Computerised identifying databanks, based on IEF patterns, for the authentication of commercial flatfish and roundfish species

Peter Bossier

Vlaams Instituut voor de Zee
Flanders Marine Institute

Ministerie van Middenstand en Landbouw, Centrum voor Landbouwkundig Onderzoek-Gen.
Departement Zeevisserij
Ankerstraat 1, 8400 Oostende, Belgium
Tel : 32 59 34 22 50, Direct tel : 32 59 34 22 66, Fax : 32 59 33 06 29, e-mail: pbossier@yucom.be

Abstract

Two computerised identifying databank of IEF protein patterns of respectively flatfish and roundfish species are presented. It was found that interspecimen similarity of the IEF patterns, as processed by digitisation, was always larger than interspecies similarity (except for two roundfish species), which allows for unequivocal authentication of unknown samples, as long as the authentic pattern is available in the databank. The databank was used to authenticate commercial frozen fish. Sometimes, authentication of frozen fish from species not present in the databank (e.g. *Gadus macrocephalus*) was assisted by high similarity to a pattern available in the databank (e.g. *Gadus morhua*). It is also demonstrated that IEF patterns available on the Internet can be incorporated into that databank. This opens the possibility to set up an integrated databank, joining IEF patterns available in various laboratories. As commercial software is now available to access such integrated databanks in a Client-Server set-up over the Internet, the construction of such an integrated databank should be considered.

Introduction

Since May 22nd 1996 a Belgian law enforces the use of official names for the labelling of fish and seafood products. Flatfish species are economically important species for the Belgian fisheries representing about 50% of the annual catch in 1996 and about 80% of the economical value at the auction. Sole (*Solea solea*) on its own, represented about 50% of the total value in 1996. Because of its fine texture and taste and its limited supply this species is highly appreciated. The auction-value of less-appreciated species like *Microstomus kitt, Pleuronectus platessa* and *Limanda limanda* is 3 to 6 times less with respect to sole (*S. solea*). Especially when these or other flatfish species are sold as fresh or frozen filets, involuntary or deliberate mislabelling can be a problem. Since filets of most flatfish species are sold unprocessed, (except for e.g. *Hippoglossus spp, or Reinhardtius hippoglossoides* which are often sold smoked) authentication of fish filets

can be done by generating species-specific protein patterns.

From the fish processing industry regular requests are received with relation to the authentication of roundfish species in frozen blocks, used to make fish sticks.

Samples containing unknown fish meat can be authenticated by iso-electric focussing (IEF) of water soluble proteins, a technique that has been around for some time now (Lundstrom, 1981, 1983; Rehbein 1995, 1998). In this article two computerised and identifying databanks containing IEF patterns (IEF: isoelectric focussing) of sarcoplasmic proteins of all flatfish and almost all lean roundfish species, described in the Belgian law mentioned above, have been constructed and are partly presented. These databanks have been successfully used to authenticate frozen fish.

Material and Methods

Authentic fish species and commercial filets

Authentic fish species were either caught during campaigns of research vessels or commercial vessels or bought in local retail shops or at the fish auction. The species were identified on the basis of external morphological characteristics according to Poll (1949) and stored frozen at –20°C. For some other species authentic material (e.g. whole fish) was not available. For these species frozen fish filets were obtained through fish processing companies. Hence the authenticity of these samples remains equivocal. Only one such sample (namely hoki) has been included in the figures.

Protein extraction

About 25 g fish meat (only white muscle) was minced and subsequently mixed with 25 ml CCl_4 in a morter. The CCl_4 was decanted after 30 min. The fish meat was transferred to a Warring blender beaker and 15 ml of glycine mix was added (glycine mix: 1% glycine in distilled water, pH 6.5; if less muscle tissue could be recovered from the fish, extraction volumes were change proportionally). After 15 seconds mixing at low speed the content was transferred to a centrifuge tube. The supernatant (10 min at 20 000 g) was filtered (S&S 595 ½). Total protein concentration was determined (Lowry method) and the filtrate was adjusted to 5 mg protein per ml using glycine extraction mix. Extracts were made either from fresh fish or frozen fish. Extracts were processed (e.g. electrophoresis) within 24h. Aliquots to which 15% of glycerol was added were stored at -80°C. These aliquots were not used to produce entries into the databanks.

Electrophoresis

Using a Ampholine PAG plates (pH range 3.5-9.5, Pharmacia) in combination with a Multiphor II electrophoresis equipment (Pharmacia), proteins (15 µg spotted on paper strips) were separated according to their pI (isoelectric point: position in the pH range where a protein has no net charge and hence zero electrophoretic mobility) at 5°C (electrophoretic conditions: max 1400 V; max 20mA; max 10 W). The distance between the electrodes was 10 cm. Fish protein samples were placed at 2 cm from the cathode. In order to facilitate standardisa-

tion of protein patterns, a pI standard (Pharmacia 3.5 to 9.3 pI range) was run on the gel as well. On one gel containing 8 samples three standard pI ladders were loaded. The pI ladder was loaded at 7 cm distance from the cathode.

Fixing and staining

The gels were fixed for 10 min (trichloroacetic acid 115 g/l, sulphosalicylic acid 34.5 g/l), soaked in destaining solution for 30 min (25% ethanol, 0.8% acetic acid), stained at 60°C (25 % ethanol, 0.8% acetic acid, Coomassie brilliant blue R250 1.15 g/l) for 10 min, destained for 24h in the same liquid in a destain bath (Biorad 556) (the destaining bath was used maximum three consecutive times) and soaked in preserving solution (10% glycerol). Gels were dried at 80°C under vacuum.

Gel processing and databank characteristics

Gels were digitalised with a flat bed scanner (HPscanJet IIcx) at 400 dpi. The images were processed and analysed with the software package Gelcompar (Applied Maths, Kortrijk, Belgium) (Bossier and Cooreman, 2000). The UPGMA (UPGMA: unweighted pair-group method of arithmetic averages) dendrograms presented are calculated on the basis of Pearson correlation coefficients or Dice distances. The software searches automatically for lanes and allows straightening them. This feature was only used on lanes that were slightly inclined. Lanes with obvious distortions were not used. The centre half parts of the lanes were further processed. Some of the software settings can be critical while constructing the databank. The software allows for background substraction, the settings of which should never be changed during the construction of the databank, as the Pearson correlation coefficient is sensitive to background staining levels. For calculating the Dice distance values, bands have to be defined. Again software settings can greatly influence the amount of bands detected. A dendrogram was made based one of these software settings. The shown dendrogram was a typical output. In total, the two databanks contain 50 species so far. These databanks were built with 32 gels, each of them containing eight samples and three pI markers. For a species to be included in the databank the IEF patterns of at least two specimens are introduced.

Results

The IEF patterns obtained were processed in two different ways, using the software GelCompar. Either the Pearson correlation coefficients or the Dice index were calculated. For calculating the Dice index, bands have to be scored (a process executed by the software for which a certain amount of settings have to be chosen. These settings can greatly influence the amount of bands being scored. The Dice index is calculated using the formula $2n_{AB}/n_A+n_B$ (in which n_{AB} number of bands common in A and B, and n_A or n_B total number of bands in respectively A or B). Using the Dice index as input, an UPGMA dendrogram was obtained in which the specimen of the same species were not necessarily clustering together. This result was obtained indepen-

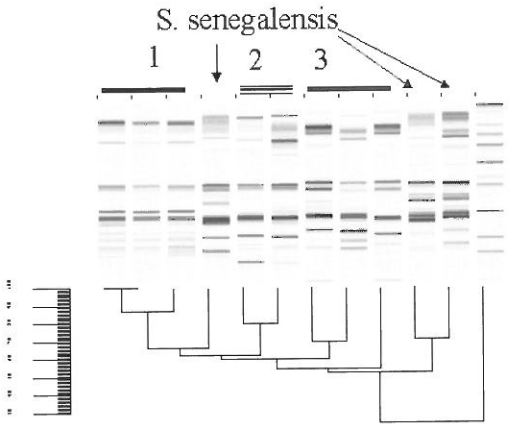

Fig. 1. UPGMA dendrogram of IEF patterns of members of the Soleidae family using Dice similarities as input. 1: Monochirus luteus, 2: Solea solea, 3: Solea lascaris 4: Solea senegalensis. Last lane: pI ladder.

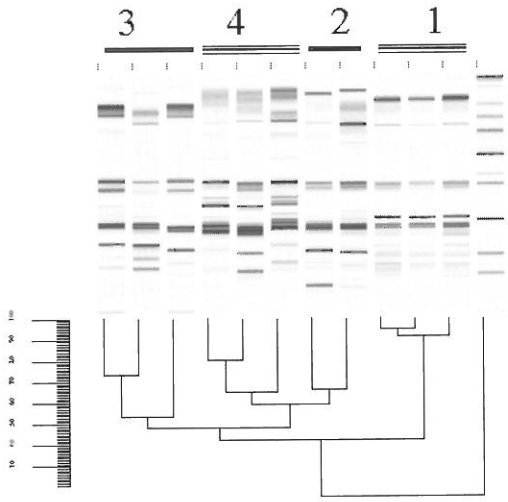

Fig. 2. UPGMA dendrogram of IEF patterns of members of the Soleidae family using Pearson correlation coefficients as input. 1: Monochirus luteus, 2: Solea solea, 3: Solea lascaris 4: Solea senegalensis. Last lane: pI ladder.

dent of the band search settings. Using the Pearson coefficient as input, UPGMA dendrograms were obtained in which specimens of the same species were always clustering together (see Fig 1 and 2). The Pearson or product-moment correlation coefficient calculates the congruence between arrays of values, in this case densitometric arrays. This similarity index is sensitive to relative band intensities, while the Dice index only takes the presence or absence of a band into account. In other words using the band intensity as a (extra) source of information rather than only its presence or absence, more consistent clustering between IEF patterns of the specimen of the same species was obtained. Hence Pearson correlation coefficients were used throughout the rest of the study.

In Fig. 3 an extract from the databank for roundfish species is shown (The databank contains 10 gadoid roundfish species and 23 other lean roundfish species, excluding salmonids as distinction between salmonids with IEF patterns processed in the described way is very difficult). Two particular elements are worthwhile commenting upon. Firstly, a commercial sample, labelled "onb" in Fig.3, was found to be closely related to *Gadus morhua* although it seemed not to be identical to it. The company had been able to buy a lot of that fish as cod (apparently at a rather low price) and was interested in authenticating it. After contacting its supplier and confronting him with the result, they confirmed that the fish was actually *Gadus macrocephalus*. Secondly, with the current state of art and in contrast with members of the

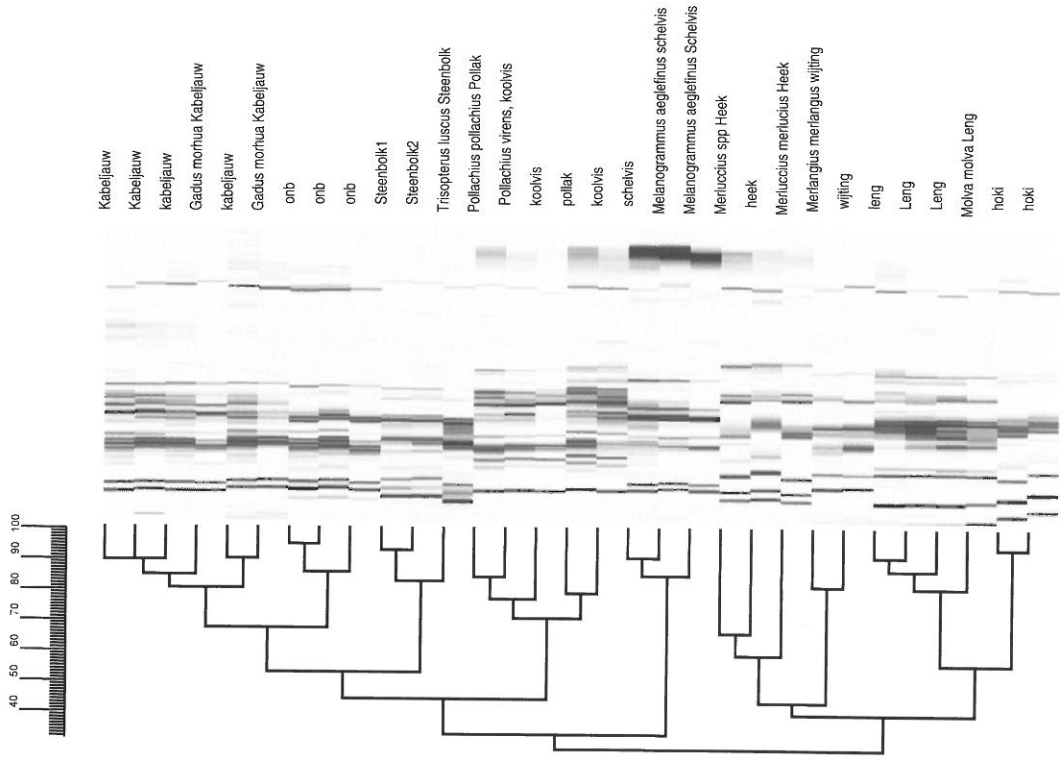

Fig. 3. *Extract from the databank of lean roundfish species. "onb": unknown sample, see text.*

genus *Solea* (Fig.2 and Bossier and Cooreman, 2000), it does not seem to be possible to distinguish between *Pollachius pollachius* and *Pollachius virens*. This might be a consequence of way in which the IEF patterns are processed and similarities calculated. This example together with the result of the *Gadus* spp.and *Solea* spp. illustrate that related species (at the genus level) seem to cluster together when the IEF patterns are processed in the described way. This opens up the possibility to start with the identification of an unknown pattern even if the authentic pattern is not available in the databank.

Discussion

Bossier and Cooreman (2000) have been validating the IEF-based authentication technique (IEF and gel processing) for a flatfish databank. The average Pearson correlation coefficient over 10 flatfish species was 79.9 (11.5, when two or three specimens were analysed. This meant that in the IEF patterns, as processed by the GelCompar software, a lot of variability was detect-

ed in between specimens. Yet, in the generated UPGMA dendrogram, IEF patterns from different specimen of the same species were always clustering together, which indicates that the variability in between species is higher than in between specimens. This finding corroborated with previous findings (e.g. Rehbein et al., 1995) that species identification with IEF is possible. These observations also emphasised that including IEF patterns of various specimens in an identifying databank facilitates the identification process. These observations are substantiated here by the results obtained with lean roundfish species. The average intraspecies similarity level of the lanes shown in Fig. 3 was 84 (7, while the average interspecies similarity level (to the nearest species) was 50 (13. In a t-student test those averages were highly significantly different ($P<0.001$). The average intraspecies variability also indicates how high the similarity of an unknown pattern to a database entry has to be to allow for an unequivocal identification.

The software GelCompar allows for the integration of whatever gel into the databank as long as it carries the same pI standard ladder (with the latest version of the software patterns produced on gels with different pI ladders can be integrated in one databank). In order to test the performance of the authentication technique, an IEF pattern available over the internet (IEF from *Platichthys stellatus* on the FDA website) was introduced into the flatfish databank (Bossier and Cooreman, 2000). In the UPGMA dendrogram, that pattern linked up with the *Platichthys flesus* pattern, as could be expected for two species from the same genus (similarity level 40%). Integration of IEF patterns produced by more than one laboratory into one databank would require a thorough standardisation of the methodology, like protein extraction and determination procedures and the standardisation of the settings for the digital processing of gels. Since standardisation of the IEF authentication procedure has proven to be possible (Rehbein et al., 1995), it should be feasible to proceed with the construction of an identifying databank of fish IEF patterns produced in various laboratories (Bossier, 1999) and make such databank available over the Internet. Again the software necessary to construct an interactive web-based identifying databank has recently become available (e.g. Applied Maths, Belgium). Such a web-based interactive databank could be a very useful tool for authenticating fish meat as it could cover on the one hand all polymorfism found within a species and on the other hand a very wide range of species, a task one particular lab would have a hard time trying to gather.

Conclusion

Our results confirm that authentication of fish filets by IEF of proteins are feasible and that a computerised identifying databank can assist in that. A thorough standardisation of the IEF procedure would open the possibility to exchange authentic IEF patterns in between laboratories and to construct a web-based interactive databank.

Acknowledgement

We thank Pieters Visbedrijf for supplying some fish species and IPIMAR (Portugal) for supplying *S. senegalensis*. We also thank J. De Riek and M. De Loose from CLO-DVP for the use of the GelCompar software and M. Vandenbouhede (CLO-DVZ) for technical assistance.

Literature

Bossier, P. and Cooreman, K. 2000. An identifying databank for the authentication of commercial flatfish (Pleuronectiformes) products at the species level by isoelectric focussing of native muscle proteins.Int J. Food Sci Technol. In press

Bossier, P. 1999. Authentication of seafood products by DNA patterns. *J. Food Sci.* 64: 189-193.

Lowry, O.H., Rosenbrough, N.J., Farr, A.L. and Randall, R.J. 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* 193: 265-275.

Lundstrom, R.C. 1981. Rapid fish species identification by agarose gel isoelectric focussing of sarcoplamic proteins. *J. Assoc. Off. Anal. Chem.* 64 : 32-37.

Lundstrom, R.C. 1983. Identification of pacific rockfish (*Sebastes*) species by isoelectric focussing. *J. Assoc. Off. Anal. Chem.* 66: 974-980.

Poll, M. 1947. *Poissons marins*. Musée Royal d'Histoire naturelle de Belgique.

Rehbein, H., Etienne, M., Jerome, M., Hattula, T., Knudsen, L.B., Jessen, F., Luten, J. B., Bouquet, W., Mackie, I.M., Ritchie, A.H. ,Martin, R., Mendes, R. 1995. Influence of variation in methodology on the reliability of the isoelectric focusing method of fish species identification. *Food Chemistry* 52:193-197.

Rehbein, H. 1998. Differenzierung roher und geräucherter Aale durch Protein- und DNA analyse. *Inf. Fischwirtsch.* 45:23-26.