JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

ARTICLE

pubs.acs.org/JAFC

# $^1$H NMR Metabolite Fingerprint and Pattern Recognition of Mullet (*Mugil cephalus*) Bottarga

Emanuela Locci, Cristina Piras, Simone Mereu, Flaminia Cesare Marincola, and Paola Scano*

Dipartimento di Scienze Chimiche, Università degli Studi di Cagliari, Cittadella Universitaria di Monserrato, S.S. 554 — Bivio per Sestu, 09042 Cagliari, Italy

**ABSTRACT:** Nuclear magnetic resonance (NMR) spectroscopy combined with multivariate data analysis (MVA) was used to investigate the molecular components of the aqueous extract of samples of bottarga, that is, salted and dried mullet (*Mugil cephalus*) roe, manufactured in Sardinia (Italy) from mullets of known and unknown geographical provenience. Principal component analysis (PCA) applied to the processed $^1$H NMR spectra indicated that samples tend to cluster according to their geographical origin and also on the basis of storage and manufacturing procedures. The most important metabolites that characterized grouping of samples are the free amino acids methionine (Met), glutamate (Glu), histidine (His), phenylalanine (Phe), tyrosine (Tyr), and isoleucine (Ile); trimethylamine (TMA) and dimethylamine (DMA), both biomarkers of degradation; nucleotides and derivatives; choline (Cho) and phosphorylcholine (P-cho); and lactate (Lac).

**KEYWORDS:** bottarga, aqueous extract, metabolite profile, $^1$H NMR, MVA

## INTRODUCTION

The eviscerated roes of striped mullet (*Mugil cephalus*) are manufactured in several countries, and the salted and dried product can be found worldwide under different names and typologies. The Mediterranean island of Sardinia (Italy) has a long tradition in manufacturing mullet roes to obtain a product called "bottarga". In recent years, Sardinian bottarga, which is sold as whole ovaries ("in baffe") or grated in jars, has become so increasingly popular in international markets that mullets of the Mediterranean sea are not enough to satisfy the request of this product. As a result, the Sardinian producers must turn their attention to other fishing areas located in different regions of the globe for roe supplies. Indeed, raw roes are purchased from distributors located mainly in FAO 31, 34, and 41 fishing areas. Even if the raw material is not necessarily original to the island, Sardinian bottarga has its own peculiar rheological and organoleptic profile due to the skills of the local producers, inheritors of an ancient tradition in processing this delicacy. For this reason, Sardinian manufacturers of bottarga are requesting a Protected Geographical Indication (PGI) designation for this product.

The globalization of food markets and the relative ease with which food commodities are transported between countries increase the awareness of consumers about the origin of the foods they eat. As far as fishery products are concerned, European legislation establishes that the FAO area in which fish was caught should be part of the information available to consumers.[1] This applies also to processed products such as bottarga. It is therefore of great importance to be able to determine the geographical origin of fish, especially when used in preparing processes to perform authentication and/or traceability studies useful to enforce labeling regulations.

Traditional methods for species authentication of fish include DNA and protein analyses.[2] Recently, the analysis of metabolite profiles by high-resolution nuclear magnetic resonance (NMR) spectroscopy has been proposed as an alternative method for the authentication of seafood.[3−8] In the past two decades $^1$H NMR has proved to be a fast and versatile technique, useful both for compositional analysis and for rapid screening of food, allowing the detection of the major metabolites in a single spectrum, and, when associated with multivariate data analysis (MVA), can provide a suitable tool for comparing, discriminating, or classifying samples on the basis of their metabolic profile.[7,9] The free metabolite pool found in the aqueous phase of animal and vegetable matrices reflects the metabolic processes of the living organism, and it can be characteristic of individuals from a specific geographical area but can also reflect the chemical and physical transformations that can take place during storage of the raw material, manufacturing, and shelf life. It includes free amino acids, nucleotides and related compounds, organic bases, sugars, and others, whereas the molecular pool soluble in organic solvent includes fatty acids and related lipids. As for the majority of marine products, which are rich in health-beneficial $\omega$-3 fatty acids, most of the investigations on salted and dried mullet roes concern the lipid components and, particularly, the lipid classes and fatty acid composition (see ref 10 and literature cited therein). To the authors' knowledge, different from the lipid components, investigations on the low molecular weight compounds of bottarga are rare in the literature.[11]

The aim of this work was to evaluate whether the $^1$H NMR low molecular weight metabolite profile of bottarga can be considered a valid tool to characterize bottarga samples having different geographical origins and production processing protocols. For this purpose, we recorded the $^1$H NMR spectra of the aqueous extract of 25 samples of bottarga, manufactured in Sardinia from mullets of known and unknown geographical origin and commercialized either in baffe or grated in jars. Principal component analysis (PCA) was applied to the $^1$H NMR spectral data to

**Received:** April 1, 2011
**Revised:** August 8, 2011
**Accepted:** August 9, 2011
**Published:** August 09, 2011

**Table 1. Summary of the Studied Bottarga Samples (Year of Mullet Catching: 2007)**

| samples | *n* | typology | identification of the catch area[a] | catch area |
|---|---|---|---|---|
| 1−3 | 3 | in baffe | FAO 34 | central-eastern Atlantic |
| 4−9 | 6 | in baffe | FAO 41 | southwestern Atlantic |
| 10−12 | 3 | in baffe | FAO 31 | central-western Atlantic |
| 13−14 | 2 | grated in jar | FAO 37.1.3 | Mediterranean Sea |
| 15−25 | 11 | grated in jar | unknown | unknown |

[a] *FAO Yearbook. Fishery Statistics. Catches*; 2000; Vol. 86/1.

explore possible grouping of samples with common characteristics in terms of origin and processing of the raw material.

## MATERIALS AND METHODS

**Samples.** Twenty-five samples of bottarga manufactured in Sardinia were used: 12 samples were in baffe and 13 were industrially grated. The typology of samples and the geographical provenience of the raw roes are summarized in Table 1. Samples from mullets of known fishing areas were kindly gifted by a manufacturer located in Sardinia, and they belong to the same batch of suppliers. Raw roes from FAO 31, 34, and 41 were transported under ice and kept frozen (−20 °C) for not longer than 6 months until manufacturing. Samples from FAO 37.1.3 underwent curing procedures soon after evisceration; no freezing procedures were adopted. Samples purchased at a local supermarket were not labeled with geographical origin. The labels reported the ingredients as mullet roe and salt.

**Chemicals.** Deuterium oxide ($D_2O$, 99.9%) was purchased from Cambridge Isotope Laboratories Inc. (Andover, MA). Sodium 3-trimethylsilylpropionate-2,2,3,3,-$d_4$ (TSP, 98 atom % D), perchloric acid ($HClO_4$, 70%), and potassium hydroxide (KOH) were acquired from Sigma-Aldrich (Milan, Italy).

**Perchloric Acid Extraction.** Water-soluble metabolites were extracted using perchloric acid on the basis of the procedure previously described by Gribbestad et al.[5] Approximately 2 g of bottarga was pulverized in a mortar and transferred into a glass dish. A 4 mL solution of perchloric acid (7% in $D_2O$) was added and the mixture continuously stirred and warmed at 50 °C until a paste consistency was obtained. The homogenate was centrifuged at 4000 rpm for 10 min at 4 °C. Then, the supernatant was adjusted to pH 7.8 with 9 M KOH in $D_2O$ and centrifuged again to remove the potassium perchlorate. The final extract was lyophilized and stored at −20 °C until analyzed. Before NMR analysis, each sample was redissolved in 1 mL of $D_2O$, and an aliquot of 600 $\mu$L was transferred into a 5 mm tube to which 50 $\mu$L of TSP/$D_2O$ solution (0.80 mM final TSP concentration) was added as internal standard. All of the extractions were performed in duplicate.

**$^1$H NMR Spectroscopy.** $^1$H NMR experiments were carried out on a Varian Unity Inova 400 spectrometer operating at 399.94 MHz. Spectra were recorded at 298 K with a spectral width of 5624 Hz, a 90° pulse of 7.5 $\mu$s, an acquisition time of 3 s, a relaxation delay of 25 s, and 64 scans. The residual water signal was suppressed by applying a presaturation technique with low-power radiofrequency irradiation for 1.5 s. The FIDs were multiplied by an exponential weighting function equivalent to a line broadening of 0.3 Hz prior to Fourier transformation. Chemical shifts were referred to the TSP single resonance at 0.00 ppm.

2D NMR $^1$H−$^1$H COSY spectra were acquired with a spectral width of 4423 Hz in both dimensions, 2048 data points, and 512 increments with 48 transients per increment. 2D NMR $^1$H−$^1$H TOCSY spectra were acquired in phase sensitive mode with a size and number of data points similar to those of the COSY and a mixing time of 150 ms.

**Data Processing and Multivariate Data Analysis.** The $^1$H NMR spectra were segmented in 191 spectral domains of 0.04 ppm width (bins) by selecting the regions 8.50−5.10 and 4.54−0.80 ppm. The spectral region between 4.54 and 5.10 ppm was excluded from statistical analysis to remove the effect of the presaturation of the water residual resonance. Bucketing was performed by MestReNova.[12] Normalization has been applied to each sample data set to minimize the effects of variable dilution of the samples. The final data set consisted of a 25 × 191 matrix, in which rows represented samples and columns the normalized area of each bin. The generated file was imported into the SIMCA-P+ program[13] and submitted to mean-centering and autoscaling before PCA was performed.[14]

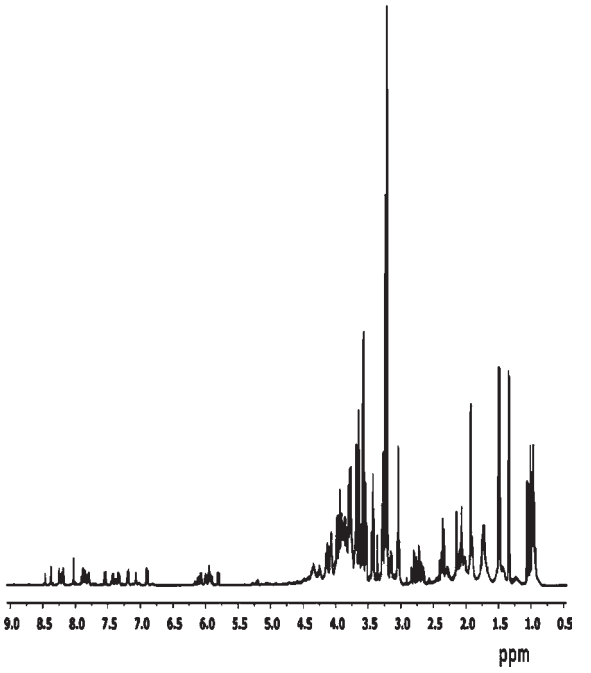


**Figure 1.** Representative $^1$H NMR spectrum of a bottarga aqueous extract in $D_2O$.

## RESULTS AND DISCUSSION

**$^1$H NMR Spectra.** The $^1$H NMR spectral profiles of bottarga aqueous extracts were similar among the examined samples, although changes in the relative intensities of some resonances were observed. Figure 1 shows a representative $^1$H NMR spectrum, and in Figure 2 the main assignments in the different spectral regions are reported.

As reported in Table 2, 42 compounds were identified on the basis of data published in the literature,[2,5,6,15,16] by performing 2D NMR experiments (COSY and TOCSY), and by recording spectra of standard compounds. In some cases, validation of the peak attribution was achieved by adding standard compounds directly to the sample solution and recording the NMR spectrum again under the same conditions.

The high-field region of the NMR spectrum (0.8−3.0 ppm) shows signals arising from aliphatic groups of free amino acids and organic acids. In particular, signals representing leucine (Leu), isoleucine (Ile), valine (Val), alanine (Ala), lysine (Lys), proline (Pro), and methionine (Met) were identified. The predominant organic acid identified is lactic acid (Lac). In addition, acetic (Ace), malic (Mal), and succinic (Suc) acids were also present. Small singlets at 2.91 and 2.73 ppm were ascribed to trimethylamine (TMA) and dimethylamine (DMA), respectively.

In the midfield region of the spectrum (3.0−5.5 ppm) the main contributions arise from the strongly overlapped signals of

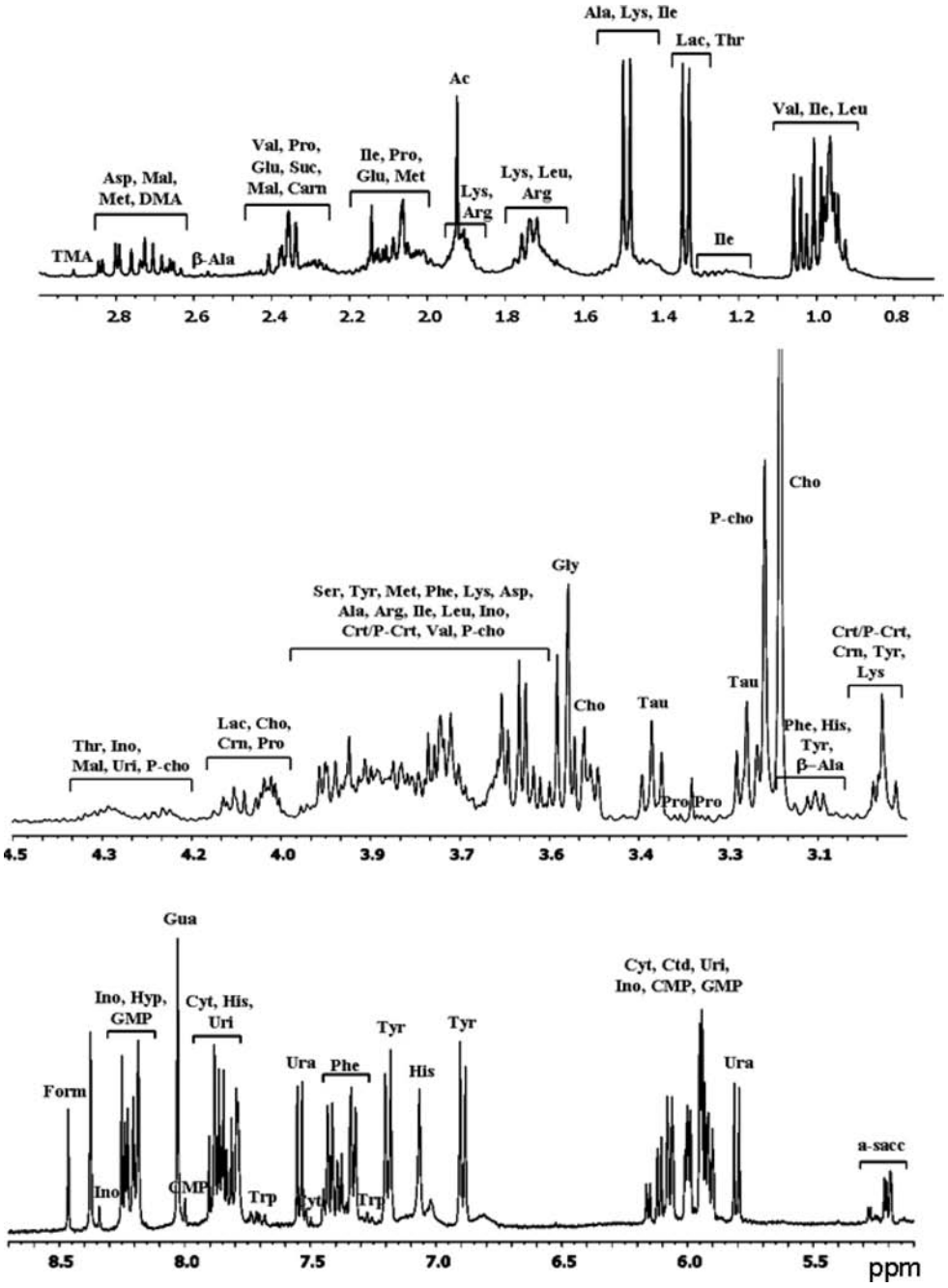


**Figure 2.** Spectral area assignments of a representative $^{1}$H NMR spectrum of a bottarga aqueous extract in $D_2O$.

the α-protons of the free amino acids and from saccharides. Moreover, the intense peaks at 3.21 and 3.24 ppm were attributed to choline (Cho) and phosphorylcholine (P-cho). Creatine (Crt) and/or phosphocreatine (P-crt) and creatinine (Crn) were also identified. Furthermore, the doublets at 5.18−5.22 ppm account for the anomeric proton of α-saccharides (a-sacc).

The signals in the low-field region (5.5−8 ppm) were assigned to the aromatic amino acids tyrosine (Tyr), phenylalanine (Phe), tryptophan (Trp), histidine (His), nucleobases (cytosine (Cyt), uracil (Ura)), nucleosides (cytidine (Ctd), uridine (Uri), inosine (Ino), guanosine (Gua)), and nucleotides. The two singlets at 8.18 and 8.21 ppm are indicative of the presence of hypoxanthine (Hyp). Formic acid (Form) resonates at 8.46 ppm.

Among the identified metabolites there are nutrients such as taurine (Tau) and carnitine (Carn), fish taste-active amino acids (i.e., Glu, Met, Gly, Ala[17,18]), nucleotide derivatives, preservatives (Lac and Mal), and biomarkers typically used for assessing fish quality. In particular, TMA accumulates in spoiling fish as a result of bacterial reduction of TMAO, and DMA, its counterpart, is diagnostic of freezing processes,[7] whereas the accumulation of Hyp and Ino is directly related to the degradation of adenosine triphosphate (ATP). The presence of these biomarkers in our samples can be ascribed to a series of chemical and physical transformations typically occurring in a biological matrix excised from the living organism, which can modify the original metabolic fingerprint. This is even more important when one is

**Table 2. Proton Chemical Shifts of the Metabolites Identified in the Aqueous Extract of Bottarga**

| compound | group | $^1H^a$ (ppm) | $^1H$ multiplicity$^b$ | correlation$^c$ (ppm) |
|---|---|---|---|---|
| acetate (Ace) | $\beta CH_3$ | 1.92 | s | |
| alanine (Ala) | $\alpha CH$ | 3.80 | q | 1.49 (C) |
| | $\beta CH_3$ | 1.49 | d | 3.80 (C) |
| $\beta$-alanine (b-ala) | $\alpha CH_2$ | 2.56 | t | 3.19 (C) |
| | $\beta CH_2$ | 3.19 | t | 2.56 (C) |
| arginine (Arg) | $\alpha CH$ | 3.77 | t | 1.92 (C) |
| | $\beta CH_2$ | 1.92 | m | 3.77, 1.70 (C) |
| | $\gamma CH_2$ | 1.70 | m | 1.92, 3.25 (C) |
| | $\delta CH_2$ | 3.25 | t | 1.70 (C) |
| aspartate (Asp) | $\alpha CH$ | 3.91 | dd | 2.68, 2.80 (C) |
| | $\beta CH$ | 2.68 | dd | 2.80, 3.91 (C) |
| | $\beta' CH$ | 2.80 | dd | 2.68, 3.91 (C) |
| carnitine (Carn) | $\alpha CH$ | 2.43 | m | 4.60 (C) |
| | $\beta CH$ | 4.60 | m | 2.43 (C), 3.49 (C) |
| | $\gamma CH$ | 3.49 | m | 4.60 (C) |
| | $N-(CH_3)_3^+$ | 3.24 | s | |
| choline (Cho) | $N-(CH_3)_3^+$ | 3.21 | s | |
| | $N-CH_2$ | 4.07 | m | 3.52 (C) |
| | $O-CH_2$ | 3.52 | | 4.07 (C) |
| creatine/phosphocreatine (Crt/P-Crt) | $N-CH_3$ | 3.04 | s | 3.93 (T) |
| | $N-CH_2$ | 3.93 | s | |
| creatinine (Crn) | $N-CH_3$ | 3.06 | s | |
| | $N-CH_2$ | 4.06 | s | |
| cytidine (Ctd) | C1′H, ribose | 5.91 | d | |
| | C5H | 6.07 | d | 7.85 (C) |
| | C6H | 7.85 | d | 6.07 (C) |
| cytosine (Cyt) | C5H | 5.98 | d | 7.50 (C) |
| | C6H | 7.50 | d | 5.98 (C) |
| cytidine monophosphate (CMP) | C5H, ring | 6.12 | d | 8.08 (C) |
| | C6H, ring | 8.08 | d | 6.12(C) |
| dimethylamine (DMA) | $N(CH_3)_2$ | 2.73 | s | |
| formate (Form) | $HCOO^-$ | 8.46 | s | |
| glutamate (Glu) | $\alpha CH$ | 3.78 | t | 2.10 (C), 2.38(T) |
| | $\beta CH$ | 2.10 | m | 2.38 (C) |
| | $\beta' CH$ | 2.10 | m | 2.38 (C) |
| | $\gamma CH_2$ | 2.38 | t | 2.10 (C), 3.78 (T) |
| glycine (Gly) | $\alpha CH_2$ | 3.57 | s | |
| guanosine (Gua) | C8H, ring | 8.03 | s | |


dx.doi.org/10.1021/jf2012979 |J. Agric. Food Chem. 2011, 59, 9497–9505

**Table 2. Continued**

| compound | group | $^1H^a$ (ppm) | $^1H$ multiplicity$^b$ | correlation$^c$ (ppm) |
|---|---|---|---|---|
| guanosine monophosphate (GMP) | C1′H, ribose | 5.94 | d | 4.76 (C) |
| | C2′H, ribose | 4.76 | m | 5.94 (C) |
| | C8H, ring | 8.20 | s | |
| histidine (His) | C2H, ring | 7.79 | s | 7.07 (T) |
| | C4H, ring | 7.07 | s | 7.79 (T) |
| hypoxanthine (Hyp) | C2H, ring | 8.21 | s | |
| | C8H, ring | 8.18 | s | |
| inosine (Ino) | C1′H, ribose | 6.11 | d | 4.82 (C) |
| | C2′H, ribose | 4.82 | m | 6.11 (C) |
| | C3′H, ribose | 4.48 | m | 4.36 (C) |
| | C4′H, ribose | 4.36 | m | 4.48, 3.88 (C) |
| | C5′H, ribose | 3.88 | m | 4.36 (C) |
| | C2H, ring | 8.34 | s | |
| | C8H, ring | 8.23 | s | |
| isoleucine (Ile) | αCH | 3.68 | m | 1.99 (C) |
| | $\beta$CH | 1.99 | m | 1.02, 1.27, 3.68 (C) |
| | $\gamma$CH | 1.48 | m | 0.94, 1.27 (C) |
| | $\gamma'$CH | 1.27 | m | 0.94, 1.48, 1.99 (C) |
| | $\gamma'CH_3$ | 1.02 | d | 1.99 (C) |
| | $\delta CH_3$ | 0.94 | t | 1.27, 1.44 (C) |
| lactate (Lac) | $\beta CH_3$ | 1.34 | d | 4.12 (C) |
| | $\alpha CH_2$ | 4.12 | q | 1.34 (C) |
| leucine (Leu) | αCH | 3.75 | t | 1.70 (C) |
| | $\beta CH_2$ | 1.70 | m | 3.75 (C) |
| | $\gamma$CH | 1.72 | m | 0.96 (C) |
| | $\delta CH_3, \delta' CH_3$ | 0.96 | d | 1.72 (C) |
| lysine (Lys) | αCH | 3.77 | t | 1.92 (C) |
| | $\beta CH_2$ | 1.92 | m | 1.43 (C), 3.03 (T), 3.77 (C) |
| | $\gamma CH_2$ | 1.43 | m | 1.73 (C), 1.92 (C), 3.03 (T) |
| | $\delta CH_2$ | 1.73 | m | 1.43, 3.03 (C) |
| | $\varepsilon CH_2$ | 3.03 | t | 1.43 (T), 1.73 (C), 1.92 (T) |
| malate (Mal) | αCH | 4.33 | dd | 2.38 (C) |
| | $\beta$CH | 2.38 | dd | 2.70, 4.33 (C) |
| | $\beta'$CH | 2.70 | dd | 2.38 (C) |
| methionine (Met) | αCH | 3.86 | t | 2.19 |
| | $\beta CH_2$ | 2.19 | m | 2.65, 3.86 (C) |
| | $\gamma CH_2$ | 2.65 | t | 2.19 (C) |
| | S-$CH_3$ | 2.14 | s | |
| phenylalanine (Phe) | αCH | 3.99 | dd | 3.13, 3.29 (C) |
| | $\beta$CH | 3.29 | dd | 3.99 (C) |
| | $\beta'$CH | 3.13 | dd | 3.99 (C) |
| | C2,6H, ring | 7.42 | m | |
| | C3,5H, ring | 7.42 | m | 7.33 (C) |
| | C4H, ring | 7.33 | m | 7.42 (C) |


dx.doi.org/10.1021/jf2012979 |*J. Agric. Food Chem.* 2011, 59, 9497–9505

**Table 2. Continued**

| compound | group | $^1H^a$ (ppm) | $^1H$ multiplicity$^b$ | correlation$^c$ (ppm) |
|---|---|---|---|---|
| phosphorylcholine (P-cho) | $N-(CH_3)_3^+$ | 3.24 | s | |
| | $N-CH_2$ | 4.32 | m | 3.68 (C) |
| | $O-CH_2$ | 3.68 | m | 4.32 (C) |
| proline (Pro) | αCH | 4.15 | t | 2.04 (C), 2.35 (C), 2.01 (T) |
| | βCH | 2.35 | m | 4.15 (C) |
| | β′CH | 2.04 | m | 4.15 (C) |
| | $\gamma CH_2$ | 2.01 | m | 3.38 (C), 3.40 (C) |
| | δCH | 3.38 | t | 2.04 (T), 2.35 (T) |
| | δ′CH | 3.40 | t | 2.04 (T), 2.35 (T) |
| serine (Ser) | αCH | 3.86 | dd | 3.96 (C) |
| | βCH | 3.96 | dd | 3.86 (C) |
| succinate (Suc) | $\alpha,\beta CH_2$ | 2.41 | s | |
| taurine (Tau) | $N-CH_2$ | 3.27 | t | 3.43 (C) |
| | $S-CH_2$ | 3.43 | t | 3.27 (C) |
| trimethylamine (TMA) | $N-(CH_3)_3$ | 2.91 | s | |
| trimethylamine oxide (TMAO) | $O-N-(CH_3)_3$ | 3.24 | s | |
| threonine (Thr) | αCH | 3.60 | d | 4.27 (C) |
| | βCH | 4.27 | m | 1.34 (C), 3.60 (T) |
| | $\gamma CH_3$ | 1.34 | d | 4.27 (C) |
| tryptophan (Trp) | C4H, ring | 7.72 | d | 7.19 (C), 7.27 (T), 7.54 (T) |
| | C5H, ring | 7.19 | t | 7.27 (C), 7.72 (C) |
| | C6H, ring | 7.27 | t | 7.19 (C), 7.54 (C), 7.72 (T) |
| | C7H, ring | 7.53 | d | 7.27 (C), 7.72 (T) |
| tyrosine (Tyr) | αCH | 3.94 | dd | 3.06 (C) |
| | βCH | 3.19 | dd | 3.06 (C) |
| | β′CH | 3.06 | dd | 3.19, 3.94 (C) |
| | C2,6H, ring | 6.89 | d | 7.19 (C) |
| | C3,5H, ring | 7.19 | d | 6.89 (C) |
| uracil (Ura) | C5H, ring | 5.81 | d | 7.54 (C) |
| | C6H, ring | 7.54 | d | 5.81 (C) |
| uridine (Urd) | CH-1′ ribose | 5.94 | d | 4.39 (C) |
| | CH-2′ ribose | 4.39 | m | 5.94 (C) |
| | C5H, ring | 5.91 | d | 7.89 (C) |
| | C6H, ring | 7.89 | d | 5.91 (C) |
| valine (Val) | αCH | 3.62 | d | 2.28 (C) |
| | βCH | 2.28 | m | 0.99, 1.04, 3.62 (C) |
| | $\gamma' CH_3$ | 1.04 | d | 0.99, 2.28 (C) |
| | $\gamma CH_3$ | 0.99 | d | 1.04, 2.28 (C) |

[a] $^1H$ chemical shifts are reported with respect to TSP signal (0.00 ppm). [b] Multiplicity definitions: s, singlet; d, doublet; t, triplet; q, quartet; dd, doublet of doublets; m, multiplet. [c] Experiment legend: C, COSY; T, TOCSY.

dealing with foodstuffs and, in particular, with fish raw materials transported across continents, stored at low temperature, as in the case of mullet roes.

**Multivariate Data Analysis.** Visual analysis of the NMR spectra does not show obvious relationships between the intensities of certain signals and the geographical origin of mullets. The

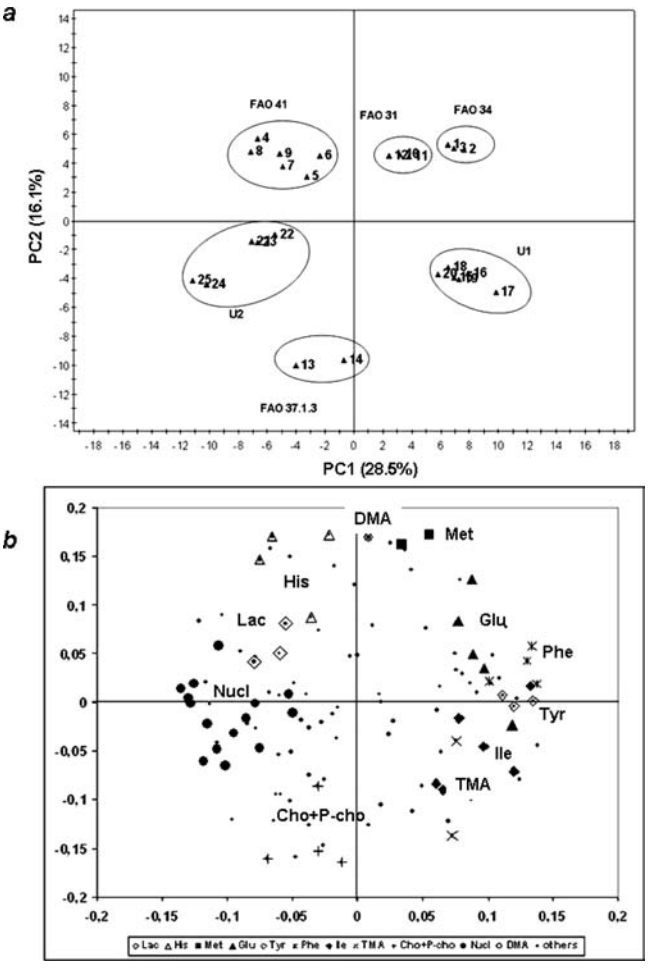


**Figure 3.** PC1 versus PC2 (a) scores plot (ellipses are arbitrarily drawn to group samples; the explained variance is given is parentheses) and (b) loadings plot (with the most significant metabolites highlighted) of PCA applied to the $^{1}$H NMR spectral data of the bottarga aqueous extracts. "Nucl" includes CMP, GMP, Ctd, Uri, Ino, Gua, Hyp, Ura, and Cyt; "others" refers to all other loadings.

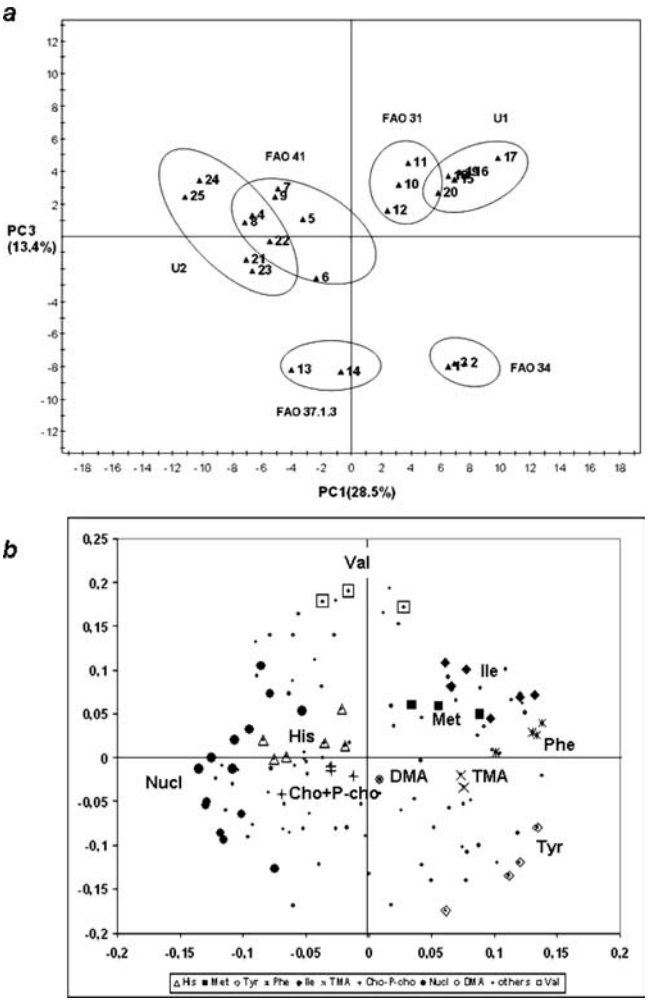


**Figure 4.** PC1 versus PC3 (a) scores plot (ellipses are arbitrarily drawn to group samples; the explained variance is given in parentheses) and (b) loadings plot (with the most significant metabolites highlighted) of PCA applied to the $^{1}$H NMR spectral data of the bottarga aqueous extracts. "Nucl" includes CMP, GMP, Ctd, Uri, Ino, Gua, Hyp, Ura, and Cyt; "others" refers to all the other loadings.

use for MVA tools is therefore warranted. An exploratory analysis of the data set was carried out by applying a PCA. PCA is an unsupervised technique and requires no information about class membership; it looks just for inherent variation in the data set. Application of PCA on the 25 samples of bottarga under investigation allowed us to reduce the large $^{1}$H NMR data set to three principal components with 28.5, 16.1, and 13.4% of total variance explained. As can be seen in Figure 3a, samples with the same known geographical origin are clustered in the same region of the scores plot of PC1 versus PC2 and distributed along the PC1, whereas samples from FAO 37.1.3 (Mediterranean sea, east Sardinia) form a cluster clearly separated from the others by PC2. PC2 contributes also to separate samples in baffe (upper side of the plot) from those that underwent industrial grating (lower side of the plot). Furthermore, samples of unknown geographical origin can be grouped in two clusters: U1 (samples 15−20) on the positive side of PC1 and U2 (samples 20−25) on the negative side of PC1. Separation of bottarga according to geographical origin of mullet is still visible in the scores plot of PC1 versus PC3 (Figure 4a). Interestingly, here samples of cluster U2 overlap with FAO 41 and those of U1 with FAO 31.

The explanation of what each PC represents in relation to the original measurements can be assessed by analyzing the coefficients by which the original variables (in our case, the spectral bins) must be multiplied to obtain the PC, that is, the "loadings". Examination of the loadings plot enables us to determine the variables with the highest impact on the variance and, thus, to identify the metabolites that contributed most to the cluster separation. However, for our data set, the latter step is not straightforward. Indeed, because the $^{1}$H NMR spectra of extracts of bottarga are very crowded with signals, several variables hold contributions from more than one metabolite, and, in some cases, the same metabolite contributes to more than one variable; thus, the resulting loadings plots are difficult to analyze. For these reasons, we simplified the appearance of the PC1 versus PC2 loadings plot, underlining only those bins composed predominantly by one metabolite and lying in the extremity of the axis, that is, giving the highest contribution to the PCs (Figure 3b). Here, contributions of metabolites to the sample clustering can be estimated considering the following rule: the position of a sample in a given direction in the score plot is influenced by the metabolites lying in the same direction in the corresponding loadings plot.


dx.doi.org/10.1021/jf2012979 |J. Agric. Food Chem. 2011, 59, 9497–9505

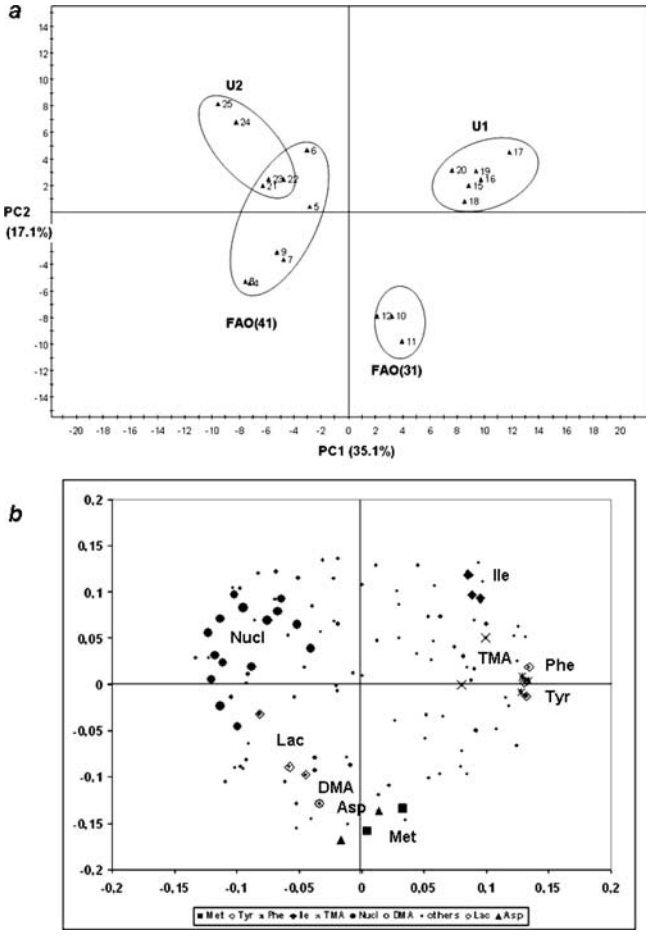


**Figure 5.** PC1 versus PC2 (a) scores plot (ellipses are arbitrarily drawn to group samples; the explained variance is given in parentheses) of PCA applied to the $^{1}$H NMR spectral data of the bottarga aqueous extracts (samples from FAO 37.1.3 and FAO 34 have been removed from the original matrix) and (b) loadings plot (with the most significant metabolites highlighted). "Nucl" includes CMP, GMP, Ctd, Uri, Ino, Gua, Hyp, Ura, and Cyt; "others" refers to all the other loadings.

Grouping of samples of different origin is mainly along PC1, and metabolites that characterize this first PC are Phe, Tyr, and, on the opposite side, nucleotides and derivatives (CMP, GMP, Ctd, Uri, Ino, Gua, Hyp, Ura, and Cyt). It is interesting to note that these latter, probably from disruption of nuclei acids, are strongly correlated.

Along PC2 we found that DMA, a biomarker of freezing, is placed on the opposite side of samples from FAO 37.1.3; this observation is in agreement with the fact that samples from FAO 37.1.3 were not keep frozen, as reported under Materials and Methods. Furthermore, Cho and P-cho, derivatives of phosphocholine (PC) in which mullet roes are rich,[19] are placed along PC2, on the same side of the grated samples. These compounds can be considered as biomarkers of hydrolytic mechanisms on PC, caused by the salting and drying procedures and probably exacerbated by the industrial grating. Samples from FAO 34 are more characterized by Met and Glu, which are taste-active compounds in fish derivatives,[17,18] whereas samples from FAO 41 are characterized by His and Lac. With regard to the samples of unknown geographical origin, the U1 cluster is rich in Ile and TMA; the latter can be linked with a more marked degradation of the raw food matrix (the uncured roes), because the salinity of the final product (bottarga) does not allow extensive bacterial growth. A comparative analysis of the full-resolution spectra confirmed the above-reported observations regarding metabolites characterizing sample grouping. The same procedure was applied to the PC1 versus PC3 loadings plot shown in Figure 4b. Opposite to PC2, in the third PC, loading values of DMA, Cho and P-cho, and His are very low, whereas Val is the metabolite that most contributes to the separation of the samples along PC3.

The results of the PCAs (Figures 3 and 4) suggested that the application of MVA to the $^{1}$H NMR data allows characterization of bottarga according to the geographical origin of the raw material and clustering of samples on the basis of their history and treatments. These encouraging results led us to investigate the possibility of classifying samples of unknown origin as belonging to a specific geographical region. To this goal, the best analytical tool would be a discriminant analysis, such as partial least-squares discriminant analysis; however, our restricted and inhomogeneous sampling, in terms of numbers of samples in each group, did not allow its performance. Therefore, we carried out a PCA removing samples from FAO 37.1.3 and FAO 34, because, by analysis of Figures 3 and 4, they do not seem to have any common characteristic with the unknown samples. The resulting scores plot, reported in Figure 5a, shows that samples 21−23 of unknown geographical origin are grouped with samples from FAO 41, whereas all other samples (15−20, 24, 25) show no similarities either with FAO 41 or with FAO 31. The explanation of this result might be that the unclassified samples do not belong to any of the fishing areas here studied and/or they are blends of roes of mullets caught in different geographic areas. The corresponding loadings plot, reported in Figure 5b, shows a trend similar to that of Figure 3b; that is, nucleosides and derivatives are strongly inversely correlated to Tyr and Phe. Moreover, they mostly contributed to the overlap of U2 group with FAO 41 samples. TMA and Ile characterize U1 samples, whereas DMA, Met, and Asp characterize samples from FAO 31.

In conclusion, the results of the present study, although a larger data set is warranted, demonstrated that the application of MVA to the $^{1}$H NMR spectral data allows bottarga to be characterized according to the geographical origin of the raw material and storage and manufacturing procedures. In fact, in the space spanned by the first three PCs, samples tend to cluster on the basis of their geographical origin and sample history. Among the molecular compounds unambiguously identified, Phe, Tyr, and nucleoside derivatives, followed by DMA, Cho and P-cho (FAO 37), His and Lac (FAO 41), Asp, Met, and Glu (FAO 31 and 34), and Ile and TMA (U1), are the metabolites that principally characterize the groups of samples. Among samples of unknown origin, the tendency of U2 to cluster together with samples of FAO 41 can be ascribed, in terms of metabolites, to the nucleoside derivatives and their inverse correlation with Phe and Tyr.

## AUTHOR INFORMATION

### Corresponding Author

*Phone: +39 070 675 4391. Fax: +39 070 675 4388. E-mail: scano@unica.it.

## REFERENCES

(1) Commission Regulation (EC) No 2065/2001, Council Regulation (EC) No 104/2000. *Off. J. Eur. Communities* **2001**, *L278*, 0006− 0008.

(2) 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.

(3) Standal, I. B.; Gribbestad, I. S.; Bathen, T. F.; Aursand, M.; Martinez, M. I. G. Low molecular weight metabolites in white muscle from cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) analyzed by high resolution $^{1}$H NMR spectroscopy. *Magn. Reson. Food Sci.* **2007**, 55–62.

(4) Aursand, M.; Standal, I. B.; Praël, A.; McEvoy, L.; Irvine, J.; Axelson, D. E. $^{13}$C NMR pattern recognition techniques for the classification of Atlantic salmon (*Salmo salar* L.) according to their wild, farmed, and geographical origin. *J. Agric. Food Chem.* **2009**, *57*, 3444–3451.

(5) Gribbestad, I. S.; Aursand, M.; Martinez, I. High-resolution $^{1}$H magnetic resonance spectroscopy of whole fish, fillets and extracts of farmed Atlantic salmon (*Salmo salar*) for quality assessment and compositional analyses. *Aquaculture* **2005**, *250*, 445–457.

(6) Mannina, L.; Sobolev, A. P.; Capitani, D.; Iaffaldano, N.; Rosato, M. P.; Ragni, P.; Reale, A.; Sorrentino, E.; D'Amico, I.; Coppola, R. NMR metabolic profiling of organic and aqueous sea bass extracts: implications in the discrimination of wild and cultured sea bass. *Talanta* **2008**, *77*, 433–444.

(7) Martinez, I.; Bathen, T.; Standal, I. B.; Halvorsen, J.; Aursand, M.; Gribbestad, I. S.; Axelson, D. E. Bioactive compounds in cod (*Gadus morhua*) products and suitability of $^{1}$H NMR metabolite profiling for classification of the products using multivariate data analyses. *J. Agric. Food Chem.* **2005**, *53*, 6889–6895.

(8) Savorani, F.; Picone, G.; Badiani, A.; Fagioli, P.; Capozzi, F.; Engelsen, S. B. Metabolic profiling and aquaculture differentiation of gilthead sea bream by $^{1}$H NMR metabonomics. *Food Chem.* **2010**, *120*, 907–914.

(9) Lee, E. J.; Shaykhutdinov, R.; Weljie, A. M.; Vogel, H. J.; Facchini, P. J.; Park, S. U.; Kim, Y. K.; Yang, T. J. Quality assessment of ginseng by $^{1}$H NMR metabolite fingerprinting and profiling analysis. *J. Agric. Food Chem.* **2009**, *57*, 7513–7522.

(10) Scano, P.; Rosa, A.; Mereu, S.; Piras, C.; Atzeri, A.; Dessì, M. A. Multivariate fatty acid and fatty alcohol profile of mullet bottarga. *Eur. J. Lipid Sci. Technol.* **2010**, *112*, 1369–1374.

(11) Chiou, T. K.; Konosu, S. Changes in extractive components during processing of dried mullet roe. *Nippon Suisan Gakkaishi* **1988**, *54*, 307–313.

(12) MestReNova (version 5.2.4) *Mestrelab R*esearch S.L.

(13) SIMCA-P+ (version 12.0) Umetrics, Umeå, Sweden.

(14) Eriksson, L.; Johansson, E.; Kettaneh-Wold, N.; Trygg, J.; Wikström, C.; Wold, S. In *Multi- and Megavariate Data Analysis Part I: Basic Principles and Applications*; Umetrics: Umeå, Sweden, 2006.

(15) Fan, T. Metabolite profiling by one- and two-dimensional analysis of complex mixtures. *Prog. NMR Spectrosc.* **1996**, *28*, 161–219.

(16) Seierstad, T.; Røe, K.; Sitter, B.; Halgunset, J.; Flatmark, K.; Ree, A. H.; Olsen, D. R.; Gribbestad, I. S.; Bathen, T. F. Principal component analysis for the comparison of metabolic profiles from human rectal cancer biopsies and colorectal xenografts using high-resolution magic angle spinning $^{1}$H magnetic resonance spectroscopy. *Mol. Cancer* **2008**, 7–33.

(17) Fuke, S.; Konosu, S. Taste-active components in some foods: a review of Japanese Research. *Physiol. Behav.* **1991**, *49*, 863–868.

(18) Hayashi, T.; Kohata, H.; Watanabe, E.; Toyama, K. Sensory study of flavour compound in extracts of salted salmon eggs (ikura). *J. Sci. Food Agric.* **1990**, *50*, 343–356.

(19) Scano, P.; Rosa, A.; Cesare-Marincola, F.; Locci, E.; Melis, M. P.; Dessì, M. A.; Lai, A. $^{13}$C NMR, GC and HPLC characterization of lipid components of the salted and dried mullet (*Mugil cephalus*) roe "bottarga". *Chem. Phys. Lipids* **2008**, *151*, 69–76.