	commercial origin	1	10
<i>Macruronus novaezelandiae</i>	MN	blue grenadier	commercial origin	1	10
<i>Salmo salar</i>	SS	salmon	commercial origin	1	10
<i>Salmo trutta</i>	ST	sea trout	commercial origin	1	10
<i>Oncorhynchus mykiss</i>	OM	rainbow trout	commercial origin	1	10
<i>Scophthalmus maximus</i>	SM	turbot	commercial origin	1	10
<i>Scophthalmus rhombus</i>	SR	brill	commercial origin	1	10
<i>Platichthys flexus</i>	PF	flounder	commercial origin	1	10
<i>Octopus vulgaris</i>	OV	common octopus	commercial origin	1	10
<i>Mytilus galloprovincialis</i>	MG	Mediterranean mussel	commercial origin	1	10

^aNs, number of sample sites per species range; Ni, total number of individuals tested per species.

given tissue, on the freshness of the tested sample, or on biochemical particularities of each species. Therefore, these methods offer a case-by-case identification performance and are especially useful as complementary tests in species diagnosis. The restriction fragment length polymorphisms (RFLPs) on mitochondrial DNA (mtDNA) amplicons developed to identify some hakes came to circumvent the main disadvantages of protein analyses (9, 10). However, new marker-dependent constraints have been unveiled. For instance, the large intraspecific variation of the left domain of the mtDNA control region (9), the lack of species-specific restriction patterns of cytochrome *b* for several pairs of closely related hakes (10), or the lack of short DNA probes to be detected in processed food constitute serious handicaps for a reliable assignation of individuals to species. The use of nuclear DNA in species identification is not as common as the mtDNA. Only Pérez et al. (11, 12) have reliably identified all hakes so far described using an rDNA-ITS1 key. Although that previous method is capable of identifying 12 hake species when a high enough amount of DNA is available from foods, the degradation status of DNA in some subproducts (e.g., pellets, surimi, pâté, soups, etc.) makes advisable the availability of a complementary test, either to reinforce or to achieve a full commercial diagnosis.

Because the nucleotide variation of cytochrome *b* is less saturated by multiple substitutions than other mitochondrial regions (13), this marker is one of the most useful DNA segments for species authentication and identification of closely related taxa, either congeneric (14) or confamilial (e.g., see ref 15). The goal of this study was to validate a diagnostic key for hake meat to solve the limitations of previous identification methodologies, mainly related to the high degradation of the DNA recovered from processed foods. For that purpose, the experimental design was focused on two sequential PCR-based keys. The first one pursued the detection of hake DNA in a sample, and the second one pursued to assign individuals to one of the 12 hake species so far recognized morphologically (2).

MATERIALS AND METHODS

Sampling and Morphological Identification of *Merluccius* Spp.

To maximize the amount of intraspecific variation detected within the cytochrome *b* gene, the 12 species of hake considered were sampled at three distant sites of their oceanographic ranges (Table 1). This worldwide sampling was performed in cooperation with commercial and research vessels that fish in Euro-African (East-Atlantic) and American (West-Atlantic and Pacific) fisheries. About 20 specimens per sample were frozen or immersed in absolute ethanol upon collection, and their GPS (global positioning system) coordinates were recorded on board. All specimens were identified using species-specific morphological traits following classification criteria previously established (1, 2). The cytochrome *b* of one specimen per sampling site (1–3 specimens per species) was sequenced for a priori screening of SNPs (single nucleotide polymorphisms) useful to calibrate the identification method. The remaining specimens (around 15 per site and 45 per species) were used to validate the identification method. Two specimens of *Gadus morhua* (Atlantic cod) were used as an outgroup taxon in the phylodiagnostic analysis.

Amplification and Sequencing of Cytochrome *b*. The methodology used for DNA extraction and purification consisted of a combination of the salting-out method (16) and the standard phenol:chloroform method (17). The 5'-end of cytochrome *b* gene was polymerase chain reaction (PCR)-amplified from total DNA using the primer pair L14735 and H15149AD described by Wolf et al. (18). The PCR reaction yielded a 464–465 bp fragment spanning from the 3'-end of the tRNA-*Glu* (36–37 nucleotides) to 428 nucleotides from 5'-end of the cytochrome *b*. Amplifications were carried out in a Mastercycler thermocycler (Eppendorf) as follows: 95 °C for 3 min and 35 cycles of 95 °C for 30 s, 63 °C for 40 s and 72 °C for 1 min, followed by a final extension step at 72 °C for 10 min. The PCR amplification mixture of 50 µL contained 50 ng of DNA template, 20 pmol of each primer, 0.3 mM dNTPs, 3 mM MgCl₂, 2 U of *Taq* DNA polymerase (Promega), and 5 µL of 10× reaction buffer. PCR products were cleaned before the sequencing reaction using NucleoSpin Extract II (Macherey-Nagel) according to the protocol of the manufacturer. The PCR amplicons of cytochrome *b* were directly sequenced on both DNA strands with the same primer pair used for PCR amplification. Sequencing was performed in an ABI Prism 3100 capilar sequencer using the BigDye Terminator Cycle Sequencing Standard (Applied Biosystems). The sequencing reaction consisted of a denaturing cycle of 95 °C for 2

min, followed by 25 cycles of 96 °C for 20 s, 55 °C for 15 s, and 60 °C for 4 min and 30 s. The final cytochrome *b* haplotypes were derived from the alignment of the forward and reverse sequences obtained for each individual, using CHROMAS software (<http://www.technelysium.com.au/chromas.html>).

Selection of Specific Cytochrome *b* Primers for the Genus *Merluccius*. The 3'-end of the tRNA-Glu gene was used to align the sequences obtained, using the SeqLab program implemented in the GCG software (19). The alignment yielded a nucleotide matrix of 465 bp and 32 cytochrome *b* sequences (three sequences from each of 10 species and one sequence from each *M. albidus* and *M. angustimanus*) suitable for the identification of conserved regions between species. This matrix was used to select a nested primer pair, namely, MerCytbNes1 and MerCytbNes2 (14), within the cytochrome *b* using the program Oligo 4.05 (20). These primers were used to generate a PCR amplicon of 122 bp, whose reaction consisted of 95 °C for 5 min, and 30 cycles at 96 °C for 1 min, 59 °C for 1 min, and 72 °C for 1 min, followed by a final extension step at 72 °C for 10 min. To check for the specificity of these primers in *Merluccius* spp., they were tested in related taxa such as tailed hakes (*M. novaezelandiae* and *M. magellanicus*) and Atlantic cod (*G. morhua*), as well as in distant taxa such as salmonids (*S. salar*, *S. trutta*, and *O. mykiss*), flatfishes (*S. maximus*, *S. rhombus*, and *P. flexus*), and molluscs (*O. vulgaris* and *M. galloprovincialis*). All of these outgroup taxa were authenticated using entire specimens and the Fishbase taxonomic web site.

Selection of Restriction Targets and Establishment of Species-Specific Restriction Patterns. Several restriction maps were elaborated on the 464–465 bp amplicons spanning from the 5'-end of cytochrome *b* for each species, using the program Webcutter 2.0 (21) that works with the enzymatic database Rebase (22) from New England Biolabs. Such restriction maps were cross-compared to select three restriction enzymes providing the maximum discriminating power between species using the interspecific SNPs detected along the cytochrome *b* sequence. To check for the consistency of the patterns predicted from the restriction maps, the cytochrome *b* of 15 authenticated individuals from each sampling location (45 individuals per species) was PCR amplified, and the amplicons were cleaned (NucleoSpin Extract II, Macherey-Nagel) and digested with three targeted restriction enzymes. Because the nested primer pair used to detect the presence of DNA from *Merluccius* spp. also amplified a homologous fragment in *G. morhua*, the cytochrome *b* amplicons of 25 individuals of this species were also enzymatically digested. All digestions were performed independently for each enzyme at 37 °C in 20 µL containing 3 U of the restriction enzyme, 0.5–1 µg of the cytochrome *b* amplicon, and 2 µL of reaction buffer. Incubations proceeded for 3 h, and the products were electrophoresed in 3% agarose gels (2 × NuSieve:1 × Seakem LE) in TBE 1 × at 70 V for 1 h. The restriction pattern characteristic of each species was established from gels upon (i) their comparison with a molecular weight marker, (ii) side by side comparisons of restriction patterns between pairs of species, and (iii) verification of the exact size of fragments as determined from the expected restriction maps of each species.

Species Identification Using Phylogenetically Informative Nucleotide Sequencing (FINS). The application of FINS (23) to identify hake species consisted of building a joint phylogeny with cytochrome *b* sequences from the authenticated sample types and cytochrome *b* sequences from different sources, using MEGA 4.0 (24). The assigned test samples consisted of either cytochrome *b* sequences of hakes randomly chosen from GenBank (i.e., EF362892, EF362889, AB248669, DQ174064, and AY946306) or hake specimens previously identified with the PCR-RFLPs key developed herein (i.e., sample 1 as *M. polli*, sample 2 as *M. senegalensis*, sample 3 as *M. merluccius*, sample 4 as *M. australis*, and sample 5 as *M. bilinearis*). All positions containing gaps and missing data were removed (complete deletion option), resulting in a final data set of 267 nucleotide sites. The reconstruction algorithm enforced was the neighbor-joining (25), and the consensus tree was inferred from 10000 bootstrap replicates (26). The phylogenetic distance between species, measured as the number of base substitutions per site, was computed using the maximum composite likelihood method (27). This analysis allowed us to assess the accuracy of the identification method at assigning samples to each species' cluster.

Species Identification Using Basic Local Alignment Search Tool (BLAST). The correct assignment of individuals to species was also tested through the calculation of the expected value of random sequence identity using the BLAST package (28). For this calculation, the tRNA-Glu-cytochrome *b* fragment of all specimens was compared to the cytochrome *b* sequence of each hake species made available in GenBank after this study (accession numbers are in Figure 1).

RESULTS AND DISCUSSION

Morphological Identification of Sample Types. Knowledge on the origin of sample types as well as their morphological assignment to a given species is indispensable prior to calibrate a method for molecular identification (29). Using a morphological key developed for hakes (1, 2), all specimens were assigned to one of the 12 well-described hake species. The exceptions were *M. angustimanus* and *M. albidus*, from which no entire specimens were available. Tissue samples from these two species were identified following two criteria: (i) the records of the trawling surveys carried out to capture them, such as the latitude and longitude coordinates falling within their distribution ranges, and (ii) previous allozyme data obtained on the same individuals for *M. albidus* (M. Roldán, personal communication). Additionally to those two exotic hakes, two more species of *Merluccius* were not available for this study, that is, the recently reported *M. patagonicus* (2) and *M. tasmanicus* (3), whose commercial relevance remains uncertain.

Preliminary Test of Presence/Absence. Identical sequences were obtained from 1–3 conspecific individuals across independent amplifications (Figure 1), and no ambiguities were present in the alignment of cytochrome *b* sequences from all of the species considered (GenBank AY323936–AY323948). The primer pair MerCytbNes1 and MerCytbNes2 (Figure 1) rendered a 122 bp PCR product of satisfactory quantity and quality in *Merluccius* spp. and in *G. morhua* (data not shown) but failed to amplify in 10 specimens from each of the eight taxa used as external taxonomic groups. On one hand, this exclusion test being positive unambiguously indicates that DNA is present in the sample either from hake or from Atlantic cod. Therefore, one can proceed to the second test for the assignment of the specimen to one of the 12 hakes or to the Atlantic cod.

On the other hand, the exclusion test being negative does not exclude the presence of meat from hake or cod in a sample, because DNA could be highly degraded (down from 122 bp) due to the mechanical, thermal, or chemical treatments applied to commercial products. However, a negative outcome of this test indicates that it would be useless to try any current molecular test requiring the amplification of a DNA amplicon larger than 122 bp. This preliminary exclusion test is very robust since the amplified fragment spans only 122 bp, a DNA amplicon shorter than those reported in previous methodologies (9, 12) and is usually amplified from most commercial samples. Moreover, in products where DNA is highly degraded due to commercial treatments, and consequently, the 464 bp amplicon might not be amplified for RFLP analyses, it would still be feasible to unambiguously identify 10 hake species (except *M. senegalensis* from *M. capensis*, both showing an identical sequence for 122 bp fragment) and cod by sequencing the short 122 bp fragment, provided it has been successfully amplified from most commercial products so far tested (ref 14 and Figure 4) and contains enough interspecific diagnostic variation (Figure 1).

Identification Key Using Hake-Specific Restriction Patterns. The SNPs detected in the tRNA-Glu-cytochrome *b* sequence between species were the basis for selecting diagnostic enzymes from the restriction map of each species. The three diagnostic enzymes selected recognized and cut the targets

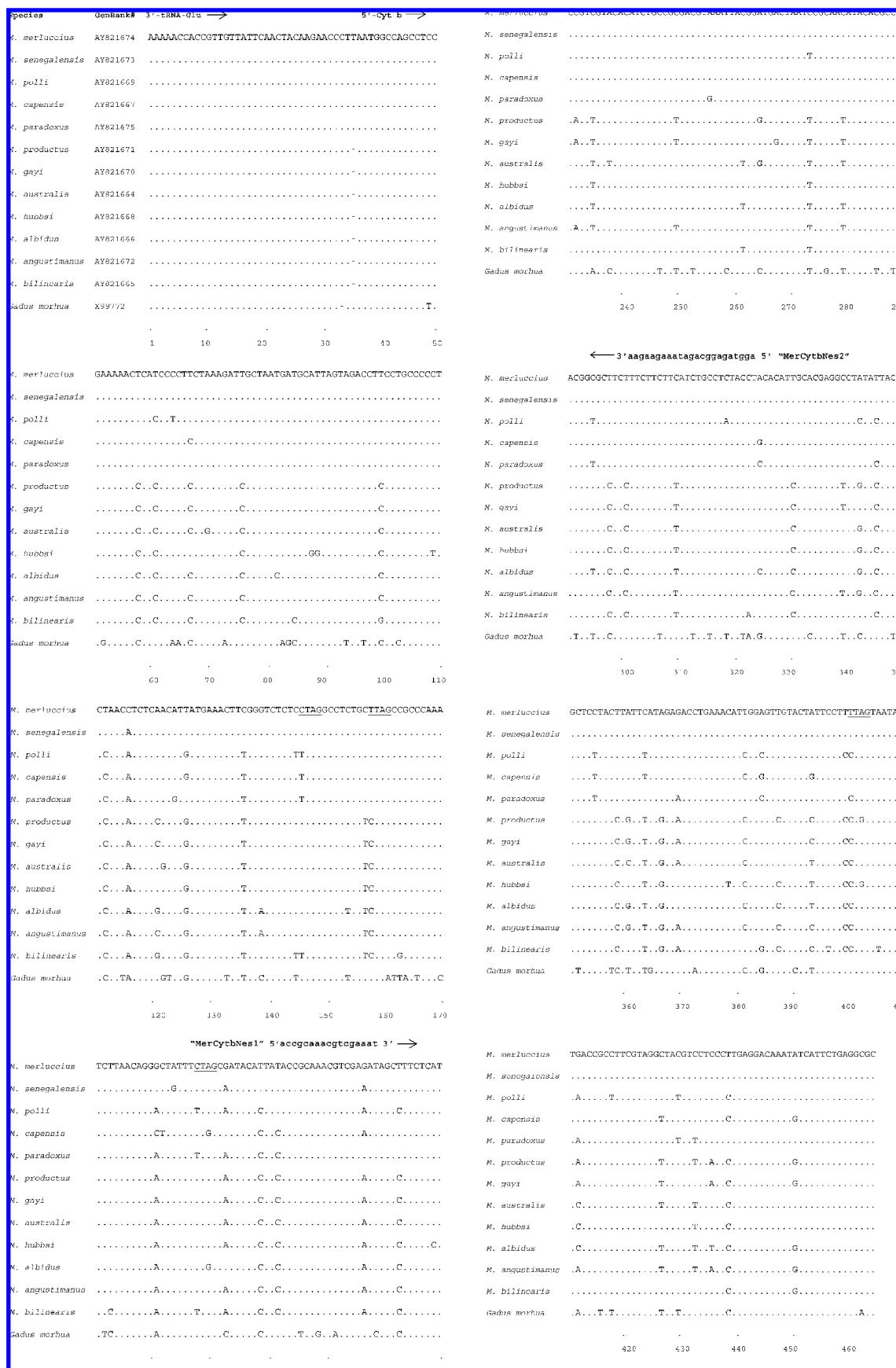


Figure 1. Alignment of tRNA-Glu-cytochrome *b* sequences from 12 hake species and *G. morhua*. The nested PCR primers MerCytbNes1 and MerCytbNes2 amplifying a 122 bp fragment used to detect the presence of DNA from *Merluccius* spp. and *G. morhua* or to identify 10 hakes and cod by sequencing (except *M. senegalensis* and *M. capensis*, which are molecularly identical to each other for this DNA fragment) are indicated in bold on the coding strand. The targets of the restriction enzyme *Bfa* I comprising useful SNPs for species identification are underlined in the reference species.

Table 2. Expected Restriction Patterns of *Merluccius* spp. and *G. morhua*^a

hake species	code	amplicon length (bp)	Mnl I			Bfa I		Afa I		
			fragment size (bp)			type	fragment size (bp)	type	fragment size (bp)	type
<i>M. merluccius</i>	ME	465	3, 7, 9, 9, 13, 36, 56, 62, 105, 165	A	42, 146, 277	A	75, 153, 237	A	AAA	
<i>M. senegalensis</i>	SE	465	3, 9, 9, 13, 43, 56, 62, 105, 165	B	42, 146, 277	A	75, 153, 237	A	BAA	
<i>M. polli</i>	PO	465	9, 9, 13, 16, 21, 27, 43, 46, 56, 57, 168	C	64, 401	B	70, 75, 83, 237	B	CBB	
<i>M. capensis</i>	CA	465	3, 9, 9, 13, 18, 43, 44, 56, 105, 165	D	465	C	75, 153, 237	A	DCA	
<i>M. paradoxus</i>	PA	465	3, 9, 9, 13, 27, 43, 56, 62, 78, 165	E	64, 401	B	75, 153, 237	A	EBA	
<i>M. productus</i>	PR	464	3, 11, 13, 18, 18, 19, 27, 33, 43, 55, 78, 146	F	12, 30, 145, 277	D	46, 75, 107, 236	C	FDC	
<i>M. gayi</i>	GA	464	3, 11, 13, 18, 18, 19, 27, 33, 43, 55, 78, 146	F	12, 30, 64, 145, 213	E	75, 153, 236	A	FEA	
<i>M. australis</i>	AU	464	3, 11, 13, 18, 18, 19, 27, 33, 39, 39, 43, 55, 146	G	12, 30, 64, 145, 213	E	46, 75, 343	D	GED	
<i>M. hubbsi</i>	HU	464	3, 13, 18, 19, 27, 51, 54, 55, 69, 77, 78	H	12, 30, 145, 277	D	46, 75, 107, 236	C	HDC	
<i>M. albidus</i>	AL	464	3, 11, 13, 18, 18, 19, 27, 33, 55, 78, 189	I	12, 64, 145, 243	F	46, 75, 107, 236	C	IFC	
<i>M. angustimanus</i>	AN	464	3, 11, 13, 18, 18, 19, 27, 33, 43, 55, 78, 146	F	12, 30, 64, 145, 213	E	46, 75, 107, 236	C	FEC	
<i>M. bilinearis</i>	BI	464	3, 9, 9, 13, 18, 19, 43, 44, 55, 105, 146	J	64, 157, 243	F	75, 153, 236	A	JFA	
<i>G. morhua</i>	GM	464	29, 93, 96, 110, 117	K	42, 145, 277	A	464	E	KAE	

^a The digested PCR amplicon comprises 36–37 bp from the 3'-end of the tRNA-*Glu* gene and 428 bp from the 5'-end of the cytochrome *b*. Bolded capitals indicate the restriction pattern that unambiguously distinguishes the corresponding species with that enzyme. The last column indicates the unique composite haplotype pattern for each species.

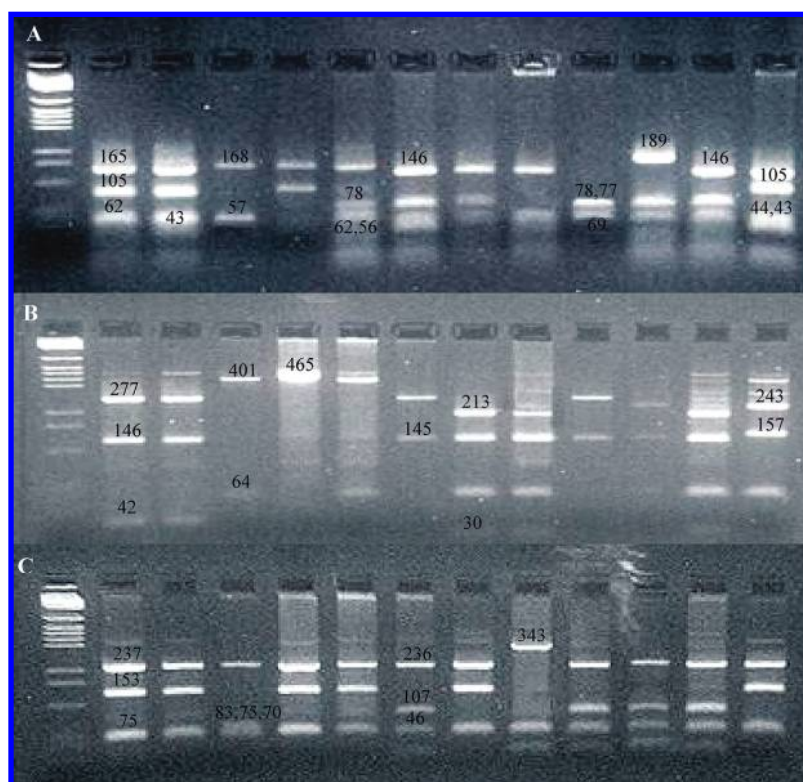


Figure 2. Agarose gel (3%) showing restriction fragments from digestion of the tRNA-*Glu*-cytochrome *b* PCR amplicon of hakes (464–465 bp) with *Mnl* I (**A**), *Bfa* I (**B**), and *Afa* I (**C**). The first lane of each panel corresponds to the molecular marker pGEM (Promega), and the following lanes correspond to the species described in **Table 2** by order. The exact size of all fragments is given in **Table 2**.

CCTC(N)7!/GGAG(N)6! (*Mnl* I), C!TAG/GAT!C (*Bfa* I), and GT!AC/CA!TG (*Afa* I). The restriction patterns generated after digestion with *Mnl* I allowed the distinction of nine hakes and Atlantic cod, that is, *M. merluccius*, *M. senegalensis*, *M. polli*, *M. capensis*, *M. paradoxus*, *M. australis*, *M. hubbsi*, *M. albidus*, *M. bilinearis*, and *G. morhua* (**Table 2**, **Figure 2A**, and **Figure 3**). Three species presented identical *Mnl* I patterns, that is, *M. angustimanus*, *M. gayi*, and *M. productus*, the latter being distinguished from the two former ones with the pattern produced by *Bfa* I (**Table 2** and **Figure 2B**). Finally, *M. angustimanus* and *M. gayi* presented different restriction patterns with enzyme *Afa* I (**Table 2** and **Figure 2C**). The application of three restriction enzymes to aliquots of the PCR amplicons

resulted in the full differentiation of 12 hake species and the Atlantic cod, their composite haplotypes being species-specific in all cases (**Table 2**). It should be noted that the identification key applied did not make use of the expected restriction fragments less than 30 bp length, which were not visualized in current agarose gels and are unnecessary for species identification. Some digestions of the PCR amplicon did not go to completion in some species (e.g., *M. senegalensis* with *Bfa* I, **Figure 2B**), leaving a faint whole fragment uncut, which did not interfere at recognizing the species pattern. The most difficult species to differentiate from each other were *M. merluccius* and *M. senegalensis*, only distinguishable with a 36 or 43 bp fragment after *Mnl* I digestion, respectively. However, as

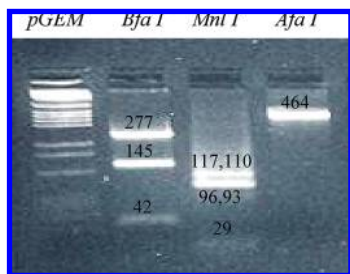


Figure 3. Agarose gel (3%) showing the restriction fragments (in base pairs) from digestion of the tRNA-*Glu*-cytochrome *b* PCR amplicon of *G. morhua* with *Bfa* I, *Mnl* I, and *Afa* I.

described below, these species were clearly differentiated using FINS and BLAST, due to the interspecific SNPs at positions 89, 167, 176, and 200 (**Figure 1**).

Methodological Validation of the PCR-RFLP Key. One of the essential properties that molecular identification methods should comply with is the intraspecific variability of the diagnostic marker calibrated. Such variability at the recognition sites of the diagnostic enzymes surely provokes a certain rate of sample misidentifications. Therefore, a compulsory task to address this problem is to carry out a validation step within and between species. Because intraspecific polymorphisms are expected in the cytochrome *b* sequence (30), the validation step consisted of applying the RFLP key to 45 individuals per species, which were previously authenticated using a morphological key. Because no exceptions were scored to the species-specific restriction pattern a priori expected and observed on two-three individuals per species, no SNPs are expected to occur frequently within the restriction targets of these diagnostic enzymes. Therefore, the PCR-RFLP identification key developed showed a high robustness for the sample sizes analyzed.

Identification of Hakes Using FINS or BLAST. Ancient DNA samples or commercial products often contain DNA that can be somehow amplified and compared to standardized sample types using methods for phylogenetic reconstruction (31). The FINS of cytochrome *b* served here at building a phylogenetic tree that showed one independent and well-supported cluster for each species (**Figure 4**). The phylogenetic reconstruction including previously identified test samples by PCR-RFLP (i.e., samples 1–5 in **Figure 4**) unambiguously assigned them to the expected single species cluster. A further assessment of the robustness of this method is patent by the correct species assignment of several cytochrome *b* sequences of hake randomly chosen from GenBank (i.e., samples EF362892, EF362889, AB248669, DQ174064, and AY946306 in **Figure 4**). Complementary to the identification methods based on RFLPs and FINS, we tested the species assignment using BLAST searches for sequence similarity. The 464–465 bp tRNA-*Glu*-cytochrome *b* sequences from all authenticated hake specimens fully matched (expected value = 0) to one of the sample types of *Merluccius* spp. or *G. morhua* indexed in GenBank after this study. Therefore, the species assignment using BLAST is an additional and fast identification strategy when either RFLP patterns or FINS reconstruction are not affordable. The genus-specific presence/absence test developed herein requires low tissue integrity because the diagnostic amplicon spans only 122 bp; therefore, it can be currently amplified from most commercial products. This preliminary test avoids the development of costly and time-consuming identification protocols when there is no hake meat in a commercial sample. The identification key that follows the exclusion test is cheaper and faster than any previous method so far developed for hakes because it only requires the use of one restriction enzyme to identify most species. The

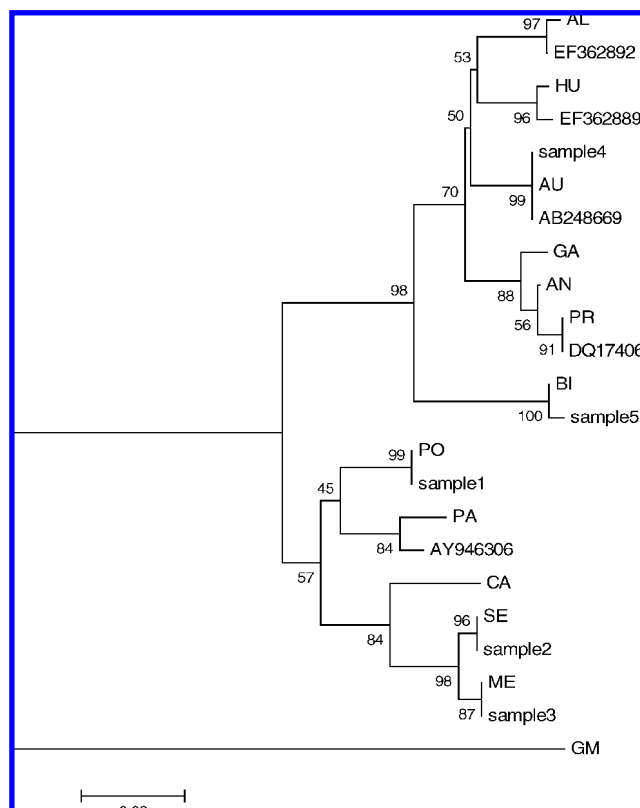


Figure 4. Neighbor-joining tree constructed using maximum composite likelihood distance. Example of a phylodiagnostic tree used to ascribe test samples (1–5) to species by using FINS. EF362892, EF362889, AB248669, DQ174064, and AY946306 are hake cytochrome *b* sequences available from the GenBank nucleotide database.

versatility of the present key permits not only the species identification using PCR-RFLP but also its confirmation by FINS or BLAST when sequence information from cytochrome *b* is available. The present molecular key together with that reported by Pérez et al. (12) are the only available methodologies that incorporate a preliminary presence/absence test and have been validated for 12 hake species. Both methods are complementary to each other because either mtDNA or nuclear DNA can be used in commercial traceability.

ACKNOWLEDGMENT

We sincerely appreciate the suggestions made by the Associate Editor, Marshall Phillips, and four anonymous reviewers to improve a previous rendition of this manuscript. We are also indebted to all of the fishing organizations that assisted in the collection of hakes in their native fisheries. Particularly, we acknowledge the sampling effort developed by J. Irisarri, B. Carrillo, I. Crespán, M. José (Valgel, S.A.), J. Ellis, L. Sampedro, 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, C. Ruiz, and M. I. Roldán. We especially thank the expert taxonomic assistance granted by C. P. Mathews. The identification methodology based on PCR-RFLPs of tRNA-*Glu*-cytochrome *b* described herein is the property of the University of Vigo, which is protected under Spanish patent #2237285/B2 and is permitted for research use only.

LITERATURE CITED

- (1) Inada, T. Studies on merlucciid fish. 1981. *Bull. Far Seas Fish. Res. Lab. Shimizu, Jpn.* **1981**, *18*, 1–172.
- (2) Lloris, D.; Matallanas, J.; Oliver, P. *Merluzas del Mundo (Familia Merlucciidae). Catálogo Comentado e Ilustrado de las Merluzas Conocidas*; FAO Catálogo de Especies Para Los Fines De La Pesca: Roma, Italy, 2003; Vol. 2, pp 1–57.
- (3) Matallanas, J.; Lloris, D. Description of *Merluccius tasmanicus* sp. nov. and redescription of *Merluccius australis* (Pisces: Merlucciidae). *J. Mar. Biol. Assoc. U. K.* **2006**, *86*, 193–199.
- (4) Roldán, M. I.; Pla, C. Species identification of two sympatric hakes by allozymic markers. *Sci. Mar.* **2001**, *65*, 81–84.
- (5) 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.
- (6) Etienne, M.; Jérôme, M.; Fleurence, J.; Rehbein, H.; Kündiger, R.; Mendes, R.; Costa, H.; Pérez-Martín, R.; et al. Identification of fish species after cooking by SDS-PAGE and urea IEF: a collaborative study. *J. Agric. Food Chem.* **2000**, *48*, 2653–2658.
- (7) 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.
- (8) Carrera, M.; Cañas, B.; Piñeiro, C.; Vázquez, J.; Gallardo, J. M. Identification of commercial hake and grenadier species by proteomic analysis of the parvalbumin fraction. *Proteomics* **2006**, *6* (19), 5278–5287.
- (9) Quinteiro, J.; Vidal, R.; Izquierdo, M.; Sotelo, C. G.; Chapela, M. J.; Pérez-Martín, R. I.; Rehbein, H.; Hold, G. L. 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.
- (10) Hold, G. L.; Russell, V. J.; Pryde, S. E.; Rehbein, H.; Quinteiro, J.; Vidal, R.; Rey-Méndez, M.; Sotelo, C. G.; et al. Development of a DNA-based method aimed at identifying the fish species present in food products. *J. Agric. Food Chem.* **2001**, *49*, 1175–1179.
- (11) 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. Aquat. Food Prod. Technol.* **2004a**, *13*, 49–57.
- (12) Pérez, M.; Vieites, J. M.; Presa, P. An ITS1 DNA-based methodology to identify world-wide hake species of the genus *Merluccius*. *J. Agric. Food Chem.* **2005**, *53*, 5239–5247.
- (13) 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.
- (14) Pérez, M.; Álvarez, C.; Balado, M.; Cabado, A. G.; Vieites, J. M.; Presa, P. Identification of South Atlantic hakes (*M. australis* and *M. hubbsi*) in processed foods by PCR-RFLPs of cytochrome b gene. *J. Aquat. Food Prod. Technol.* **2004b**, *13* (2), 59–67.
- (15) Meyer, A. Evolution of mitochondrial DNA in fishes. In *Biochemistry and Molecular Biology of Fishes*; Hochachka, P. W., Mommsen, P., Eds.; Elsevier Press: Amsterdam, The Netherlands, 1993; Vol. 2, pp 1–38.
- (16) 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.
- (17) Sambrook, J.; Fritsch, E. F.; Maniatis, T. 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, 1998; pp E3–E4.
- (18) Wolf, C.; Burgener, M.; HuKbner, P.; LuKthy, J. PCR-RFLP analysis of mitochondrial DNA: Differentiation of fish species. *Lebensm.-Wiss. Technol.* **2000**, *33*, 144–150.
- (19) Edelman, I.; Olsen, S.; Devereux, J. Program Manual for the Wisconsin Package, version 1.0; Genetics Computer Group (GCG): Madison, Wisconsin, 1996.
- (20) 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.
- (21) Heiman, M. Webcutter 2.0, 1997; <http://www.firstmarket/cutter/cut2.html>.
- (22) Roberts, R.; Macelis, D. REBASE-restriction enzymes and methylases. *Nucleic Acids Res.* **2001**, *29*, 268–269.
- (23) 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.
- (24) Tamura, K.; Dudley, J.; Nei, M.; Kumar, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **2007**, *10.1093/molbev/msm092*.
- (25) Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425.
- (26) Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *39*, 783–791.
- (27) Tamura, K.; Nei, M.; Kumar, S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *PNAS* **2004**, *101*, 11030–11035.
- (28) 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.
- (29) 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.
- (30) Grant, W. S.; Clark, A. M.; Bowen, B. W. Why RFLP analysis of control region sequences failed to resolve sardine (*Sardinops*) biogeography: Insights from mitochondrial DNA cytochrome b sequences. *Can. J. Fish. Aquat. Sci.* **1998**, *55*, 2539–2547.
- (31) 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 & Hall: New York, 1996; pp 10–49.

Received for review June 3, 2008. Revised manuscript received August 15, 2008. Accepted August 29, 2008. This research was supported with Grant RZ00/020 from the Spanish Ministerio de Ciencia y Tecnología (INIA).

JF801700X