Detection of aliment adulteration by means of near infrared spectroscopy—a feasibility study based on open-source R Project

György Bázár,[a] György Kövér,[b] László Locsmándi,[a] András Szabó[a] and Róbert Romvári[a]

[a]*Laboratory of Animal Product Qualification, Faculty of Animal Science, Kaposvár University, Guba S. str. 40. H-7400 Kaposvár, Hungary.*
E-mail: bazarg@eventing.hu
[b]*Department of Mathematics and Physics, Faculty of Economic Science, Kaposvár University, Guba S. str. 40. H-7400 Kaposvár, Hungary*

Introduction

The determination of the origin of food raw materials, and the detection of adulteration are major issues for the food industry and are attracting research topics. The native Mangalica pig breed is one of the most important issues in this matter in the Hungarian meat industry. Mangalica pigs are used for extensive rearing on prairie-like areas under severe conditions, utilising only the available feed. This pig has a dark and fatty, marbled meat, giving a very favourable taste, and it is beneficial for culinary techniques because it increases the level of sensory qualities. Also this meat is excellent to produce the original Hungarian salami sausage or meat products, with extremely prolonged ripening, such as Serrano ham. Among other reasons, the high commercial value of meats of high consumer popularity leads to an expressed need for fast and reliable methods to identify the animal species and type, either in carcass meats, or even more in meat products. Near infrared (NIR) spectroscopy is one of the most progressive methods frequently used for discriminating between different meats.[1–6] Apart from the mostly used "target softwares", the use of special, open-source software-packages (e.g. R Project for Statistical Computing, www.r-project.org) is spreading. The aim of our study was to evaluate a NIRS method to discriminate between pork groups originated from industrially reared commercial genotypes, and organically reared autochthonous Mangalica. Our goal was to build and test the method with the R project and its GPLS package.

Materials and methods

Ninety-one (91) meat samples were analysed in this study, comprising 27 Mangalica, 26 Landrace, 27 Large White and 11 Landrace × Large White crossbreed (last three groups are referred to as intensive breeds) meat samples. Mangalica were reared and fed under extensive conditions, while the other genotypes originated from industrial rearing systems, consuming commercially available feeds. Intensive pigs were slaughtered at an average weight of 104 kg, while Mangalica were slaughtered at 157 kg. Homogenised loin (*m. longissimus dorsi*) samples were scanned freshly and freeze-dried. NIR spectra were collected in reflectance mode using a NIRSystems 6500 spectrometer (FOSS NIRSystems, Silver Spring now Laurel, MD, USA) equipped with a sample transport module and small ring cup cuvette (IH-0307). Reflectance spectra were recorded from the 1100 to 2500 nm region and recorded as –log(*R*) at 2 nm intervals, with the WinISI II version 1.5 spectral analytical software (InfraSoft International LLC, Port Matilda now State College, PA, USA).

Pure spectral databases were exported from WinISI in *.txt format without being transposed. Separate files were saved for each group. Classification was developed with R Project for identifying breeds. Generalised Partial Least Squares method (GPLS package) was used, which is based on an extension of PLS in the context of generalised linear regression.[7–9] No spectral pre-treatment was applied. Full leave-one-out cross-validation and independent validations were carried out to test the system.

All 27 Mangalica and 39 randomly selected intensive samples were analysed by wet-chemistry. The intramuscular fat (ether extract) content of samples was determined by the Soxhlet method. Hydrochloric acid digestion and a Kjel-Foss Fast Nitrogen Analyser was used for the determination of the nitrogen content; protein content was obtained by multiplying these data with 6.25. Chemical data are used on a dry matter basis [DM%], thus values obtained can be applied correctly both for fresh and freeze-dried samples.

Results and discussion

In Trial 1, all samples were involved in the discrimination. The successful classification (Table 2 – Trial 1) was attributed to the considerable differences in intramuscular fat content of the two groups (Table 1).

Table 1. The ether extract and protein content [DM%] of investigated meat samples (Bázár *et al.*[10]).

		Ether extract content (dry matter based) [%]				Protein content (dry matter based) [%]			
Genotype	*n*	Mean	*SD*	Minimum	Maximum	Mean	*SD*	Minimum	Maximum
Mangalica	27	19.1[a]	5.1	11.6	33.0	78.5[a]	5.2	64.6	85.9
Intensive	39	9.4[b]	2.7	4.4	14.8	89.2[b]	3.1	82.3	95.5
Landrace	8	10.4[b]	2.3	6.2	14.0	87.9[b]	2.6	83.6	92.2
Large White	20	8.9[b]	3.0	4.4	14.8	89.3[b]	3.7	82.3	95.5
Landrace × Large White	11	9.3[b]	2.1	6.1	12.7	90.0[b]	1.7	87.8	93.9

[a, b]: $P < 0.05$.

Table 2. Summarised results of presented trials (*Bázár et al.*[10]).

Trial	Fresh Samples				Freeze-dried Samples			
	Cross-validation		Independent validation		Cross-validation		Independent validation	
	Factor No.	Hits	Factor No.	Hits	Factor No.	Hits	Factor No.	Hits
1	7	100%	–	–	5	100%	–	–
2	4	90%	4	97.2%	4	96.6%	4	94.4%
3	4	100%	4	91.7%	4	100%	4	94.4%
4	4	100%	5	90.5%	4	100%	5	95.2%

Trial 1: all samples involved for generating discriminator equation ($n = 27 + 64$), without independent test;
Trial 2: discriminator equation generated on overlapping groups ($n = 15 + 15$), independent test with highly different groups ($n = 12 + 24$) (Figure 1);
Trial 3: discriminator equation generated on highly different groups ($n = 15 + 15$), independent test with overlapping groups ($n = 12 + 24$) (Figure 2)
Trial 4: randomly selected samples used for discriminator equation ($n = 20 + 50$), independent test with remaining samples ($n = 7 + 14$)

To overcome the possibly robust impact of the chemical composition of meats on the performance of the discrimination system, special sample sorting based on chemical composition was applied.

In the first check (Table 2 – Trial 2, Figure 1), 15 Mangalica samples with the lowest fat content (15.5 ± 2.3 DM%) were chosen and 15 with the highest fat level (11.9 ± 1.6 DM%) from the intensive samples. Thus, we composed two groups for calibration with overlapping fat contents. The equation used for discrimination was tested on the remaining 12 Mangalica and 24 intensive samples (intramuscular fat contents: 23.7 ± 3.9 DM% and 7.7 ± 1.7 DM%, respectively). 27 of 30 fresh samples were classified correctly during the cross-validation, and there was only one misclassified of the 36 independent samples. The better result obtained for the fresh samples during the independent validation can be explained by the special sorting conditions. Samples of the groups in the independent validation set ($n = 36$) were more different from one another than the samples in the cross-validation set ($n = 30$). As for freeze-dried samples, one and two samples were misclassified during cross-validation and independent validation, respectively.

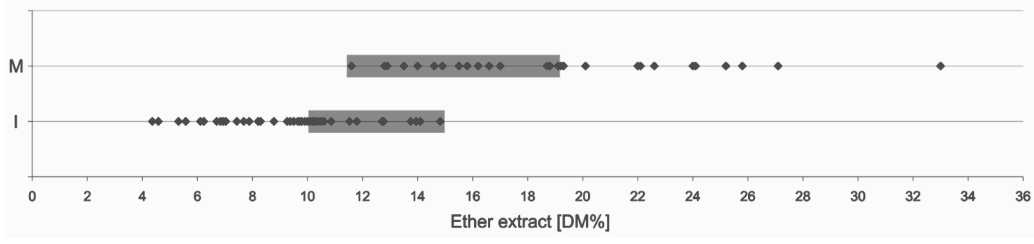

Figure 1. Trial set-up by using overlapping groups for calibration, representing identical samples (Trial 2) M: Mangalica genotype; I: intensive genotypes; box: samples used for calibration (Bázár *et al.*[10]).

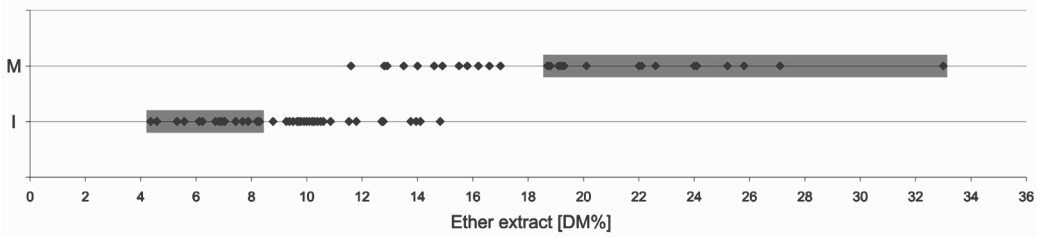

Figure 2. Trial setup by using furthest groups for calibration, representing identical samples (Trial 3) M: Mangalica genotype; I: intensive genotypes; box: samples used for calibration (Bázár *et al.*[10]).

The check was repeated by choosing 15 Mangalica samples with extremely high fat content (22.7 ± 4.0 DM%) and 15 intensive samples with extremely low (6.6 ± 1.2 DM%) (Table 2 – Trial 3, Figure 2).

The equation generated on these furthest groups was tested on the remaining 12 Mangalica and 24 intensive samples with overlapping intramuscular fat contents (14.6 ± 1.7 DM% and 11.0 ± 1.7 DM%, respectively. After a faultless cross-validation there were three unidentified fresh samples in the independent test. The mistake rate dropped by two mis-classified samples by using the freeze-dried forms. It seems that the system classified the samples not only by fat content, but the total multicomponent structure also has a great impact.

A final discriminatory equation (Table 2 – Trial 4) was generated using 20 randomly selected Mangalica and 50 randomly selected intensive pork samples. The equation was tested with the remaining 7 Mangalica and 14 intensive samples. Cross-validation was free of falsely classified samples. Using an independent test, two fresh samples and one freeze-dried sample were classified improperly.

Conclusion

The use of NIR spectroscopy combined with the open source R Project seemed to be effective for the recognition of the pork samples of different pig breeds (Mangalica and commercial genotypes) and production systems, without any spectral pre-treatments. The technique applied as in this trial is helpful for practical use in the meat industry, since meat with high value could be separated from meat with a lower quality level. Further studies are needed to test the method with mixtures of different meats, and also with meat products, in order to get closer to practical application.

Acknowledgement

This work was established by the support of the National Office for Research and Technology (BGY_NIRS). The support of the Bolyai János Research Grant (Bo_108_07) to A. Szabó by the Hungarian Academy of Sciences is gratefully acknowledged. Thanks are due to Professor Károly Kaffka for his kind guidance.

References

1. G. Downey and D. Beauchene, *Meat Sci*. **45,** 353 (1997).
2. K. Thyholt, T. Isaksson, I.K. Hildrum, M.R. Ellekjaer and O.C. Eide, *J. Near Infrared Spectrosc.* **6,** A361 (1998).
3. J. McElhinney, G. Downey and T. Fearn, *J. Near Infrared Spectrosc.* **7,** 145 (1999).
4. D. Alomar, C. Gallo, M. Castaneda and R. Fuchslocher, *Meat Sci.,* **63,** 441 (2003).
5. D. Cozzolino, A Vadell, F. Ballesteros and G. Galietta, *Anal. Bioanal. Chem*. **385,** 931 (2006).
6. F.G. Del Moral, A. Guillén, L.G. Del Moral, F. O'Valle, L. Martínez and R.G. Del Moral, *J. Food Eng.* **90,** 540 (2009).
7. D.J. Cummins and C.W. Andrews, *J. Chemometr.* **9,** 489 (1995).
8. M.I. Griep, I.N. Wakeling, P. Vankeerberghen and D.L. Massart, *Chemometr. Intell. Lab. Syst*. **29,** 37 (1995).
9. B. Ding and R. Gentleman, *Bioconductor Project Working Papers*, Working Paper 5 (2004). http://www.bepress.com/bioconductor/paper5
10. G. Bázár, G. Kövér, L. Locsmándi, G. Andrássy-Baka and R. Romvári, *J. Near Infrared Spectrosc.* **17,** 119 (2009). This paper is a short version of G. Bázár *et al.*, *J. Near Infrared Spectrosc*. **17,** 119 (2009), reprinted with permission of IM Publications LLP.