JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Classification of Gilthead Sea Bream (*Sparus aurata*) from ^{1}H NMR Lipid Profiling Combined with Principal Component and Linear Discriminant Analysis

SERGE REZZI,[†,‡] IVAN GIANI,[§] KÁROLY HÉBERGER,[†,||] DAVID E. AXELSON,[⊥] VITTORIO M. MORETTI,[§] FABIANO RENIERO,[†] AND CLAUDE GUILLOU*[,†]

Physical and Chemical Exposure Unit, Institute for Health and Consumer Protection, Joint Research Centre, European Commission, BEVABS T.P. 281, I-21020 Ispra (VA), Italy, Department of Veterinary Science & Technology for Food Safety, University of Milan, Via Trentacoste 2, 20134 Milano, Italy, Chemical Research Center, Hungarian Academy of Sciences, P.O. Box 17, H-1525 Budapest, Hungary, and MRi_Consulting, 8 Wilmot Street, Kingston, Ontario, Canada K7L4V1

The combination of ^{1}H NMR fingerprinting of lipids from gilthead sea bream (*Sparus aurata*) with nonsupervised and supervised multivariate analysis was applied to differentiate wild and farmed fish and to classify farmed specimen according to their areas of production belonging to the Mediterranean basin. Principal component analysis (PCA) applied on processed ^{1}H NMR profiles made a clear distinction between wild and farmed samples. Linear discriminant analysis (LDA) allowed classification of samples according to the geographic origin, as well as for the wild and farmed status using both PCA scores and NMR data as variables. Variable selection for LDA was achieved with forward selection (stepwise) with a predefined 5% error level. The methods allowed the classification of 100% of the samples according to their wild and farmed status and 85–97% to geographic origin. Probabilistic neural network (PNN) analyses provided complementary means for the successful discrimination among classes investigated.

KEYWORDS: NMR profiling; fish; gilthead sea bream (*Sparus aurata*); chemometrics; multivariate methods; linear discriminant analysis; principal component analysis; food authenticity; geographic origin

INTRODUCTION

Nowadays availability and international trade of fish and seafood are strongly conditioned by food safety norms. Several European Directives have introduced safety standards into the chain for fisheries and aquaculture products with the concept "from farm to fork", usually based on the Codex Alimentarius provisions (*1*).

A labeling regulation for fishery and aquaculture products came into effect in the European Union in 2001, requiring identification of the official commercial and scientific name, the origin of the fish, and its production method (farmed or wild) (*2*). This regulation aims to provide consumers with basic information on characteristics of such products and is enforced at a national level, like in Italy by the Ministry of Agriculture Decree No. 27.03.02 on the labeling of fish products. On these bases there is a need for research delivering both a product-specific and a general analytical traceability system for fish products, verifying existing paper traceability schemes on production methods and the geographic origin of fish.

Gilthead sea bream (*Sparus aurata*) is an economically very important fish species, cultured in Italy and, in general, in the Mediterranean basin (*3*). The market demand and, as a result, the price for fresh sea bream have increased markedly over the past decade because of the desirable aroma and quality attributes of this fish; consequently, its farming is deemed to be a profitable business. Thus, the European fish farmers have gradually expanded their annual production from 3148 t in 1990 to 61 284 t in 2001, but on the other hand intensive production of sea bream has raised concerns over the quality of cultured fish in comparison with wild fish.

Within this frame, there is a need to develop reliable analytical methods to assess both wild and farmed fish as well as their geographic origin. Different analytical strategies can be applied to address these authenticity issues depending on the available technique (*4*). One possibility is to study the composition of lipids extracted from the fish muscle. High resolution nuclear magnetic resonance (NMR) spectroscopy has been used for the compositional study of complex mixtures of lipids providing

* To whom correspondence should be addressed. Fax: +39 0332789303. E-mail: claude.guillou@jrc.it.
[†] European Commission.
[‡] Present address: Nestlé Research Center, Bioanalytical Science, CH-1000 Lausanne 26, Switzerland.
[§] University of Milan.
[||] Hungarian Academy of Sciences.
[⊥] MRi_Consulting.

10.1021/jf070736g CCC: $37.00
Published on Web 10/31/2007

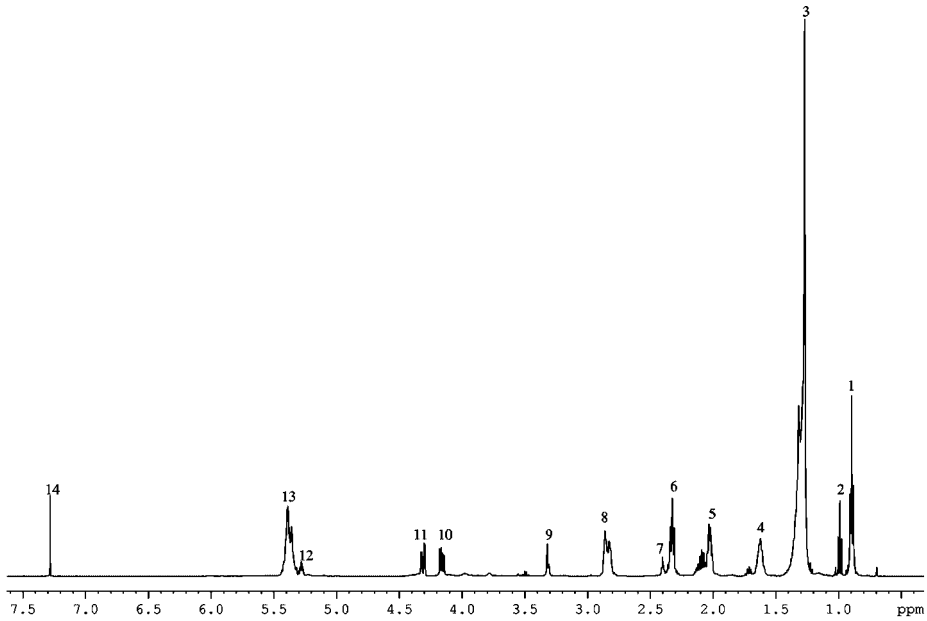

Figure 1. ^{1}H NMR spectrum of Gilthead sea bream lipid extract. 1, All fatty acids (f.a.) —(CH_3) except *n* – 3 f.a.; 2, *n* – 3 —(CH_3); 3, all f.a. —(CH_2)*n*— except 20:5 and 22:6; 4, f.a. —CH_2—CH_2—COOH except 22:6 and 20:5; 5, unsaturated f.a. —CH_2—CH=CH; 6, all f.a. —CH_2—COOH except 22:6; 7, 22:6 =CH—CH_2—CH_2—COOH; 8, polyunsaturated f.a. =CH—CH_2—CH=; 9, phosphatidylcholine —N(CH_3)$_3$; 10 and 11, glyceryl C1,3 protons; 12, glyceryl C2 protons; 13, unsaturated f.a. —C*H*=C*H*—; and 14, C*H*Cl$_3$.

both qualitative and quantitative information (*5–8*). ^{13}C NMR provides, for instance, a reliable method to unambiguously identify several fatty acids and to establish the stereochemistry of unsaturated fatty acids as well as their position on the glycerol backbone in fish lipids (*9–12*). Furthermore, ^{1}H NMR allows a rapid acquisition of the lipid profile containing information on different lipid classes, the global unsaturation level, ω-3 polyunsaturated fatty acids, and eventual oxidative deteriorations (*9, 13–17*). Lipid extracts from fish have also been studied by ^{1}H and ^{2}H NMR (SNIF NMR) (*18*). The combination of the ^{13}C/^{12}C isotope ratio measured by isotopic ratios mass spectrometry (IRMS) and D/H measured by ^{2}H NMR have been used for multivariate analysis to discriminate wild and farmed salmons (*19*).

^{1}H NMR spectroscopy has extensively been used in other areas mainly in metabolomic studies for diagnostic purposes (*20–23*). The aim of this study was to apply ^{1}H NMR spectroscopy in combination with multivariate statistics to lipid extracts from Gilthead sea bream (*S. aurata*) to achieve the distinction of wild and farmed specimens and to classify them according to the geographic origin of production.

MATERIALS AND METHODS

Collection and Preparation of Samples. From September 2003 to June 2004, 46 species of farmed sea bream (average weight = 326.7 g) were purchased from four different productive countries (Italy, Greece, Croatia, and Turkey). At the same time 19 wild sea bream (average weight = 241.4 g), caught in the Mediterranean Sea, were collected from the wholesale fish market of Milan. The collected samples were stored at −20 °C until extraction of lipids.

Extraction of Muscle Lipids. Lipids were extracted from muscular tissue according to the method described in ref *24*. Briefly, oil was extracted from 10 g of tissue with a chloroform/methanol mixture in a pyrex test tube using an ultraturrax (24 000 rpm under ice cooling) for 2 min. After extraction the lipids were stored at −20 °C.

^{1}H NMR. Oil samples (50 mg) were diluted in 700 μL of deuterated chloroform ($CDCl_3$, Aldrich, 99.8% deuteration) in a 5 mm NMR tube for analysis. NMR spectra were registered on a Bruker (Rheinstetten, Germany) DRX-500 instrument operating at 500.13 MHz for ^{1}H observations using a 5 mm inverse detection probe maintained at 300 K.

^{1}H NMR spectra were digitized into 64K data points over a spectral width of 10 000 Hz with an acquisition time of 3.3 s. An additional relaxation delay of 2 s was included, making a total recycling time of 5.3 s. A 30° pulse was used, and 128 transients were coadded before Fourier transformation. The number of dummy scans was 16.

Spectra were Fourier transformed applying a line broadening apodization function of 0.1 Hz. Chemical shifts were calibrated against the residue of protonated chloroform at 7.28 ppm. All spectra were manually phased, and baseline correction was then applied. Spectra were integrated using the AMIX software (Bruker) into a series of 165 bins (buckets) of 0.04 ppm over the range of chemical shifts from 7.2 ppm to −0.5 ppm. Buckets were expressed either as raw data or reported to the overall intensity of the spectrum (scaled). These buckets represent the descriptors for multivariate analysis. Spectral regions from 3.1 to 3.95 ppm and from 1.07 to 1.24 ppm were removed from the integration procedure because of the presence of signals from ethanol residues coming from the extraction step. By this method, the choline signal exhibiting a significant shift of its resonance (around 3.4 ppm) was also not integrated to exclude any extra variance in our data set of variables. A typical ^{1}H NMR spectrum is reported in **Figure 1**.

Chemometric Techniques. *Principal Component Analysis (PCA).* The dimensionality of the data can effectively be reduced by using PCA, particularly because the bucketed ^{1}H NMR data are highly correlated. Abstract principal components (PCs, factors) are formed from the original variables so that the new PCs should be orthogonal. The number of *n*-dimensional points can be projected in a smaller dimension hyperplane. Clusters and eventual outliers can be observed in a smaller dimension easily. In many cases, PCA models can be used for a successful classification, if the class memberships are known in advance. However, PCA optimizes the directions of largest variability (variance) and not the largest class separation ability. Two integration (bucketing) methods were applied: raw bucketing and scaling to the total intensity. Scaling the buckets to the overall intensity involves expressing each variable as a percentage of the whole lipid profile, except for obviously the excluded regions. The PCA was carried out on the correlation matrix; that is, standardization (mean centering and scaled to unit standard deviation) has been applied as data pretreatment. The model building has been performed by Statistica 6.0 program package (*25*) using the factor analysis module.

Linear Discriminant Analysis (LDA) and Canonical Correlation Analysis (CCA). In LDA a linear function of the variables is sought, which maximizes the ratio of between-class variance and minimizes

the ratio of within-class variance. LDA can also be considered as a dimension reduction method. It projects points to a smaller dimension hyperplane. LDA selects directions, which achieves maximum separation among the given classes. The latent variable (canonical variate, root) obtained by CCA is a linear combination of the original variables. If we have more variables than k classes, $k - 1$ canonical variates can be determined. A variant of this method is the stepwise discriminant analysis that permits the variables with a major discriminant capacity to be selected. The description of the LDA and CCA algorithm can be found in refs *27–29*. Discriminant models have been developed by the Statistica 6.0 program package (*26*) using the general discriminant analysis, LDA, and CCA modules.

Probabilistic Neural Network (PNN) Analysis. PNN networks have three layers: input, pattern, and summation. The input layer has as many elements as there are individual parameters needed to describe the samples to be classified. In the present case the input parameters were scores derived from PCA of the original spectroscopic data. The PNN calculations used systematically decreasing numbers of input scores to derive the best compromise between the use of a minimum number of inputs and the best predictive models. Training for PNN nets using the genetic adaptive option tests a whole range of smoothing factors (multipliers for each input), trying to optimize a combination that works best on the network created from the training set. The stopping criterion involved minimizing the average percentage of incorrect classifications over all categories. Leave-one-out cross validation (LOO) was used for the training set in all calculations to derive an optimized network. This means that LOO is used for both the random selection and the Kennard–Stone selection samples when optimizing the net. In both cases the trained network was applied to the samples that were held out of the network for the optimization process. Therefore, these "test" samples (random and Kennard–Stone) should really be viewed as "external validation" samples because they were in no way involved in the network optimization. Further details may be found in ref *29*. PNN models have been built using the Neuroshell Classifier (*30*).

Data Sets. The input data were arranged in a matrix form. Two input matrices were built: (i) raw buckets and (ii) scaled buckets (each bucket was calculated relative to the overall intensity of the spectrum). Farmed samples ($n = 46$) were derived from Italy ($n = 15$), Greece ($n = 18$), Croatia ($n = 10$), and Turkey ($n = 3$). Wild fish samples ($n = 19$) were derived from the Mediterranean basin. Five- and four-group (not considering the wild fish samples) classifications have been made.

Random sample selection and Kennard–Stone sample selection have also been used for external validation. Kennard–Stone selection maximizes the minimal Euclidean distances between already selected objects and the remaining objects (*31*). The distance between all samples is calculated from the descriptor values, and the first two points selected are those that are furthest apart. The next point selected is the one that is furthest away from these two. Ideally the number of points selected should be sufficient to cover the variable space for each class. There were 44 samples in the training set and 21 in the test data set unless otherwise noted.

Special attention was paid not to surpass the limit of $(n - g)/3$, where g is the number of groups (*32*). Therefore, a maximum of 20 and 15 variables can be incorporated in the models in the cases of five-group and four-group classifications, respectively.

The data sets are available from the authors upon request.

RESULTS AND DISCUSSION

Results based on PCA. First, we compare the PCA results from raw and scaled data. The purpose of the scaling was to overcome differences in the amount of tissue from which the oils were originally extracted and the variability in extraction.

PCA exhibits a good discrimination ability of lipid samples from wild versus farmed specimens (**Figure 2**). We would expect that raw data should perform worse in such conditions because many sources of analytical variation have not been removed. In fact, we observed that PCA analysis of scaled buckets gives better results than the PCA with raw data for the distinction of wild and farmed fish.

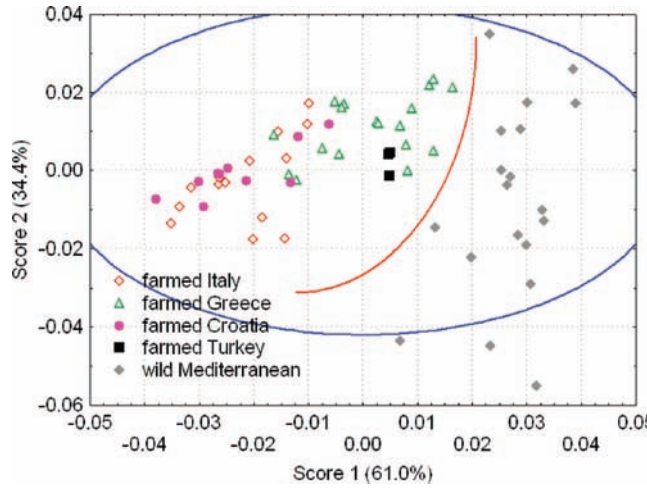

Figure 2. Results of PCA for scaled buckets, mean centered data. Discrimination of farmed and wild fish samples. Score 1 against score 2 (the total variance explained by the given PCs can be found in brackets).

The first three PCs calculated from scaled buckets explain in total 97.5% of the total variance. The examination of the loadings allows the identification of the variables mainly responsible of the observed clustering. The spectral regions defined by the buckets at 0.88, 1.28, 1.32, 2.84, and 5.4 ppm explain the main variation on the first two components (**Figure 2**). The typical spectrum assignment of lipids from fish reported in **Figure 1** explains the separation of wild fish by the higher intensities of methyl (0.88 ppm) and methylene protons (1.32 ppm) together with methylene (2.84 ppm) and methyne protons (5.4 ppm) in unsaturated fatty acids.

The farmed fish samples showed higher group homogeneity, as can be expected for standardized feeding and growing conditions, than the samples of wild fish for which largest variability and more outliers are observed as a natural consequence of differences in sampling conditions (size, age, etc.). Although a relative compactness of points for farmed fish can be observed (cf. **Figure 2**) the discrimination according to geographic origin cannot be clearly achieved with PCA. The PCA of scaled buckets provides much better discrimination ability even in two dimensions.

The next step was to use PCA scores without pretreatment (raw) and scaled in an LDA. For LDA five fish groups were defined (Italy, Greece, Croatia, Turkey, and wild from the Mediterranean Sea). As the number of columns is bigger than the number of rows in the input matrix, LDA cannot be applied directly. A logical solution is to use the PCA scores in further discrimination studies. Two kinds of splits, random and Kennard–Stone, were applied for the data set. The results are summarized in **Table 1**.

The scores were used as a variable pool. The best discriminating variables have been selected in a forward stepwise manner either using all scores (64) or using a truncated set of scores (the first 32). The better results on the "truncated" test set might be surprising at first sight because the 32 scores are a subset of the 64 scores. However, the information loss is negligible, the removed scores comprising mainly the noise. Generally, the forward selection algorithm could select more scores from among 64 than from 32 (see the numbers in brackets in **Table 1**). Worse results were realized on the testing set; however, somewhat better results were realized on the training sets (raw data). Similar tendencies have earlier been found for olive oil discrimination using both total and truncated PCA

Table 1. Results of LDA of PC Scores[a]

	raw RS (64 scores)	raw KS (64 scores)	scaled RS (64 scores)	scaled KS (64 scores)
training	97.7 (12)	100.0 (16)	90.9 (10)	100 (11)
testing	33.3 (12)	77.3 (16)	71.4 (10)	59.1 (11)

	raw RS (32 scores)	raw KS (32 scores)	scaled RS (32 scores)	scaled KS (32 scores)
training	95.5 (9)	97.7 (13)	93.2 (8)	90.7 (6)
testing	66.7 (9)	81.8 (13)	76.2 (8)	86.4 (6)

[a] The Scores were derived from PCA of correlation matrices for raw or scaled spectra. Either all scores (64) were used as a Variable pool or the truncated set of scores was used (the first 32 scores). Correct classifications are in percent. The number of scores incorporated in the LDA model selected by forward selection algorithm can be found in brackets. RS = random selection; KS = Kennard–Stone selection.

Table 2. Results of LDA of PC Scores from *Mean Centered* Spectra[a]

	raw RS (64 scores)	raw KS (64 scores)	scaled RS (64 scores)	scaled KS (64 scores)
training	93.2 (10)	97.7 (10)	97.7 (5)	95.3 (10)
testing	66.6 (10)	90.9 (10)	90.5 (5)	95.5 (10)

	raw RS (32 scores)	raw KS (32 scores)	scaled RS (32 scores)	scaled KS (32 scores)
training	93.2 (10)	97.7 (10)	90.9 (4)	83.7 (7)
testing	66.7 (10)	90.9 (10)	90.5 (4)	90.9 (7)

[a] Either all scores (64) were used as a variable pool or the truncated set of scores was used (the first 32 scores). correct classifications are in percent. The number of scores incorporated in the LDA model selected by forward selection algorithm can be found in brackets. RS = random selection; KS = Kennard–Stone selection.

scores in LDA (*29*). A similar table can be constructed if scaling the buckets to total intensity as data pretreatment (**Table 2**).

Again, the same if not better results can be provided using the truncated pool of scores (32) than using the pool for all scores (64). The forward selection algorithm has included fewer variables (PC scores) in case of mean centering than in the case of scaling. This may indicate a slight overfit for scaled data. Keeping in mind the philosophy of parsimony, mean centering is advisable for the bucketed NMR data.

Although the numbers of scores involved are far from being close to the limit set by Defernez and Kemsley (*32*), some overfit can be suspected in the case of using a pool for all scores. It should be noted that the wild-farmed discrimination is close to perfect even if the classification of all five groups is not. The worse case scenario can be seen on **Figure 3**: the canonical scores (roots) are plotted against each other. Only two samples for wild fish were misclassified.

However, it is relatively easy to find boundaries for classification as shown in **Figure 3** and also in **Figure 4** in which the outlying observations for wild samples lay on the outskirt of the point cloud for farmed fish. **Figure 4** demonstrates an intentionally wrong data pretreatment (without mean centering) and not the best discrimination ability (score 1 not involved). Even in this case the wild and farmed status can be distinguished.

Moreover, the compactness of points for farmed fish can be observed in this projection, whereas a larger variability is observed for wild fish.

LDA of Raw and Scaled NMR Profiles. Results of five-group and four-group classifications are summarized below.

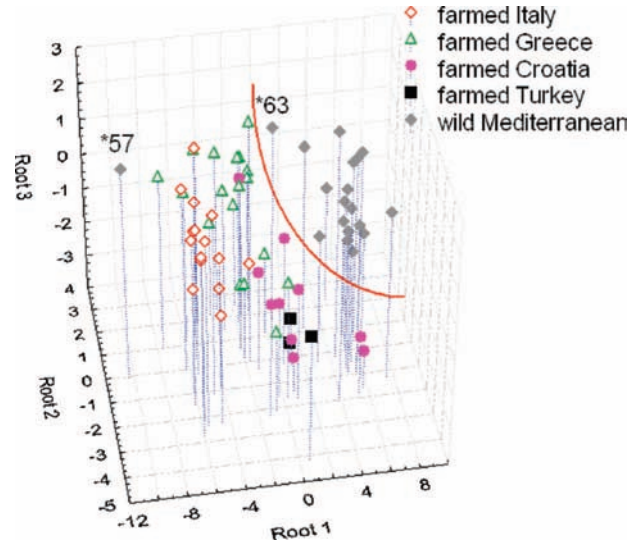

Figure 3. Results of stepwise LDA of PC scores calculated from scaled data. Three-dimensional plot of canonical scores (plot of root 1, root 2, and root 3). Points for misclassified wild samples are indicated by a star and a number.

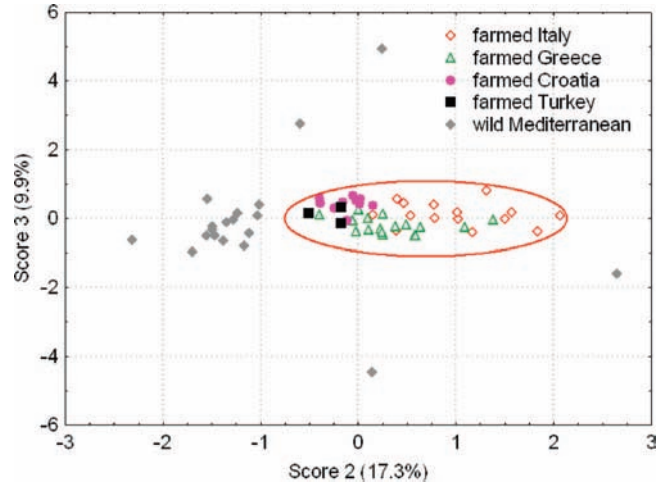

Figure 4. Results of PCA for scaled buckets (without mean centering). Score 2 is plotted against score 3.

Table 3. Results of LDA for Raw and Scaled NMR Profiles (without PCA)[a]

	raw RS	raw KS	scaled RS	scaled KS
Five-Group Classifications				
training	93.2 (6)	93.0 (10)	100.0 (8)	95.3 (8)
testing	76.2 (6)	90.9 (10)	90.5 (8)	100.0 (8)
Four-Group Classifications				
training	100.0 (8)	96.7 (7)	100.0 (6)	90.0 (4)
testing	88.2 (8)	93.8 (7)	70.6 (6)	93.8 (4)

[a] Correct classifications are in percent. The Number of chemical shifts selected by forward selection algorithm can be found in brackets. RS = random selection; KS = Kennard–Stone selection.

Four types of analyses have been made: raw buckets and scaled buckets have been used both with random and with Kennard–Stone splits of training and test sets. The correct classifications (in %) are summarized in **Table 3**.

From this table it can be concluded that, as expected, the scaling is a somewhat better option than analyzing the raw buckets. The average correct classifications (training + test sets) amounted from 88 to 97% depending on data (raw and scaled),

Table 4. Results of PNN Analysis for Bucketed and Scaled Mean Centered NMR Profiles[a]

	RS	KS
Wild vs Farmed		
training	95.4 (2)	100 (2)
testing	95.2 (2)	100 (2)
Country (Four Groups)		
training	93.2 (10)	97.7 (10)
testing	76.2 (10)	81.8 (10)

[a] Correct classifications are in percent. Numbers in brackets are PCA scores used. RS = random selection; KS = Kennard–Stone selection. Training = 44 samples; test = 21 samples.

on the split (random or Kennard–Stone), and on number of groups to be classified (four or five).

Table 4 summarizes some PNN (*33*) results for the randomly and Kennard–Stone selected training and test samples using 2 (wild vs farmed) or 10 (country of origin) PCA scores for input. These data further reinforce the findings that such class discriminations are readily feasible as a consequence of the use of such highly diagnostic methods as ^{1}H NMR. Although the prediction accuracy is similar to that of the LDA method, fewer scores are required for the wild versus farmed distinction.

^{1}H NMR lipid profiling combined with multivariate statistical analysis is able to discriminate wild and farmed fishes. Although the differentiation of geographic origins is somewhat worse, the potential is clear: by taking more samples and better design (balanced data sets, i.e., approximately equal number of samples for each country), the desired task (discrimination of samples according to geographical origin) can be made.

Wild and farmed discrimination can be found in the literature (*19*, *34*, *35*). The previous NMR approach reported in the literature is based on ^{2}H NMR, which requires time-consuming data acquisition. Moreover, D/H ratios have to be computed through curve fitting of the NMR signals, which is in contrast to the present approach, in which chemometrics is used to identify the relevant information in the spectra without any a priori knowledge.

The mentioned sources (*34*, *35*) extracted various phospholipid classes and fatty acids. After extraction and quantification PCA and quadratic discriminant analysis have provided the given separation using as low as four fatty acids (*34*). All labor-intensive extraction–quantification procedures became unnecessary when using the ^{1}H NMR profiling approach reported here. In the present case, the relevant information to address fish authenticity issues can be found in the main NMR signals of the lipid spectrum. This allows us to reasonably imagine a transfer of this method to a flow NMR version using a lower number of scans (*29*). Such a flow NMR method would drastically increase sample throughput.

^{1}H NMR profiling of lipid extracts combined with discriminant analysis holds all relevant information to unambiguously classify wild and farmed samples of Gilthead sea bream. Besides, the lipid profiles also allow extracting information that could be correlated to geographic origin. In this case, classification results must be interpreted carefully in the view of using them as robust information to establish the geographic origin of fish. It is indeed clearly established that the feeding process would affect significantly the lipid profile of fish. Various production practices, especially including the use of different feeds, may contribute to a compositional variability of the fish extracts. Then, it is important to underline the need to deeply assess the effect of feedings all across Mediterranean farms to unambiguously assign market fish to their geographic origin of production. These encouraging results need thus to be confirmed by the analysis of a greater number of samples including feeding trials.

ACKNOWLEDGMENT

We are grateful to Dr. Zhen-Yu Chen, associate editor of *J. Agric. Food Chem.*, and to three anonymous reviewers for their helpful suggestions for improvement of our manuscript.

Supporting Information Available: Additional data (sample, origin, production mode, month and year of sampling, total weight, and sampled from). This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Moretti, V. M.; Turchini, G. M.; Bellagamba, F.; Caprino, F. Traceability issues in fishery and aquaculture products. *Vet. Res. Commun.* **2003**, *27*, 497–505.

(2) Commission Regulation (EC) No. 2065/2001 of 22 October 2001 laying down detailed rules for the application of Council Regulation (EC) No. 104/2000 in regard to informing consumers about fishery and aquaculture products.

(3) Grigorakis, K.; Alexis, M. N.; Taylor, K. D. A.; Hole, M. Comparison of wild and cultured gilthead sea bream (*Sparus aurata*); composition, appearance and seasonal variations. *Int. J. Food Sci. Technol.* **2002**, *37*, 477–484.

(4) Martinez, I.; Aursand, M.; Erikson, U.; Singstad, T. E.; Veliyulin, E.; van der Zwaag, C. Destructive and non-destructive analytical techniques for authentication and composition analyses of foodstuffs. *Trends Food Sci. Technol.* **2003**, *14*, 489–498.

(5) Gunstone, F. D. ^{13}C NMR studies of mono-, di- and triacylglycerols leading to qualitative and semiquantitative information about mixtures of these glycerol esters. *Chem. Phys. Lipids* **1991**, *58*, 219–224.

(6) Gunstone, F. D. High resolution NMR studies of fish oils. *Chem. Phys. Lipids* **1991**, *59*, 83–89.

(7) Gunstone, F. D. High resolution ^{13}C NMR. A technique for the study of lipid structure and composition. *Prog. Lipid Res.* **1994**, *33*, 19–28.

(8) Gunstone, F. D. Information on the composition of fats from their high-resolution ^{13}C NMR spectra. *J. Am. Oil Chem. Soc.* **1993**, *70*, 361–366.

(9) Aursand, M.; Rainuzzo, J. R.; Grasdalen, H. Quantitative high-resolution ^{13}C and ^{1}H NMR of 3 fatty acids from white muscle of Atlantic salmon (*Salar salar*). *J. Am. Oil Chem. Soc.* **1993**, *70*, 971–981.

(10) Aursand, M.; Grasdalen, H. Interpretation of the ^{13}C NMR spectra of omega-3 fatty acids and lipid extracted from the white muscle of Atlantic salmon (*Salmo salar*). *Chem. Phys. Lipids* **1992**, *62*, 239–251.

(11) Aursand, M.; Jørgensen, L.; Grasdalen, H. Positional distribution of omega-3-fatty-acids in marine lipid triacylglycerols by high-resolution ^{13}C nuclear magnetic resonance spectroscopy. *J. Am. Oil Chem. Soc.* **1995**, *72*, 293–297.

(12) Medina, I.; Sacchi, R.; Aubourg, S. ^{13}C NMR monitoring of free fatty acid release after fish thermal processing. *J. Am. Oil Chem. Soc.* **1994**, *71*, 479–482.

(13) Sacchi, R.; Medina, I.; Aubourg, S. P.; Addeo, F.; Paolillo, L. Proton nuclear magnetic resonance rapid and structure-specific determination of ω-3 polyunsaturated fatty acids in fish lipids. *J. Am. Oil Chem. Soc.* **1993**, *70*, 225–228.

(14) Igarashi, T.; Aursand, M.; Hirata, Y.; Gribbestad, I. S.; Wada, S.; Nonaka, M. Nondestructive quantitative determination of docosahexaenoic acid and n-3 fatty acids in fish oils by high-resolution H-1 nuclear magnetic resonance spectroscopy. *J. Am. Oil Chem. Soc.* **2000**, *77*, 737–748.

(15) Saito, H.; Nakamura, K. Application of the NMR method to evaluate the oxidative deterioration of crude and stored fish oils. *Agric. Biol. Chem.* **1990**, *54*, 533–534.

(16) Sacco, A.; Brescia, M. A.; Liuzzi, V.; Reniero, F.; Guillou, C.; Ghelli, S.; van der Meer, P. Characterization of Italian olive oils based on analytical and nuclear magnetic resonance determinations. *J. Am. Oil Chem. Soc.* **2000**, *77*, 619–625.

(17) Fauhl, C.; Reniero, F.; Guillou, C. ^{1}H NMR as a tool for the analysis of mixtures of virgin olive oil with oils of different botanical origin. *Magn. Reson. Chem.* **2000**, *38*, 436–443.

(18) Aursand, M.; Mabon, F.; Martin, G. J. High-resolution H-1 and H-2 NMR spectroscopy of pure essential fatty acids for plants and animals. *Magn. Reson. Chem.* **1997**, *35*, 91–100.

(19) Aursand, M.; Mabon, F.; Martin, G. J. Characterization of farmed and wild salmon (*Salmo salar*) by combined use of compositional and isotopic analysis. *J. Am. Oil Chem. Soc.* **2000**, *77*, 659–666.

(20) Atherton, H. J.; Bailey, N. J.; Zhang, W.; Taylor, J.; Major, H.; Shockcor, J.; Clarke, K.; Griffin, J. L. A combined ^{1}H-NMR spectroscopy- and mass spectrometry-based metabolomic study of the PPAR-{alpha} null mutant mouse defines profound systemic changes in metabolism linked to the metabolic syndrome. *Physiol. Genom.* **2006**, *27*, 178–186.

(21) Bertram, H. C.; Kristensen, N. B.; Malmendal, A.; Nielsen, N. C.; Bro, R.; Andersen, H. J.; Harmon, D. L. A metabolomic investigation of splanchnic metabolism using ^{1}H NMR spectroscopy of bovine blood plasma. *Anal. Chim. Acta* **2005**, *536*, 1–6.

(22) Pears, M. R.; Rubtsov, D.; Mitchison, H. M.; Cooper, J. D.; Pearce, D. A.; Mortishire-Smith, R. J.; Griffin, J. L. Strategies for data analyses in a high resolution ^{1}H NMR based metabolomics study of a mouse model of Batten disease. *Metabolomics* **2007**, *3*, 121–136.

(23) Beger, R. D.; Schnackenberg, L. K.; Holland, R. D.; Li, D.; Dragan, Y. Metabonomic models of human pancreatic cancer using 1D proton NMR spectra of lipids in plasma. *Metabolomics* **2006**, *2*, 125–134.

(24) Bligh, E. G.; Dyer, W. Y. A rapid method to total lipid extraction and purification. *Can. J. Biochem. Physiol.* **1959**, *37*, 911–917.

(25) *STATISTICA* (data analysis software system), Version 6; StatSoft, Inc.: Tulsa, OK, U.S.A., 2004 (www.statsoft.com).

(26) Massart, D. L.; Vandeginste, B. G. M.; Buydens, L. M. C.; De Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J. Supervised pattern recognition. *Handbook of Chemometrics and Qualimetrics: Part B*; Elsevier: Amsterdam, The Netherlands, 1998; Chapter 33, p 207.

(27) Otto, M. *Chemometrics Statistics and Computer Application in Analytical Chemistry*; Wiley-VCH: Weinheim, Germany, 1999.

(28) Hastie, T.; Tibshirani, R.; Friedman, J. Linear Methods for Classification. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*; Springer: New York, 2001; Chapter 4, pp 79–114.

(29) Rezzi, S.; Axelson, D. E.; Héberger, K.; Reniero, F.; Mariani, C.; Guillou, C. Classification of olive oils using high throughput flow ^{1}H NMR fingerprinting with principal component analysis, linear discriminant analysis and probabilistic neural networks. *Anal. Chim. Acta* **2005**, *552*, 13–24.

(30) *Neuroshell Classifier/Neuroshell2/AI Trilogy*; Ward Systems Group, Inc.: Frederick, MD, 2004.

(31) Kennard, R. W.; Stone, L. A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11*, 137–148 (American Society for Quality Control, American Statistical Association).

(32) Defernez, M.; Kemsley, E. K. The use and misuse of chemometrics for treating classification problems. *Trends Anal. Chem.* **1997**, *16*, 216–221.

(33) Specht, D. F. Probabilistic neural networks, International Neural Network Society. *Neural Networks* **1990**, *3*, 109–118.

(34) Sérot, T.; Gandemer, G.; Demaimay, M. Lipid and fatty acid compositions of muscle from farmed and wild adult turbot. *Aquacult. Int.* **1998**, *6*, 331–343.

(35) Tritt, K. L.; O'Bara, C. J.; Wells, M. J. M. Chemometric discrimination among wild and cultured age-0 largemouth bass, black crappies, and white crappies based on fatty acid composition. *J. Agric. Food Chem.* **2005**, *53*, 5304–5312.

Received for review March 13, 2007. Revised manuscript received September 18, 2007. Accepted September 18, 2007. K.H. is thankful for the visiting scientist grant at the Joint Research Centre.

JF070736G