Identification of animal species by Cluster Analysis

R. VERBEKE, D. VAN DE SOMPEL and H.F. DE BRABANDER
Laboratory of Chemical Analysis of Food from Animal Origin
Veterinary Faculty of the University of Ghent, Casinoplein,24 B-9000 Ghent, Belgium

INTRODUCTION

There is a need for simple and reliable analytical methods to determine the animal species from which meat products are made. Curing, processing and intensive heat treatment results in a denaturation of sarcoplasmic proteins hampering species specific detection by serological or electrophoretic methods. The only valuable approach in identifying the meat species in heat processed products should be based on a determination of the species specific heat stable components in the meat products. Chicken meat in heated pork has been detected by the anserine/carnosine ratio (1). However, admixture of beef or other species decreases the sensibility of the method. Characterisation of animal fats by fatty acid analysis, using "typical" fatty acid ratios, has been repeatedly reported (2,3). However,since the feeding regime may significantly affect the fatty acid composition, the discrimination of species on basis of fatty acid analysis only is of doubtful value.

In earlier investigations we demonstrated that the distribution pattern of the major fatty acids over the triglycerides of pork, beef, horse and chicken fat was species specific(4,5,6,7). The fatty acid composition of the triglycerides and the monoglycerides, isolated after a lipase treatment of the fats, was determined gaschromatographically. The fats can be classified in their correct groups through successive use of 4 discriminant parameters derived from the gaschromatographic data. With this procedure the relative percentage of the fat species in a binary mixture may be calculated. With the evolution of modern computertechnology, specialised software became available which make the interpretation of multiple observations (such as our fatty acid data) with classification procedures or cluster analysis more easy. A computerprogram should be able to detect structures in a given multi-dimensional data set. Since we have at our disposal the data of fats obtained from diverse anatomical locations from animals on different feeding regimes we wondered if a computerprogram would be able to detect "species" structures in this data set.

In this investigation our data set was analysed with a cluster analysis computer-program CLUSTAN2 . The results are compared with our earlier observations.

EXPERIMENTAL

Materials and methods

In earlier papers (5,6) the pig, beef, horse and hen fats analysed were described. The fats were extracted from meat samples in chloroform-methanol and the triglycerides isolated by TLC-chromatography (silicagel 60). The fat tissue samples were homogenised, melted and filtered at 80°C. The clear fat was stored in the freezer (-20°C) until used.

Fats were transesterified by incubating 20 mg fat in presence of 1 ml sodium methylate solution (0.025 N) in methanol at 90°C during 1 h. The fatty acid composition in position 2 of the triglycerides was determined by a modification of the method described before (6). Pancreatic lipase (100 mg; E.C. n° 3.1.1.3.; Sigma type II) was homogenised with 1 ml 1 M TRIS-buffer (pH = 8.2). On a piece of ground glass of 1.5 x 7 cm (e.g. a cover of a tank) 250 µl lipase solution was applied. A homogeneous lipase reaction band was formed on silicagel plates (10 x 20 cm) by gently pushing the plate against the ground glass piece. 100 µl of a fat solution (80 mg of fat in 1 ml n-hexane) was evenly applied over the lipase reaction band. The silicagel plate was placed at 2 cm above the water surface. After 10 min. incubation the plate was removed and dried carefully. The lipid mixture was concentrated into a narrow band by developping the plate three times with diethylether-formic acid (98:2, v/v) over a distance of ± 5 cm. The lipase reaction band was removed by cutting off the part of the plate. The remainder of the plate was developped in n-hexane-diethylether-formic acid (80:20:2, v/v/v). After drying, the monoglyceride fraction was transferred into a small column (0.6 mm I.D.) and elution was performed with 2, 1 and 1 ml freshly distilled, dry diethylether. The ether was evaporated under a jet of nitrogen. The lipids were transesterified with 200 µl sodium methylate solution. The gaschromatograph used was a Varian 3700. A capillary column (50 m; 0.25 mm I.D.; R.S.L.; Belgium) coated with SILAR 10C was used .The carrier gas was H2 at 2 ml/min. The temperature of the column, the injector and the detector was at 160, 210 and 220°C respectively.

Data analysis

All calculations were performed on a SIEMENS 7551 mainframe computer using BS 2000 as operating system. Communication with the computer and introduction of the data was through a ADDS Regent 200 terminal. The data were analysed with a cluster analysis program CLUSTAN2 (Release 2.1, Computing laboratory of the University of St-Andrews, FIFE KY16 9SX Scotland). The eigenvectors and factor scores or principal components were also calculated. Only the first two principal components were plotted. Four datasets were presented to the computer. All the sets

consist of 70 fat data (21 pork, 14 horse, 15 beef and 20 chicken fats). In a first set the fatty acid composition of the total triglycerides (TG = 7 parameters: the concentration, in mole % of the following fatty acids : 14:0 16:0 16:1 18:0 18:1 18:2 18:3) were analysed . A second set consisted of the fatty acid composition of the monoglycerides isolated by the procedure described above (MG = 7 parameters). In a third set the fatty acid composition of both the triglycerides and the monoglycerides were used (TG & MG = 14 parameters).In the fourth set the proportion in 2-position (calculated as (MG/TG)*33.33333)(PROP = 7 parameters) was analysed. Scatter diagrams of the first two principal components found and dendrograms (method of Ward) were plotted with a CALCOMP OFF-LINE plotter.

RESULTS AND DISCUSSION

Scatter diagrams of the principal components

The scatter diagrams of the first two principal components obtained for the four datasets are given in Figs 1,2,3 and 4. Fig 1 shows the results of the dataset consisting of the triglyceride composition (TG) of the four species. Horse fat forms a cluster clearly separated from the cluster formed by the beef,pork and hen fats. Those three fats form clusters which are not distinctly separated. Especially one beef fat and four pork fats could easily be misclassified as hen fats. The first two principal components correspond with a cumulative variance of 75 %. In Fig 2 the results of the MG (monoglyceride) data set are given. The scatter diagram shows three clearly separated clusters: horse, pig and beef+hen. The beef samples form a group marginally confused with the hen samples. Especially 5 hen fats could be misclassified as beef fats. The first two principal components correspond here with a cumulative variance of 66 %. In Fig 3 the computerprogram was run on the TG & MG data.set The general appearance of Fig 1 and 3 is alike. Horse fat may clearly be discriminated from beef, pork and hen fats, appearing as three marginally separated clusters. Again one beef fat and four pork fats could be misclassified as hen fat. The cumulative variance of the first two principal components is 62 %. In Fig 4 the results for the proportion in 2 position (given by the formula PROP-2 = (MG/TG)*33.33333) are presented. In this figure the pork fat cluster is.differentiated from the other three fats. Hen fat is marginally separated from beef and horse fat forming one cluster.One hen fat is positioned half-way the pork and the hen cluster. The first two principal components correspond with a cumulative variance of 82 %. The best separation was obtained with the MG data set (Fig 2)

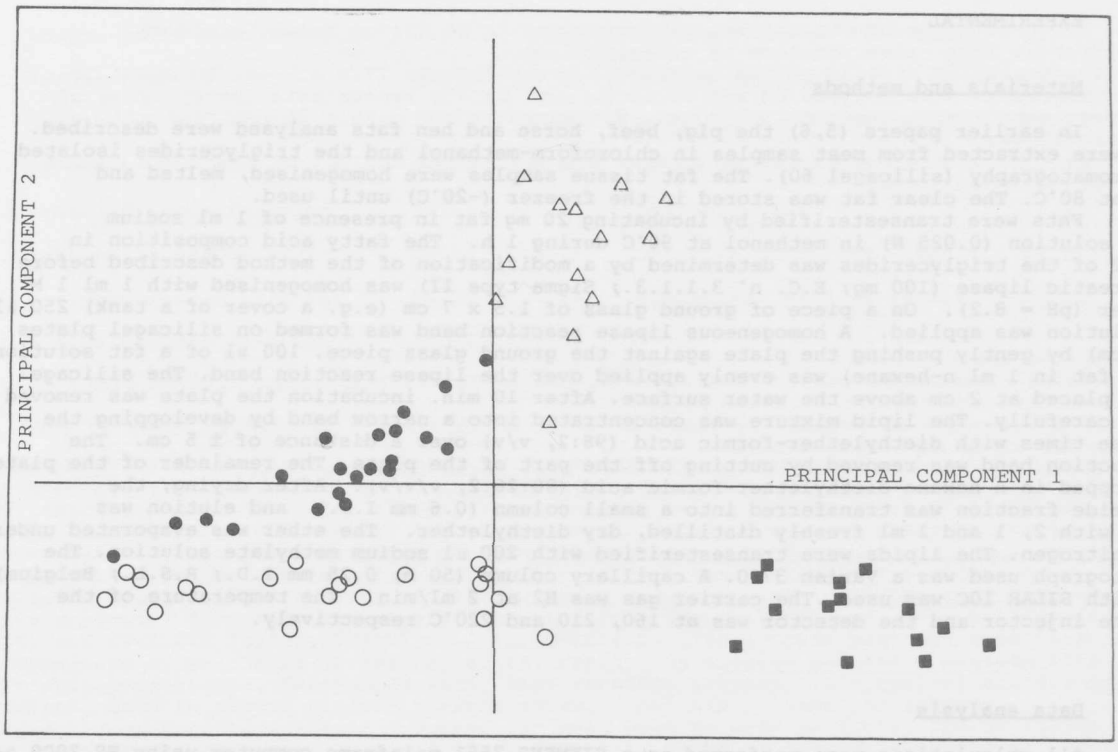

FIG 1 : Scatter diagram of the first two principal components found in the triglyceride composition of pork(●)(n=21) beef(△)(n=15) horse (■)(n=14) and hen(○)(n=20) fats. The cumulative variance of these two principal components is 75 %

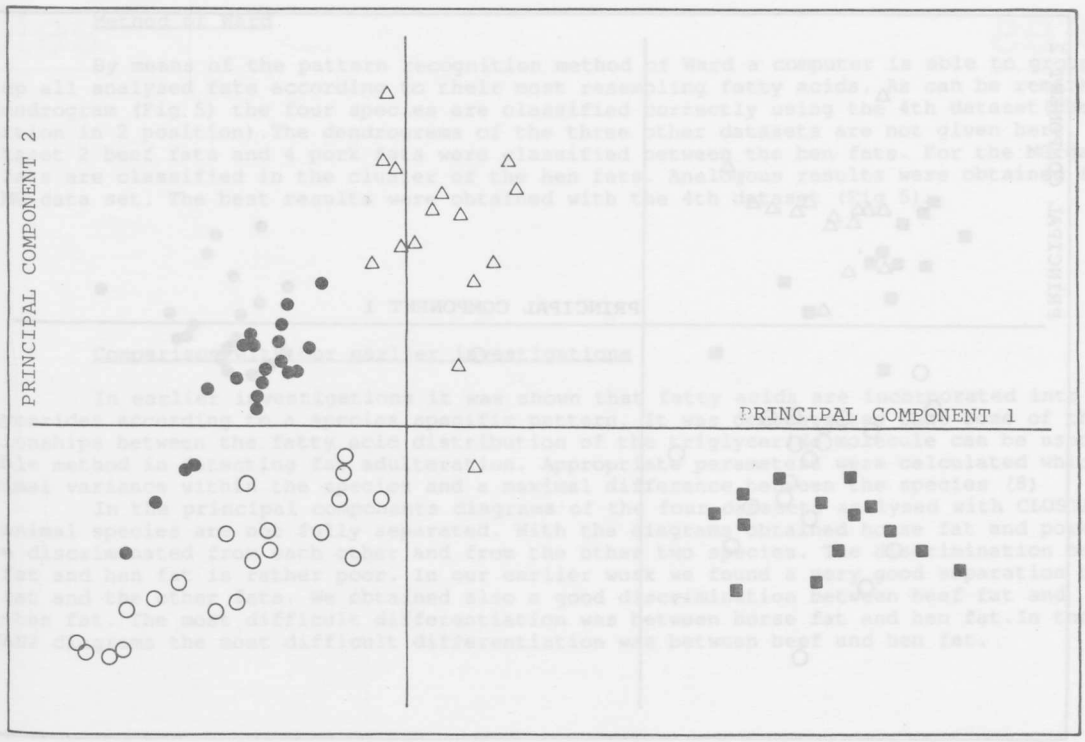

FIG 3 : Scatter diagram of the first two principal components found in the triglyceride + monoglyceride dataset of pork(●)(n=21) beef(△)(n=15) horse (■)(n=14) and hen(○)(n=20) fats. The cumulative variance of these two principal components is 62 %.

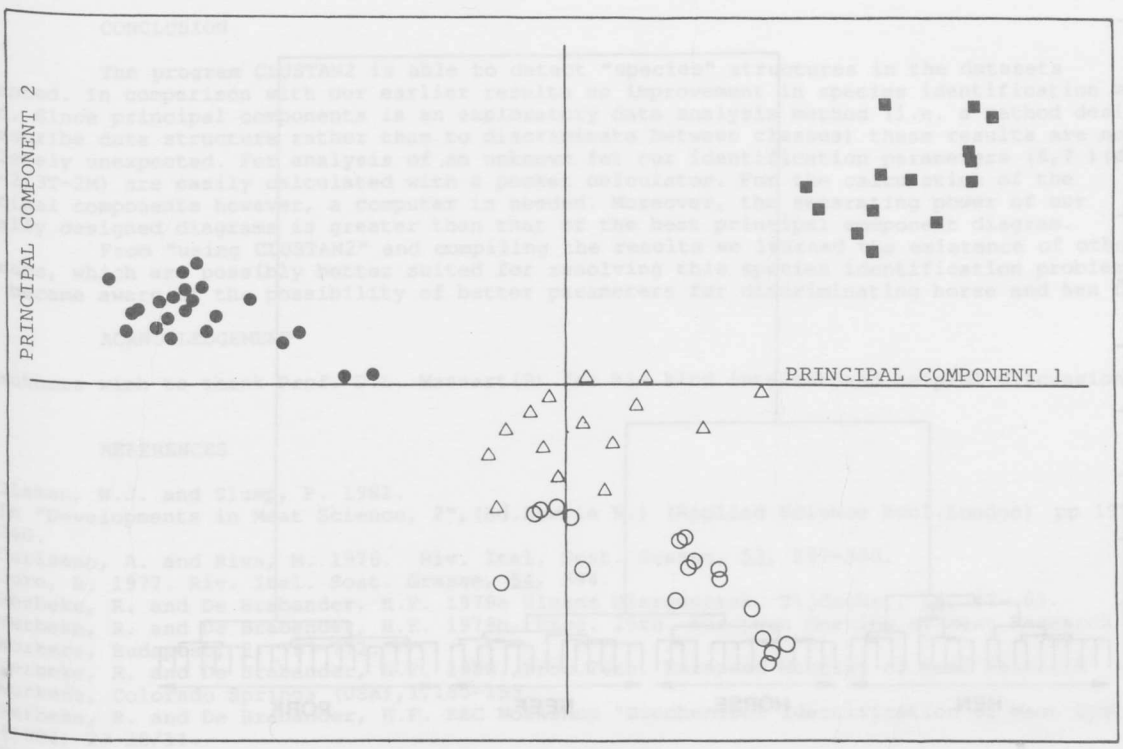

FIG 2 : Scatter diagram of the first two principal components found in the monoglyceride composition of pork(●)(n=21) beef(△)(n=15) horse (■)(n=14) and hen(○)(n=20) fats. The cumulative variance of these two principal components is 66 %.

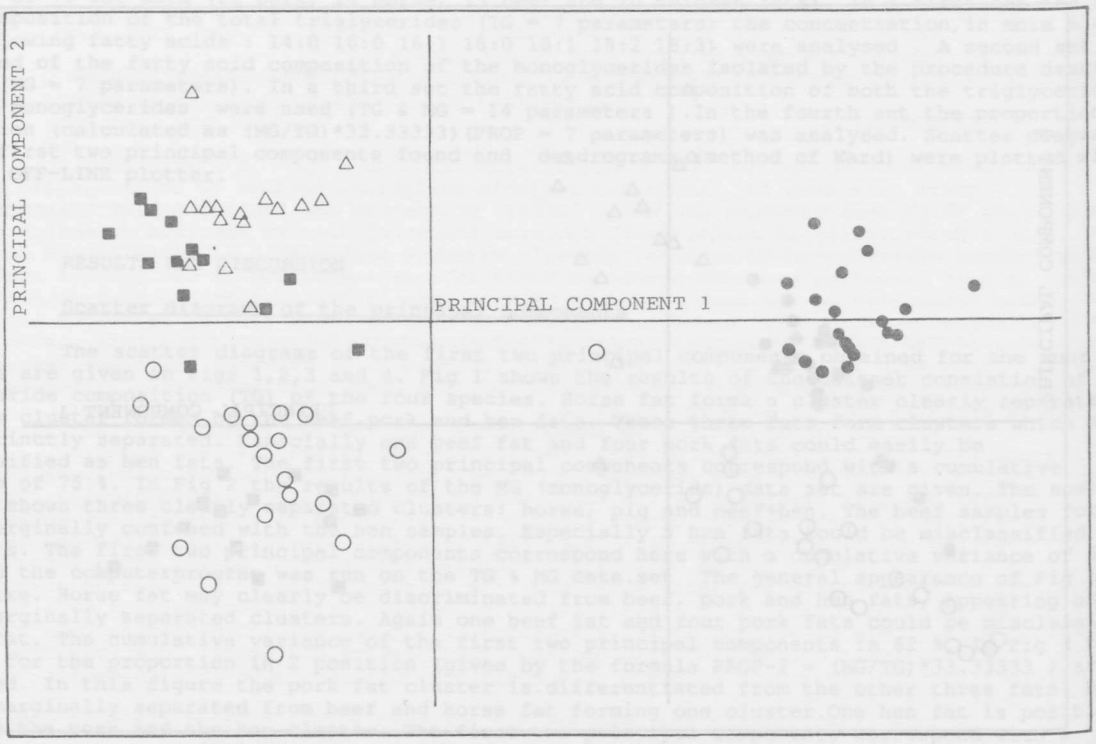

FIG 4 : Scatter diagram of the first two principal components found in the proportion in 2-position dataset (PROP 2 = (MG/TG)* 33.33) of pork(●)(n=21) beef(Δ)(n=15) horse (■)(n=14) and hen(O)(n=20) fats.
The cumulative variance of these two principal components is 82 %.

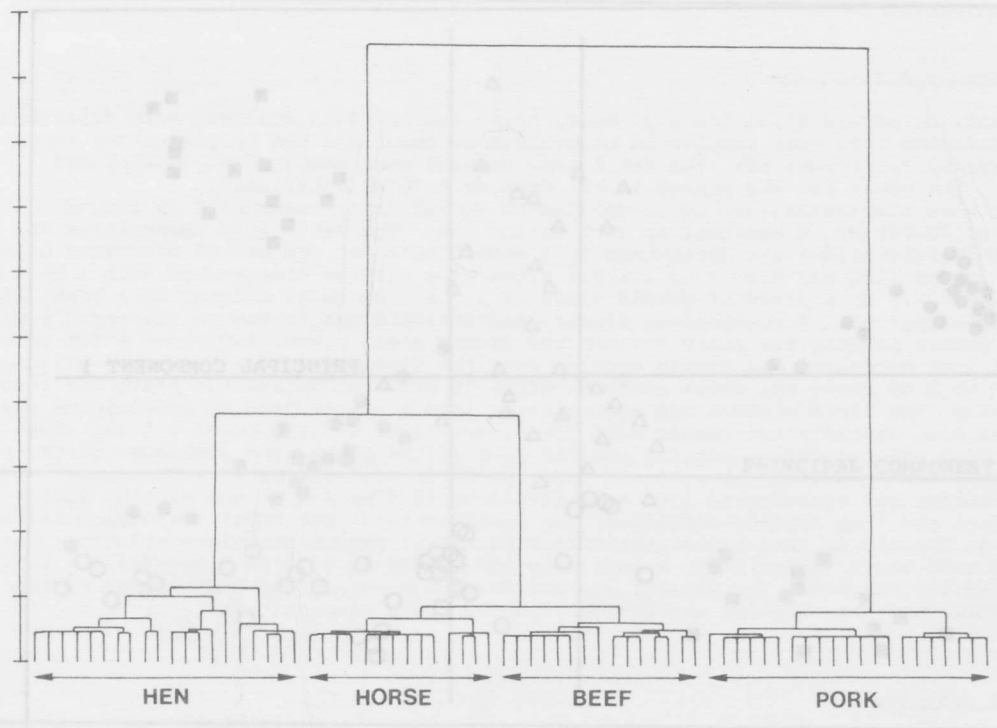

FIG 5 : Dendrogram (method of Ward) of the proportion in 2-position dataset.

Method of Ward

By means of the pattern recognition method of Ward a computer is able to group step by step all analysed fats according to their most resembling fatty acids. As can be remarked from the dendrogram (Fig 5) the four species are classified correctly using the 4th dataset(the proportion in 2 position).The dendrograms of the three other datasets are not given here. For the TG dataset 2 beef fats and 4 pork fats were classified between the hen fats. For the MG dataset 2 beef fats are classified in the cluster of the hen fats. Analogous results were obtained for the TG & MG data set. The best results were obtained with the 4th dataset (Fig 5).

Comparison with our earlier investigations

In earlier investigations it was shown that fatty acids are incorporated into the triglycerides according to a species specific pattern. It was demonstrated that some of the relationships between the fatty acid distribution of the triglyceride molecule can be used as a reliable method in detecting fat adulteration. Appropriate parameters were calculated which gave a minimal variance within the species and a maximal difference between the species (8)

In the principal components diagrams of the four datasets analysed with CLUSTAN2 the four animal species are not fully separated. With the diagrams obtained horse fat and pork fat may be discriminated from each other and from the other two species. The discrimination between beef fat and hen fat is rather poor. In our earlier work we found a very good separation between pork fat and the other fats. We obtained also a good discrimination between beef fat and horse+hen fat. The most difficult differentiation was between horse fat and hen fat.In the CLUSTAN2 diagrams the most difficult differentiation was between beef and hen fat.

CONCLUSION

The program CLUSTAN2 is able to detect "species" structures in the datasets presented. In comparison with our earlier results no improvement in species identification was found. Since principal components is an exploratory data analysis method (i.e. a method designed to describe data structure rather than to discriminate between classes) these results are not completely unexpected. For analysis of an unknown fat our identification parameters (6,7)(e.g. PROP-2,3T-2M) are easily calculated with a pocket calculator. For the calculation of the principal components however, a computer is needed. Moreover, the separating power of our manually designed diagrams is greater than that of the best principal component diagram.

From "using CLUSTAN2" and compiling the results we learned the existence of other programs, which are possibly better suited for resolving this species identification problem .We also became aware of the possibility of better parameters for discriminating horse and hen fat.

ACKNOWLEDGEMENT

The authors wish to thank Prof. D.L. Massart(9) for his kind interest and helpful discussions.

REFERENCES

(1) Olsman, W.J. and Slump, P. 1981. In "Developments in Meat Science, 2",(Ed.Lawrie R.) (Applied Science Publ.London) pp 195-240.
(2) Carisano, A. and Riva, M. 1976. Riv. Ital. Sost. Grasse, 53, 297-300.
(3) Doro, B. 1977. Riv. Ital. Sost. Grasse, 54, 394.
(4) Verbeke, R. and De Brabander, H.F. 1979a Vlaams Diergeneesk. Tijdschr., 48, 47- 63.
(5) Verbeke, R. and De Brabander, H.F. 1979b. Proc. 25th. European Meeting of Meat Research Workers, Budapest, 2, 767-772.
(6) Verbeke, R. and De Brabander, H.F. 1980.,Proc.26th. European Meeting of Meat Research Workers, Colorado Springs (USA),1,150-153.
(7) Verbeke, R. and De Brabander, H.F. EEC Workshop "Biochemical Identification of Meat Species" (1984) 27-28/11.
(8) Fisher, R.A. 1936. Ann. Eugenics, London, 7, 179-188.
(9) Massart D.L. and Kaufman L., The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis, (1983) John Wiley & Sons, New York.