Pyrolysis-mass spectrometry characterisation of ovine fat tissues according to diet

I. Sebastian, Christine Viallon, Pascal Tournayre, Philippe Berge, J.L. Berdague

Pyrolysis-mass spectrometry characterisation of ovine fat tissues according to diet

I. Sebastian, C. Viallon, P. Tournayre, P. Berge and J.-L. Berdagué[*]

Laboratoire Flaveur, Station de Recherches sur la Viande, INRA de Theix,
63122 Saint-Genès-Champanelle, France

Abstract. The lipids of subcutaneous fat tissues from 120 lambs were analysed by Curie point Pyrolysis - Mass Spectrometry (Py-MS) after extraction of the lipids with hexane. The animals came from six European countries with different production systems. The classification of the fat tissues according to the predominant type of feed the lambs consumed (milk, concentrate or pasture) was obtained by Discriminant Analysis (DAs) and Artificial Neural Networks (ANNs). The classification models were validated by "leave-one-out"-type cross-validations. The results showed that the Py-MS spectra of the fat tissue lipids permitted recognition of feed type. The percentages of correct classifications determined by cross-validation were 78 % by DA and 92 % by ANNs.

Keywords. Curie point Pyrolysis-Mass Spectrometry – lamb carcasses – type of diet – rapid classification.

Introduction

Increasing social demand has impelled the development of new methods for assuring the traceability of food products from source through processing. In the meat products sector the development of characterisation techniques is often problematic owing to heterogeneity of animal tissues and diversity of modes of production. The objective of this work was to test Curie point pyrolysis combined with mass spectrometry as a method for characterising commercial lamb carcasses derived from different production systems using three predominant types of animal feed; pasture, concentrate, and milk. For this purpose, samples of fat tissue, relatively homogeneous and easy to take, were collected from lamb carcasses.

The Curie point pyrolysis – mass spectrometry method was chosen as a rapid way to "fingerprint" the overall composition of the matrix of products. These "fingerprints" were obtained by filtering positive ions produced after thermal fragmentation (pyrolysis) and electron impact fragmentation of the matrices studied. Combined with real-time data analysis, this method has proved to be a powerful characterisation tool, and has already helped to solve problems of classification and recognition in various agrifood sectors [1-4].

Materials and methods

Animal material

The study was conducted on 120 commercial lambs of 12 different breed types and grown under a wide range of conditions representative of the current European productions systems. Six European countries supplied the lambs: France, Greece, Iceland, Italy, Spain and the United Kingdom. These lambs were distributed into three groups according to the predominant type of feed consumed before slaughter: group C: Concentrate ($n = 40$), group P: Pasture ($n = 60$) and group M: Milk ($n = 20$). Table I presents the distribution of the animals according to type of feed, country of origin, breed type and sex condition. There was a large within-feed type variability, the lamb age ranging from 12 to 1 year.

Sampling procedure and storage

From each lamb carcass, 25 g of fat tissue were taken from the subcutaneous fat in the rump region. Samples were wrapped in aluminium foil, vacuum-packed in polyethylene bags and stored at –20 °C until analysis.

Sample preparation for Py-MS

Extraction of lipids

To obtain homogeneous samples representative of fat tissues, analyses were carried out on a solvent-extracted lipid fraction. The extraction (3 h at 20 °C) was carried out on 0.5 g of tissue placed in a 20 ml glass tube containing 15 ml of hexane.

Deposition of samples on pyrolysis foil

This operation consisted in depositing a few microlitres of liquid sample on a clean iron-nickel foil (SS Scientific Ltd;

Original articles

Table 1. Distribution of the lambs according to the predominant type of feed consumed before slaughter, country of origin, type of breed and sex condition.

Type of feed	Country (number)	Breed	Sex condition
Group C (Concentrate, $n = 40$)	France ($n = 10$)	Lacaune	female
	Greece ($n = 10$)	Karagouniko	entire male
	Spain ($n = 10$)	Rasa Aragonesa	entire male
	Italy ($n = 10$)	Appenninica	entire male
Group M (Milk, $n = 20$)	Greece ($n = 10$)	Karagouniko	entire male
	Spain ($n = 10$)	Churra	entire male
Group P (Pasture, $n = 60$)	France ($n = 10$)	Texel, Ile-de-France	female
	Iceland ($n = 10$)	Icelandic	entire male
	Iceland ($n = 10$)	Icelandic	female
	Italy ($n = 10$)	Bergamasca	castrated male
	UK ($n = 10$)	Suffolk x Mule	castrated male
	UK ($n = 10$)	Welsh Mountain	entire male

Hellingly BN27 4DU, UK). The foil was then dried in an oven and placed in silica pyrolysis tube (Ets. Maillière; 63170 Aubière, France) using a stainless steel depth gauge to position it 10 mm from the tube outlet. For air-tightness during pyrolysis, a Viton O-ring (SS Scientific) was placed around the tube. Different methods for direct and indirect deposition of tissue lipids were tested to obtain a signal (Total Ion Count or TIC) between 1 and 2.10^6 counts. By direct deposition, 1 µl of lipids in hexane solution (*i.e.* an average lipid mass of approximately 30 µg) was placed on the iron-nickel foil. This deposit was then oven dried for 2 min at 100 °C to eliminate residual hexane before pyrolysis. For the indirect deposits, 2 µl of an aqueous suspension of talc, calcium carbonate, silica or magnesia, or solution of sucrose and ovalbumin was first deposited on the foil. After drying the deposit for 7 minutes at 100 °C, 1 µl of the lipids in hexane was placed on the dried spot and the foil was dried again for 2 min at 100 °C.

Pyrolysis - mass spectrometry

The analyses were performed over a period of one week, taking care to distribute the samples of the three feed types evenly over the analysis days. For each lipid sample three consecutive replicate pyrolysis runs were carried out. The pyrolysis mass spectrometer used was a Cp-PyMS RaPyD 400 Horizon Instrument. The tube carrying the foil was heated at 530 °C for 3 s during the pyrolysis, with a temperature rise time of 0.6 s. This pyrolysis temperature was chosen because it gave a balanced fragmentation of glucidic, lipid and protein fractions. The pyrolysate then entered an expansion chamber (heated at 160 °C), whence it diffused through a molecular beam tube (170 °C) to the ionisation chamber of the mass spectrometer (180 °C). To minimise secondary fragmentation of the pyrolysate, the ionisation method used was low-voltage electron impact ionisation (34 eV). Non-ionised molecules were retained on a cold trap cooled by liquid nitrogen. The ionised fragments were focused by the electrostatic lens of a set of source electrodes, accelerated, and directed into a quadrupole mass filter. The ions were separated by the quadrupole according to their mass-to-charge ratio, and detected and amplified with an electron multiplier. The mass spectrometer scanned the ionised pyrolysate 65 times during the pyrolysis. Data were collected over the range $49 < m/z < 241$ atomic mass units (a.m.u.). Spectral information on ion counts for the individual masses scanned, and the total ion count for each analysed sample were recorded. The mass spectrometer was calibrated using the chemical standard perfluorokerosene (ref. number 32,835-9; Sigma - Aldrich).

Data analysis

The analyses yielded a primary data matrix, M0, containing 191 variables or fragments in the range $49 < m/z < 241$ Dalton and 360 observations or analyses (120 samples $\times$ 3 replicates per sample).

Pre-treatment of data

Before constructing fat tissue classification models, a sequence of four steps, described in figure 2, was necessary, namely normalisation, elimination of aberrant data, elimination of drift due to the "analysis day" effect, and median filtration of replicates. At each of these steps, principal component analysis (PCA) was used to visualise the structure of the data set.

The normalisation of the data involved expressing, for each mass spectrum, the individual mass fragments F_i in percentages of total ion count. The procedure, used classically for the analysis of Py-MS data, corrects for differences in the mass of the sample deposited manually on the metallic foil. As the total abundance of the spectra is very closely correlated with the mass pyrolysed, normalisation allows the shape of the signal to be considered rather than its intensity.

Matrix M1 containing 191 normalised mass fragments and 360 observations was thus obtained.

The elimination of aberrant data was performed from a PCA of the normalised data. All data points more than three standard deviations away from the mean of the first three principal components were thus eliminated. By this criterion, one fat sample was found to be aberrant for all three pyrolysis replicates, and two other samples for two of the three pyrolysis replicates. In all, nine mass spectra corresponding to all three replicates of all three aberrant samples were eliminated from the data set. Matrix M2 containing 191 normalised mass fragments and 351 observations was thus obtained from matrix M1.

The elimination of the drift linked to the "analysis day" effect was easily achieved because the analysis of fats from the different countries and three feed types were spread evenly over the different analysis days. To eliminate the daily drift of the abundance of each of the 191 normalised mass fragments $F_{i (N)}$, the R_i ratio of the overall mean of the 351 analyses to the average of the daily analysis of each $F_{i (N)}$ was calculated. The abundance of each $F_{i (N)}$ was multiplied by the corresponding ratio R_i. Matrix M3 containing 191 normalised mass fragments corrected for the daily drift $F_{i (N, DD)}$ and 351 observations was thus obtained from matrix M2.

Median filtering of the replicates per mass fragment was performed to retain the more robust median of each of the three replicates rather than the mean [5]. Matrix M4 containing the median of the replicates of each of the 191 normalised mass fragments corrected for the daily drift $F_{i (N, DD, Med)}$ and 117 observations was thus obtained from matrix M3, and used to build classification models for the three animal feed types.

Classification operations

After a step to select the relevant information the samples of each group were classified by DA_s and ANN_s.

A limited subset of relevant mass fragments extracted from the M4 matrix, was selected by stepwise discriminant analysis (SDA) [6] to classify fat tissues according to the three types of animal feed. The selection ($p < 0.05$) afforded a matrix $M5_{input}$ containing 20 mass fragments and 117 observations. This matrix was used to build classification models based on linear (DA_s) or non-linear (ANN_s) methods.

DA_s [7] yielded the percentage of samples correctly classified by cross-validation of the models, and visualised the classification performance from the coordinates of the samples on the canonical axes ($M6_{DAestimate}$).

For classification by ANN_s; the algorithm used was standard back propagation [8]. The structure of the network used to classify fat tissue samples consisted of three layers containing 20 inputs (20 selected variables of the $M5_{input}$ matrix), and one hidden layer containing 15 nodes and three outputs ($M6_{output}$ matrix). The $M6_{output}$ matrix is a disjunctive table containing three binary variables for the three classes of feed types. The architecture of the network was 20-15-3. The training of this type of network consisted in an iterative adjustment of the weights of the connections between input- ($M5_{input}$ matrix) and output nodes ($M6_{output}$ matrix) in order to minimise the error of prediction of the $M6_{output}$ matrix. Computations of the networks during cross-validation steps were stopped when the validation error of the networks stopped decreasing and with a maximum of 300 iterations of the back propagation algorithm. The tuning of each network training run was: learning rate: 0.30; back-propagation momentum: 0.10, and additional noise regularly decreasing from 0.20 to 0.05. The activation function was a logistic. An illustration of the cross-validation results has been obtained with the sample plot of a PCA calculated from the $M6_{ANNestimate}$ matrix obtained by cross-validation of the neural model [9].

Cross-validations were of the "leave-one-out" type [6]. They consisted in successively removing each of the $n = 117$ observations from the M5 matrix and predicting its class (animal feed type) from the remaining 116 observations. For the DAs, 117 different analyses were thus carried out from 117 different sub-populations containing 116 observations, and the corresponding results of classification of each observation were presented in confusion matrices. Similarly, for the ANN_s, 117 different training runs were carried out from the $n = 117$ different sub-populations containing 116 observations. The result of the classification of each observation by these 117 neural networks was presented in a confusion matrix.

Results and discussion

The total mean time needed for the manual preparation of each ready-for-pyrolysis foil (including pipetting of the glycoprotein support and sample, drying steps and placing the foil in the pyrolysis tube) was about 3 min (54 foils prepared in 2 h). The pyrolysis of lipids without a support resulted in instantaneous vaporisation of the lipids followed by re-condensation on the walls of the tube as well as in the expansion chamber of the pyrolyser. In these conditions pyrolysis of lipids is incomplete, yielding a very weak total ion count of about 8.5×10^4 (Fig. 1). In contrast to the analyses carried out by Goodacre *et al*. [1] on olive oil, or by Berdagué *et al*. [2] on pork backfat, we chose not to pyrolyse the lipids directly because of the risk of soiling the instrumentation. Instead we developed a reactive support that eliminated projection of material and increased the lipid signal during pyrolysis. The best results were obtained using a deposit of 2 µl of a sucrose (10 mg.ml^{-1}) and ovalbumin (5 mg.ml^{-1}) mixture. These conditions allowed the strongest lipid signal and were chosen to analyse ovine fats. The pyrolysis of a sugar and a protein causes Maillard reactions that favour the thermal fragmentation of lipids. However, although the total ion count assigned to the lipids was 3.4×10^5, the ratio of this signal to the total signal (foil + sucrose + ovalbumin + lipids) was only about 1/4. None of the other

types of support tested were retained because of poor adhesion to the foil or persistent projection of lipid material.

A visualisation of the effect of the different steps of the data pre-treatment is presented in figure 2. The normalisation procedure gives a less flattened shape (*i.e.*, less one-dimensional) to the scatter and shows up the aberrant data points (Fig. 2A and B). The elimination of aberrant data (Fig. 2C) brings out an "analysis-day" effect that is relatively classical in Py-MS and is partly attributable to fluctuations in the quality of the source vacuum. These fluctuations are not measurable with conventional vacuum gauges, but certainly contribute to memory effects from previous analyses. These memory effects slightly perturb the fragmentation yields of certain ions. In our experiments, the analysis day effect also integrates uncontrolled variance due to sample preparation. These effects result in data drift, which we have visualised in a simplified way with analyses from days 2 and 7 only (Fig. 2C). The elimination of the analysis day effect corresponds to a homothetic linear correction in 191-dimensional space. This correction, which could only be considered because the three animal feeding groups were distributed evenly over the different analysis days, proved very efficient (Fig. 2D). However, this correction is not applicable for industrial quality control operations, because an even daily distribution of animal samples according to feed type is unlikely in production situations. To day, several authors have proposed solutions for numerical correction of instrumental drifts in mass spectrometry [10-11]. However, because of the numerous causes of these drifts (quality of the vacuum in the source, pollution of transfer lines and/or source, ageing of the electron multiplier...) and because of their relatively stochastic nature, all corrections applied only provide local solutions in a particular context or period of time. Finally, to monitor and eliminate this type of instrumental drift, periodic use of reference samples during analysis is advocated. The use of signal processing with floating references to guarantee the consistency of Py-MS data is a possibility that remains to be investigated. Median filtering of the replicates resulted in a relatively Gaussian distribution of the data points in the space (Fig. 2E) that is much more amenable to the different algorithms for the selection of variables of ascending or descending type that we have tested (results not presented). It is also important to note that after the data pre-treatment operations, no structure related to "animal feed type" yet appears. The different pre-treatment steps used proved to be necessary for the subsequent calculations. Only after the four pre-treatments that yielded $M5_{input}$ matrix was acceptable discrimination achieved, with a coefficient of determination (r^2) of 0.66. The absence of any one of the pre-treatment steps caused the discriminant analysis classification performance to collapse, the best coefficients of determination obtained being less than or equal to 0.37.

The linear variable selection procedure is suboptimal relative to other procedures (genetic algorithms for example [12]), especially in the case of neural network modelling. However, we used this procedure for its speed and its satisfactory performance. The selection retained 20 ions out of
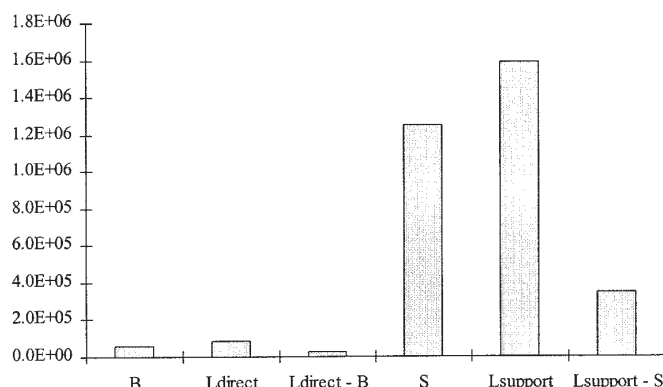

Total Ion Count

Figure 1. Comparison of the mean values ($n = 5$) of the Total Ion Count obtained by Py-MS for:
- B = blank (foil + tube + o-ring), L_{direct} = direct deposit of 30 µg of lipids (foil + tube + o-ring + lipids), $L_{direct} − B$ = estimate of the lipids signal obtained with a direct deposit,
- S = support (foil + tube + o-ring + sucrose and ovalbumin support), $L_{support}$ = deposit of 30 µg of lipids on the sucrose and ovalbumin support (foil + tube + o-ring + sucrose and ovalbumin support + lipids) and $L_{support} − S$ = estimate of the lipids signal obtained with a lipid deposit on the sucrose and ovalbumin support.

The ($L_{support} − S$) signal obtained with the reactive support (TIC = $3{,}4 \cdot 10^5$) was 13 time greater than the ($L_{direct} − B$) signal obtained without the reactive support (TIC = $2{,}7 \cdot 10^4$). The ($L_{support} − S$)/$L_{support}$ ratio was approximately 1/5.

the 191 ions of the pyrolyis spectra. The masses of these ions, classified by order of entry in the discriminant functions by the procedure were as follows: m/z = 191, 208, 190, 118, 64, 75, 219, 100, 112, 126, 218, 80, 72, 146, 85, 82, 53, 181, 93, 97. Given that the ions produced by pyrolysis are from two successive fragmentations (thermal fragmentation of the matrix, electronic impact fragmentation of pyrolysates), it is difficult to connect the masses of the fragments to the molecular species analysed. In this work the spectra are thus considered as fingerprints of products with no attempt made to identify the chemical origin of the differences observed between the fat from the three diets. The classification performance obtained by cross-validation by DA_s and ANNs were respectively 77.8 % (91/117) and 92.3 % (108/117) of lambs correctly assigned to their group (Fig. 3). With low numbers of samples, cross-validation may be adequate to westimate the error of classification [13]. However, when the number of samples is higher, a better estimate of the error value is obtained using bootstrap methods or, better, with a test sample drawn at random from the population [14].
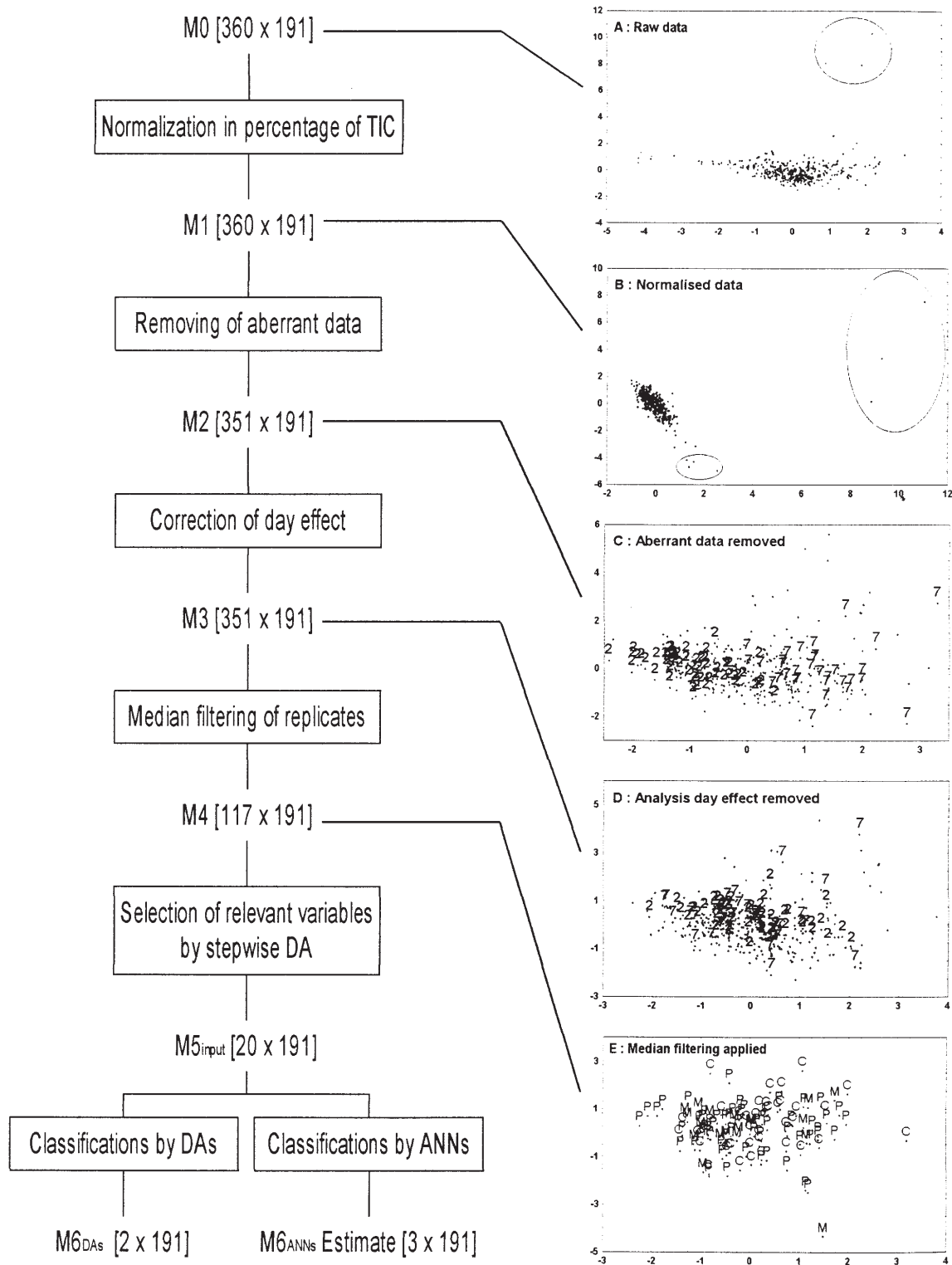
Figure 2. Description of the data analysis procedure.
Visualisation of the four data treatment steps before classification was obtained by principal component analysis.
Fig. 2A: Projection of raw data.
Fig. 2B: Normalisation of data and visualisation of aberrant data points.
Fig. 2C and 2D: Demonstration and correction of the analysis day effect (2 = day 2, 7 = day 7).
Fig. 2E: Visualisation after median filtering of replicates shows no apparent structure of data according to diet (C = Concentrate, P = Pasture, M = Milk).
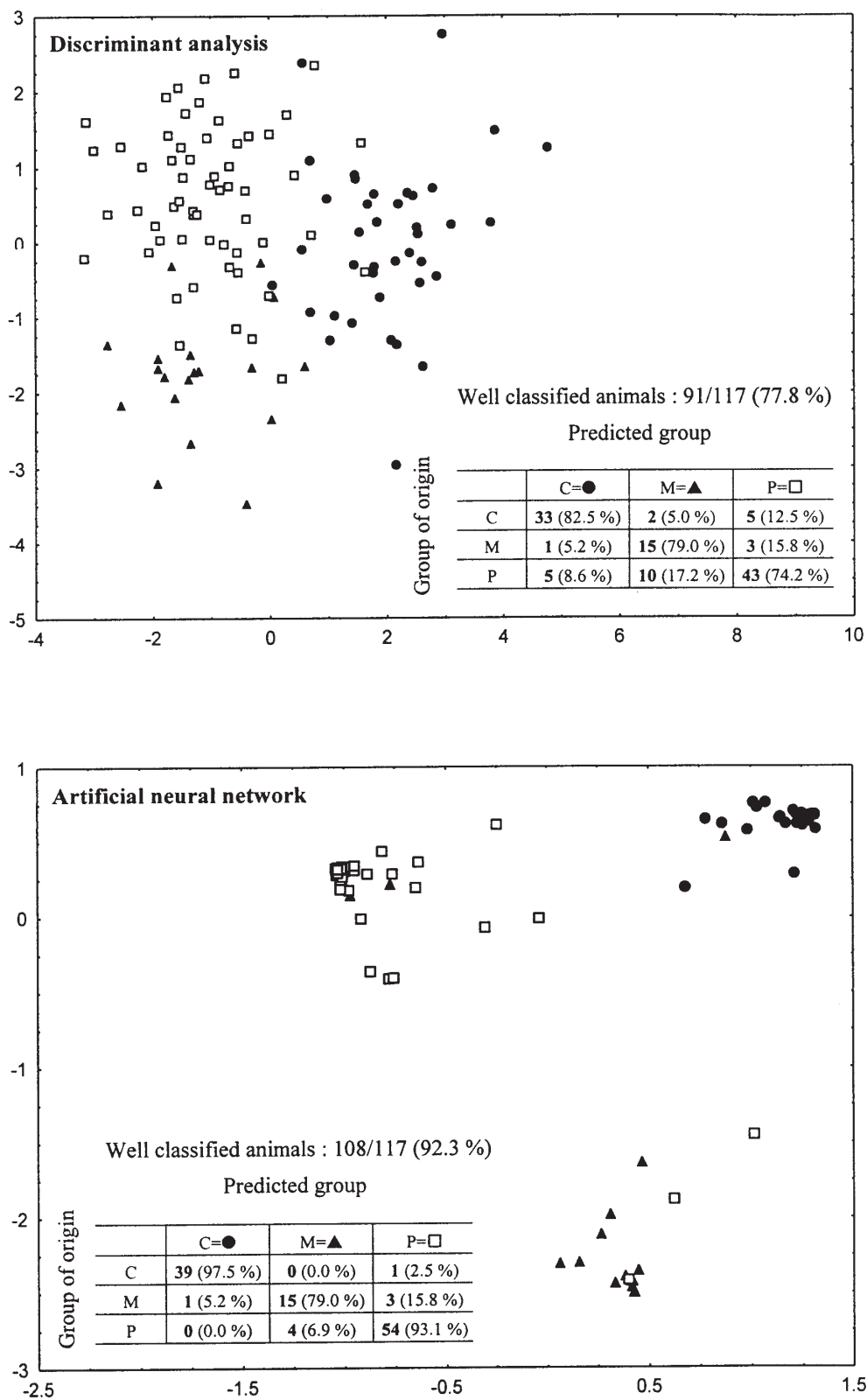
Figure 3. Plots of the cross-validation results obtained by Discriminant Analyses (A) and Artificial Neural Networks (B)
The samples are identified according to the diet: Concentrate (C = ●), Pasture (P = □) and Milk (M = ▲).
Results of cross-validations are presented in confusion matrix for each analysis.

Conclusion

The development of a method of analysis, from sample preparation (extraction of lipids, reactive support to amplify lipid pyrolysis) to the implementation of a data analysis strategy (pre-treatment of the data and extraction of relevant information), has successfully overcome the problem of post-mortem classification of commercial animals on the basis of their feeding background. However, the on-line application of this method in a control laboratory will require the optimisation of the mode of preparation of the samples before pyrolysis, and the provision of a robust means of correcting the time drift of the signal.

Acknowledgements

References

1. Goodacre, R.; Kell, D.B.; Bianchi, G. *J. Sci. Food Agric.* **1993**, *63*, 297-307.

2. Berdagué, J.L.; Rabot, C.; Bonneau, M. *Sciences des Aliments* **1996**, *16*, 425-433.

3. De Barruel, A.; Vernat, G.; Martin, J.F.; Berdagué, J.L. *Viandes & Produits Carnés* **1996**, *17 (6)*, 365-368.

4. Cardinal, M.; Viallon, C.; Thonat, C.; Berdagué, J.L., submitted for publication in *Aquaculture Research.*

5. Feinberg, M.; Ducauze, C. *Analusis* **1984**, *12 (1)*, 26-31.

6. SAS Institute Inc, In *SAS/STAT® User's Guide*, Release 6.03 Edition., 1988, 359-447.

7. SAS Institute Inc, In *SAS/STAT® User's Guide*, Release 6.03 Edition., 1988, 902-922.

8. McClelland, J.L.; Rumelhart, D.E. In *Explorations in parallel distributed processing; A handbook of models, programs and exercises;* MIT Press Ed.: Cambridge, 1988.

9. STATISTICA *A comprehensive system for statistics, graphics and application development*, Version 5.1 for Windows; Statsoft: Charenton-le-Pont, France, 1997.

10. Goodacre, R.; Kell, D.B. *Anal. Chem.* **1996**, *68*, 271-280.

11. Goodacre, R.; Timmins, E.M.; Jones, A.; Kell, D.B.; Maddock, J.; Heginbothom, M.L.; Magee, J.T. *Anal. Chim. Acta.* **1997,** *348*, 511-532.

12. Mitchell, M. In *An introduction to genetic algorithms*; MIT Press, Ed: Cambridge, 1996.

13. Efron, B. *Journal of the American statistical association* **1983**, *78 (382)*, 316-331.

14. Efron, B.; Gong, G. *The American Statistician* **1983**, *37*, 36-48.