Application of DNA-based Techniques for Intraspecies Differentiation of Hake Fish

Zora Hubálková¹, Petr Králík², Janka Kasalová³, Eva Renčová¹

¹Department of Analytical Biotechnology, Veterinary Research Institute, Brno, Czech Republic
²Department of Food and Feed Safety, Veterinary Research Institute, Brno, Czech Republic
³Department of Meat Hygiene and Technology, University of Veterinary and Pharmaceutical Sciences, Brno, Czech Republic

Received October 30, 2008
Accepted May 18, 2009

Abstract

Fish species identification is important as the interest of consumers in sea-fish meat is increasing. The aim of this study was to determine hake species distribution on the Czech market by the PCR-RFLP and sequencing of mt cyt b, and to develop and optimise an alternative system for determination of hake species by sequencing and/or PCR-RFLP using Pan I sequence. Among 20 samples of hake obtained on the Czech market three species were identified: North Pacific hake (Merluccius productus), Argentine hake (Merluccius hubbsi) and South Pacific hake (Merluccius gayi). The approaches tested in our study represent a significant tool either for the differentiation of hake species from other gadoid species (PCR) or for intraspecies identification of different hake species (PCR-RFLP, sequencing). This knowledge can be applied in detection of fish species substitution within the consumers’ rights protection.

Merluccius, PCR-RFLP technique, pantophysin I gene, cytochrome b gene, genomic DNA

Hake belong to an important group of gadoid fish (order Gadiformes, family Merlucciidae) imported to the Czech Republic in large numbers. Most often fish are imported as fillets, compressed blocks of meat or even processed products, which makes species identification based on morphological features impossible. Therefore, developing methods for species identification is crucial for the prevention of species substitution. Substitution of hake species for other gadoid species and other hake species is often noted. This is the reason for ensuring authenticity of the product, and preventing adulteration and improper labelling.

Methods based on protein analysis present an important tool for hake species identification. The use of isoelectric focusing (IEF) was described for Merluccius merluccius, M. australis, M. hubbsi, M. gayi and M. capensis identification (Pineiro et al. 2000). Another method, i.e. two-dimensional electrophoresis (2DE) was used for the detection of Merluccius merluccius, M. australis, M. hubbsi and M. capensis (Pineiro et al. 1998; Pineiro et al. 2001). For the identification of Merluccius merluccius urea-IEF (Etienne et al. 1999), sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (Rehbein et al. 1999; Pineiro et al. 1999) and chromatography (Pineiro et al. 1997) were employed. However, all protein-detection methods are very laborious, time consuming and require a well defined standard for comparison with unknown samples.

On the other hand, the introduction of molecular genetic methods, mainly PCR and its modifications (PCR-RFLP) or sequencing provides a major improvement in fish species differentiation. The use of PCR-RFLP which consists of PCR amplification of the mitochondrial cytochrome b (mt cyt b) gene following enzymatic restriction was described for the identification of a broad spectrum of Merluccius sp. (Hold et al. 2001; Quinterio et al. 2001). For hake species differentiation another genetic marker, the 5S rRNA gene was used in PCR based on amplification of the non-transcribed spacers (NTSs) within this gene (Perez and Garcia-Vazquez 2004). The PCR method based on the partial pantophysin I (Pan I) genomic sequences was developed for the identification of six gadoid fish species (Hubalkova et al. 2008).
The objective of this work was (1) to determine hake species distribution on the Czech market following our previous study (Hubalkova et al. 2008), by PCR-RFLP and sequencing of mt cyt b, and (2) to develop and optimise an alternative system for the determination of hake species by sequencing and/or PCR-RFLP using Pan I sequence.

Materials and Methods

Samples of hake and DNA analysis
Twenty samples of frozen fish products (declared as codfish) were obtained on the Czech market. The samples were either fillets or compressed blocks of meat. All samples were stored at –20 °C until DNA preparation. DNA was isolated using DNeasy Tissue Kit (Qiagen, Hilden, Germany) as described in Hubalkova et al. (2008).

Conventional PCR analysis of pantophysin I (Pan I) gene
The conventional PCR method was carried out according to Hubalkova et al. (2008) as a monoplex PCR reaction using hake-specific primer pairs based on the partial Pan I genomic sequence amplification. The PCR reaction was performed in a final volume of 20 μl using the ThermoCycler Machine DYAD (Bio-Rad, Munich, Germany). The reaction mix contained 10 μl of PPP Master Mix (Top Bio, Prague, Czech Republic), 10 pmol of each primer set, and 4 μl of the DNA template. The amplification was carried out with an initial denaturing step at 95 °C for 3 min, followed by 35 cycles and each cycle with denaturation at 95 °C for 10 s, annealing at 65 °C for 20 s, extension at 72 °C for 2 min, and then a final extension at 72 °C for 5 min.

PCR-RFLP analysis of mt cyt b gene
On account of the results of the conventional PCR method, the PCR-RFLP technique based on the mt cyt b gene was chosen for closer intraspecies (Merluccius spp.) determination. The PCR and restriction cleavage with HaeIII, HinfI and AciI were carried out as described previously (Hold et al. 2001).

Sequencing of PCR products from Pan I and mt cyt b genes
From each sample, two independent PCR reactions of Pan I and mt cyt b genes were carried out to eliminate Taq polymerase errors in the evaluation of the sequences. Obtained amplicons were cloned into the vector pDrive Cloning Vector (Qiagen, Hilden, Germany) and inserted into Escherichia coli Top10F’ competent cells and screened by the blue-white test. Randomly chosen white colonies were checked by PCR for the presence of the insert using specific and M13 universal primers (Hubalkova et al. 2008). Positive clones were purified by QIAprep Spin Miniprep kit (Qiagen, Hilden, Germany) and the PCR products were sent for sequencing to a commercial facility (MWG Biotech, Ebersberg, Germany).

PCR-RFLP analysis of Pan I gene
Based on the analysis of restriction sites by WebCutter (http://rna.lundberg.gu.se/cutter2/index.html) of partial Pan I sequences of Merluccius spp. gained by sequencing, restriction enzyme HincII (New England Biolabs, Ipswich, USA) was selected for possible differentiation of Merluccius spp. PCR products of Pan I partial sequence were purified using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and 14 μl of each PCR product was digested with 5 units of enzyme in a final volume of 20 μl at 37 °C overnight.

Results and Discussion

Conventional PCR analysis of Pan I gene
With the conventional PCR technique using the specific primer pairs derived from the partial genomic sequence of the Pan I gene all 20 investigated samples have amplified the same length of the PCR product of 201–202 bp. Obtained PCR products were cloned and sequenced to reveal the potential intraspecies variability among particular hake sequences.

Determination of Merluccius spp. by mt cyt b PCR-RFLP
For closer intraspecies differentiation among Merluccius spp., the PCR-RFLP analysis of the mt cyt b gene (464 bp) was chosen (Hold et al. 2001). Using a combination of the restriction enzymes HaeIII, HinfI and AciI (Table 1, Plate XVII, Fig. 1), we identified North Pacific hake (Merluccius productus), Argentine hake (Merluccius hubbsi) and South Pacific hake (Merluccius gayi). Nevertheless, the restriction patterns obtained after digesting the 464 bp amplicon of the samples 10–20 were quite weak (Table 2), so that the identification of these species was not absolutely confirmative using this method. On that account we aimed to devise an alternative system more suitable for hake species differentiation.
Table 1. The RFLP profiles (bp) from cytochrome b sequence (464 bp) and pantophysin I sequence (202 bp)

<table>
<thead>
<tr>
<th>Species</th>
<th>Cytochrome b gene (464 bp)</th>
<th>Pantophysin I gene (202 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hinf I</td>
<td>Hae III</td>
</tr>
<tr>
<td>M. gayi</td>
<td>U</td>
<td>39, 109, 315</td>
</tr>
<tr>
<td>M. productus</td>
<td>81, 382</td>
<td>39, 109, 315</td>
</tr>
<tr>
<td>M. hubbsi</td>
<td>81, 382</td>
<td>39, 109, 126, 189</td>
</tr>
</tbody>
</table>

U - unaffected

Table 2. Intraspecies identification of 20 hake species using different molecular biological methods

<table>
<thead>
<tr>
<th>Sample No.</th>
<th>Conventional PCR method (Pan I gene – 202 bp)</th>
<th>PCR-RFLP analysis (mt cyt b gene); REs: Hinf I, Hae III, AcI I (Hold et al. 2001)</th>
<th>Sequencing of mt cyt b gene</th>
<th>PCR-RFLP analysis (Pan I gene); RE: Hinc II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Merluccius sp. M. gayi</td>
<td>M. gayi</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Merluccius sp. M. productus</td>
<td>M. productus</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Merluccius sp. M. productus</td>
<td>M. productus</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Merluccius sp. M. productus</td>
<td>M. productus</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Merluccius sp. M. hubbsi</td>
<td>M. hubbsi</td>
<td>M. hubbsi</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Merluccius sp. M. productus</td>
<td>M. productus</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Merluccius sp. M. hubbsi</td>
<td>M. hubbsi</td>
<td>M. hubbsi</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Merluccius sp. M. gayi</td>
<td>M. gayi</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Merluccius sp. M. gayi</td>
<td>M. gayi</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Merluccius sp. M. hubbsi *</td>
<td>M. hubbsi</td>
<td>M. hubbsi</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Merluccius sp. M. productus *</td>
<td>M. productus</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Merluccius sp. M. productus *</td>
<td>M. productus</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Merluccius sp. M. hubbsi *</td>
<td>M. hubbsi</td>
<td>M. hubbsi *</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Merluccius sp. M. hubbsi *</td>
<td>M. hubbsi</td>
<td>M. hubbsi</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Merluccius sp. M. productus *</td>
<td>M. productus</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Merluccius sp. M. hubbsi *</td>
<td>M. hubbsi</td>
<td>M. hubbsi</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Merluccius sp. M. hubbsi / M. productus *</td>
<td>M. hubbsi</td>
<td>M. hubbsi</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Merluccius sp. M. hubbsi *</td>
<td>M. hubbsi</td>
<td>M. hubbsi</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Merluccius sp. M. gayi / M. hubbsi *</td>
<td>M. productus</td>
<td>M. gayi/M. productus</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Merluccius sp. M. hubbsi *</td>
<td>M. hubbsi</td>
<td>M. hubbsi</td>
<td></td>
</tr>
</tbody>
</table>

The RFLP profile not clearly identifiable

**Sequencing**

The corresponding PCR products obtained from the mt cyt b gene of all 20 samples were also sequenced. Nucleotide sequences were compared and aligned with sequences of the mt cyt b gene published in the GenBank. The samples that were not unambiguously identified by the PCR-RFLP technique, were determined (Table 2) by sequencing of the mt cyt b gene. Sequencing confirmed the results of PCR-RFLP analysis in all but two samples (17 and 19). According to these results, the sequences of 20 samples of the Pan I gene were aligned, and three sequences of the Pan I gene for North Pacific hake (*Merluccius productus*), Argentine hake (*Merluccius hubbsi*) and South Pacific hake (*Merluccius gayi*) were deposited in the GenBank (Acc. Nos. EU301630, EU301631, EU301632). Alignment of sequences showed that the similarity among them varies from 97.0 to 99.5% (Fig. 2).

**PCR-RFLP analysis of Pan I gene**

In order to develop an alternative technique for *Merluccius* spp. differentiation, we employed PCR-RFLP analysis of the Pan I partial genomic sequence. *In silico* analysis
revealed that it will probably not be possible to distinguish between *M. gayi* and *M. productus* because there is a SNP in position 166 in *M. productus* where G could be replaced by A (Fig. 2). This also causes amino acid substitution (Asparagine to Serine). No other possible targets for restriction analysis were found. By the PCR-RFLP of the *Pan* I gene we were able to reliably determine *M. hubbsi* only. For these purposes, the restriction enzyme *Hinc* II, that can distinguish *M. hubbsi* (unaffected by *Hinc* II) from *M. gayi* and *M. productus* (both cleaved with *Hinc* II into two fragments in position of site 118) was generated (Table 1). These results were compared with the results of the PCR-RFLP and sequencing of the mt cyt *b* gene whereas hake intraspecies identification using *Hinc* II cleavage of the *Pan* I gene was confirmed (Table 2). Concerning the discrepancy between the results of PCR-RFLP and sequencing of the mt cyt *b* gene of samples 17 and 19, the PCR-RFLP results of the *Pan* I gene of samples 17 and 19 were consistent with the results of mt cyt *b* gene sequencing. The RFLP profiles of the mt cyt *b* gene of the samples 17 and 19 did not fully correspond with the restriction patterns (Table 2) as described by Hold et al. (2001). Visual interpretation of the signals could simulate the presence of a mixture of two *Merluccius* species; however, this was disproved by the sequencing of mt cyt *b* gene.

The aim of this study was to illustrate the usefulness of different methods of molecular biology for hake species discrimination. For the differentiation of hake species from other gadoid species the conventional PCR method based on the amplification of a partial genomic sequence of the *Pan* I gene can be used (Hubalkova et al. 2008). Nevertheless, for intraspecies identification of hake species the application of other methods (e.g. the PCR-RFLP technique) is advisable. The PCR-RFLP of the mt cyt *b* gene may be used for intraspecies identification of hake species (Hold et al. 2001), although our results were not always of predicative significance. The innovative contribution of this study lies in the application of the partial *Pan* I sequences of three hake species (Fig. 1) as an alternative tool for detection of the above mentioned species in fish meat and fillets. PCR-RFLP of the *Pan* I gene using *Hinc* II can differentiate *M. hubbsi* from *M. productus* and *M. gayi*. The most direct means of obtaining information from PCR products is by direct nucleotide sequencing employing universal primers (Kocher et al. 1989). Sequencing
offers unambiguous identification of individuals of a single species of hake. Considering the higher cost and time consumption, sequencing is not suitable for routine analytical purposes. All the above described and approved techniques represent an important tool for disclosing fish food adulteration in terms of species substitution detection. Among other things, the study also gives an overview of the spectrum of hake species offered on the Czech market.

**Acknowledgements**

The work was supported by the Ministry of Agriculture of the Czech Republic (Grants No. QG60090/2005, MZE 0002716201 and MSM 6215712402 Veterinary Aspects of Food Safety and Quality). We thank Dr. Grant H. Pogson (University of California, Santa Cruz, USA) for providing sequences of hake species originated from his research study. We gratefully acknowledge David McShane (University of Aberdeen, Scotland) for critical reading of the English grammar.

**References**


Fig. 1. RFLP profiles from mt cyt b gene (464 bp) fragmented with enzymes *Hinf*I, *Hae*III and *Acil* (1 - *M. gayi*, 2 - *M. productus*, 3 - *M. hubbsi*); 50 bp marker (M)