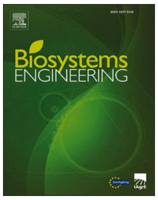
Biosystems Engineering 262 (2026) 104366

Contents lists available at ScienceDirect

Biosystems Engineering

journal homepage: www.elsevier.com/locate/issn/15375110



Research Paper

Discrimination of autochthonous pig breeds from meat and fat samples by FT-NIR spectra

S. Parrini [a,*], C. Dadousis [b,c], F. Sirtori [a], M.C. Fabbri [a], M. Čandek-Potokar [d], J.M. Garcia-Casco [e], B. Lebret [f], R. Nieto [g], C. Aquilani [a], R. Bozzi [a]

[a] Department of Agriculture, Food, Environment and Forestry, University of Florence, Piazzale delle Cascine 18, 5014, Florence, Italy
[b] School of Health Sciences, Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom
[c] Veterinary Health Innovation Engine, School of Veterinary Medicine, University of Surrey, Guildford, United Kingdom
[d] Kmetijski Inštitut Slovenije, Hacquetova ulica 17, 1000, Ljubljana, Slovenia
[e] Centro de I+D en Cerdo Ibérico, Dept. Animal Breeding, INIA-CSIC, Ctra Ex101 km 4.7, 06300, Zafra, Spain
[f] PEGASE, INRAE, Institut Agro, 35590, Saint-Gilles, France
[g] Department of Nutrition and Sustainable Animal Production, Estación Experimental del Zaidín, Spanish National Research Council, CSIC, Profesor Albareda, 1, 18008, Granada, Spain

ARTICLE INFO

Keywords:
Swine
Authentication
Near-infrared spectroscopy
Local breeds
Identification
Classification

ABSTRACT

The combination of Fourier transform near-infrared spectroscopy (FT-NIRS) of meat and fat samples and principal component discriminant analysis (DAPC) has been proposed as a tool for discriminating the local production of autochthonous pig breeds. Spectral samples (n = 272) belonging to 11 local European pig breeds, Longissimus muscle, and subcutaneous fat (both intact and minced) are collected. Classification accuracy based on DAPC was applied on FT-NIRS to predict breed of origin in i) semi-external cross-validation, splitting the data into training (80 %) and testing (20 %) sets; ii) external validation, in which one breed at a time was excluded from model training and classified in one of the remaining breeds. The effect of varying sample sizes from 50 % to 100 % of the data was assessed. Almost all breeds' spectra variability was summarised into two principal components for tissue and sample preparation. In cross-validation, intact fat yielded higher classification accuracies than intact meat, with less pronounced differences in minced samples. Success assignment rates of ~81–83 % were obtained for two breeds in intact meat samples and were higher than 83 % for five breeds in fat samples. For minced samples, correct assignments between 80 % and 100 % were possible for five breeds, both in meat and fat samples. Sample size marginally affected the results. External validation confirmed similarity among some breeds, with greater accuracy for fat samples. The assignments success provides encouraging results for discriminating local pig production, mainly based on fat, using a rapid, eco-friendly FT-NIRS method, which could serve as tool for quality assurance.

Nomenclature table

Abbreviations	Description
BI	Bisara
BS	Black Slavonian
DAPC	Principal component discriminant analysis
CS	Cinta Senese
CV	Semi-external cross-validation
FT-NIRS	Fourier transform near-infrared spectroscopy
GA	Gascon
IB	Iberian
KR	Krškopolje
K	number of clusters
LIW	Lithuanian White Old Type

(*continued on next page*)

This article is part of a special issue entitled: Technologies in Livestock published in Biosystems Engineering.

* Corresponding author.

E-mail addresses: silvia.parrini@unifi.it (S. Parrini), c.dadousis@surrey.ac.uk (C. Dadousis), francesco.sirtori@unifi.it (F. Sirtori), mariachiara.fabbri@unifi.it (M.C. Fabbri), meta.candek-potokar@kis.si (M. Čandek-Potokar), garcia.juan@inia.csic.es (J.M. Garcia-Casco), benedicte.lebret@inrae.fr (B. Lebret), rosa.nieto@eez.csic.es (R. Nieto), chiara.aquilani@unifi.it (C. Aquilani), riccardo.bozzi@unifi.it (R. Bozzi).

https://doi.org/10.1016/j.biosystemseng.2025.104366
Received 2 August 2024; Received in revised form 10 December 2025; Accepted 15 December 2025

(*continued*)

Abbreviations	Description
LWOT	Lithuanian Indigenous Wattle
MB	Negre Mallorqui
NIRS	Near-Infrared spectra
PC	principal component
PCs	principal components
TRN	training dataset
TST	validation data set
SHS	Schwäbisch-Hällisches Schwein
TU	Turopolje

Science4Impact statement

The study proposed a solution to quality assurance issues in the traceability process for European autochthonous pig breeds. Near-infrared spectra, coupled with discriminant analysis of principal components, are proposed as a tool to discriminate among different autochthonous pig breeds using meat or fat, both intact and ground.

Scientific evidence of its approach under practical application could allow for the distinction between the meat and fat of some local breeds without knowing either the quality characteristics of the samples or the specifications of the pigs. The discrimination system based on NIRS (a tool included in Precision Livestock Farming) and the DAPC application could represent an innovative, rapid, low-cost, and eco-friendly approach useful for all stakeholders in the supply chain (consumers, producers, and farmers). A traceability monitoring system for local breed products could support consumers in choosing products derived from local breeds and increase consumers' confidence.

1. Introduction

Monitoring systems that guarantee the safety, quality, and authenticity of animal products, such as meat and its derived products, especially from autochthonous breeds, are attracting increasing interest due to the growing importance of a "green image" for consumers (Acciaro et al., 2020). Autochthonous breeds play a crucial role in the production of high-quality animal products, as their specific genetic traits and adaptation to local environments inherently contribute to high sensory quality and nutritional value, which are primarily linked to traditional rearing systems (Bonneau & Lebret, 2010; Prieto et al., 2017). These local breeds and their derived products are also associated with positive consumer perceptions of animal welfare and product authenticity, aspects often reinforced by their association with traditional extensive rearing systems. In this context, tools to ascertain the product's origin and production conditions as a traceability monitoring system based on discrimination analysis could be useful to help control fraud, protect local pig breeds of limited use, and thereby increase consumer confidence in the swine food chain, and support the local economy (Alonso et al., 2020). European countries, such as Italy, Spain, France, Portugal, and Greece, are leading in food authentication research, particularly on geographical origin, adulteration, mislabelling, and food safety. This trend is likely driven by sustained interest in food authentication in European legislation (Danezis et al., 2016). However, administrative traceability systems remain vulnerable to errors and fraud. In recent decades, molecular approaches have gained increasing attention over traditional physicochemical methods. Several molecular methods have shown promise for agri-food surveillance: the most commonly cited in the literature are microsatellites, simple sequence repeats, single-nucleotide polymorphisms, and high-throughput tools such as the PorcineSNP60 BeadChip. Nevertheless, their large-scale implementation in meat traceability is still limited, mainly due to the need for specialised laboratories, complex procedures, and high per-sample costs (Fanelli et al., 2021).

Other analytical methods for product authentication include a direct approach to verify the geographical origin, which is mainly based on the determination of the matrix's chemical composition (Danezis et al., 2016). This approach has often been used to determine typical components of specific areas or production methods using molecular techniques when different breeds are used. DNA analysis is the most common method for identifying and quantifying meat species (Mabood et al., 2020; Muñoz et al., 2020), though these technologies are time-consuming, labour-intensive, expensive, and hazardous due to toxic solvents. Depending on the animal product, the difficulty of applying traditional discriminative methods is linked to the complexity of the matrix's chemical composition, as a single or a few components cannot adequately describe the sample (Xu et al., 2011). Furthermore, these analyses are rarely applied to small populations, as is the case with autochthonous breeds characterised by local production, due to the costs and logistical difficulties involved in controlling and tracing individuals, products, or traditional processes (Ortiz et al., 2020).

Near-infrared spectroscopy (NIRS) and chemometric analysis could serve as a detection tool for developing authenticity and traceability models, given their ease of use, cost-effectiveness, and environmentally friendly, non-invasive characteristics (Acciaro et al., 2020; Prieto et al., 2017). This technology requires a chemometric approach that, at the quantitative level, computes a matrix between the spectral data of the sample (absorbance) and the data from the reference analyses of the samples (Prieto et al., 2017). In qualitative analysis, discriminant analysis can be used to investigate the existence of clusters within a dataset to develop a prediction tool for new samples belonging to groups (Oliveri et al., 2021).

Of particular interest could be the combined use of principal component discriminant analysis (DAPC) (Jombart, 2008a; Jombart et al., 2010) and Fourier transform near-infrared spectroscopy (FT-NIRS) to trace the origin of meat based on specific NIR spectra of each breed. Originally developed for population genetics, DAPC is well-suited for high-dimensional data, where the number of variables (in this case, wavelengths) often exceeds the number of samples. DAPC has proven effective in other contexts involving complex multivariate data (Thia, 2023), including spectroscopic studies, as it addresses the limitations of standard classification methods, which typically apply discriminant analysis directly to spectral data without accounting for the high collinearity among wavelengths.

Autochthonous pig breeds are characterised by significant variability (within and between breeds) due to their diverse genetic background and environmental and productive factors (Čandek-Potokar et al., 2019). They could represent an exploratory field for applying NIRS for discrimination processes. In addition, to maximise the utility of NIRS technology in terms of reducing analysis time and costs, the use of fresh, intact samples that do not include preparation, such as grinding or homogenisation, could serve as an alternative, even if it has been shown that sample preparation does represent an advantage for the accuracy of NIRS predictions (Meza-Márquez et al., 2010).

The present study is part of a larger project devoted to promoting sustainable pork chains based on European autochthonous pig breeds, including meat quality and authentication. The objective of the present study was to investigate the combination of FT-NIRS and discriminant analysis on principal components (DAPC) applied to intact or minced meat and fat samples as a potential tool to discriminate the local production of autochthonous pig breeds.

2. Materials and methods

2.1. Sampling

The study included meat and fat samples collected on pigs from 11 local European breeds from the TREASURE project (https://treasure.kis.

si/), namely, Bisara (BI), Black Slavonian or Crna Slavonska (BS), Cinta Senese (CS), Gascon (GA), Iberian (IB), Krškopolje pig (KR), Lithuanian White Old Type (LIW), Lithuanian Indigenous Wattle (LWOT), Negre Mallorqui or Majorcan Black (MB), Schwäbisch-Hällisches Schwein or Swabian Hall pig (SHS) and Turopolje (TU), as reported in Table 1. The pig samples used in this study originate from eight countries (Table 1). Pigs of each autochthonous breed are reared in their respective traditional production areas using a semi-extensive system, exploiting available natural resources and integrating feed supplementation. Sample collection is performed using the same protocol and under comparable conditions (e.g., 1–2 days post-slaughter, refrigerated carcasses) to minimise potential seasonal or environmental variation. A total of 287 samples of longissimus lumborum muscle and subcutaneous fat (backfat) from the left side of the carcasses are collected, as previously reported in (Ortiz et al., 2020) for meat samples and in (Parrini et al., 2023) for fat tissue. Each partner country provided vacuum-packed, frozen samples at −20 °C, which were sent to the University of Florence for FT-NIR spectrum acquisition.

2.2. Spectra collection

For each meat and fat sample, after thawing at 4 ± 1 °C for 24 h, two aliquots of intact tissue are scanned by using an FT-NIRS Antaris II instrument (Thermo Fisher Scientific Inc., Waltham, MA, USA) in absorbance mode, considering the infrared region with wavenumbers between 3999 and 9999 cm^{-1}. Homogeneous sample areas for scanning are selected to ensure consistent sample thickness. Each sample is exposed to an electromagnetic scan in absorbance mode using a circular quartz cup spinner, with absorbance reported as log 1/R (R, 100 % reflectance). Each spectral measurement is obtained from 32 scans performed at a wavenumber resolution of 4 cm^{-1}. The instrument automatically corrects the spectra against the background signal under standardised room environment conditions. Additionally, laser focal depth and measurement parameters are standardised across all scans to maintain a constant sampling volume.

Subsequently, two aliquots of each sample are uniformly minced by an electric grinder before the NIRS data collection, following the same procedure to reduce heterogeneity prior to spectral acquisition. Therefore, both fat and meat samples are scanned by FT-NIRS initially as intact (n = 252) and subsequently as minced samples (n = 272). The increase in the number of minced samples is due to CS samples being available only as minced meat and fat (Table 1).

Table 1
Information on meat and fat samples from 11 local European pig breeds, used respectively as intact and minced, for FT-NIRS spectra.

Breed	Name code	Country of origin	No. meat samples		No. fat samples	
			intact	minced	intact	minced
Bisara	BI	Portugal	34	34	34	34
Black Slavonian (Crna Slavonska)	BS	Croatia	16	16	16	16
Cinta Senese	CS	Italy	–	20	–	20
Gascon	GA	France	38	38	38	38
Iberian	IB	Spain	36	36	36	36
Krškopolje	KR	Slovenia	18	18	18	18
Lithuanian Indigenous Wattle	LWOT	Lithuania	24	24	24	24
Lithuanian White Old Type	LIW	Lithuania	18	18	18	18
Negre Mallorqui (Majorcan Black)	MB	Spain	16	16	16	16
Schwäbisch-Hällisches Schwein	SHS	Germany	33	33	33	33
Turopolje	TU	Croatia	19	19	19	19
Total			252	272	252	272

2.3. Multivariate discriminant model development

For each sample, the final spectrum is obtained by averaging the spectra of the two aliquots; therefore, four datasets (2 tissues x 2 modes) are available, each composed of 3112 absorbance data points. Data analyses are performed with the R software (R Foundation for Statistical Computing, 2021). Before model development, quality control of the individual spectra is conducted. The full NIRS range is used in all analyses. Spectra are pre-processed using multiplicative scatter correction (MSC) to minimise scattering effects and baseline variations. Original spectra and MSC-treated spectra for intact meat and fat, as well as for minced meat and fat, are reported in the Supplementary material (Fig. S1–S4). Subsequently, the spectral data are mean-centred and scaled to ensure equal weighting of all variables before multivariate analysis. Outlier detection is performed using the Mahalanobis distance, applying a threshold of five standard deviations from the mean. In the first step, principal component analysis (PCA) is applied to visualise potential breed separation.

DAPC is performed on the spectra using the R package adegenet (Jombart, 2008b; Jombart et al., 2010(R Foundation for Statistical Computing, 2021) . For breed spectra discrimination, two approaches are considered: semi-external cross-validation (semi-supervised learning) and external validation (unsupervised learning). Using semi-external cross-validation (CV), the data are randomly (without replacement) split at 80:20 for the training (TRN) and validation sample (TST) sets, respectively. In this case, all breeds are present in both sets (stratified sampling). The effect of dataset size is tested at 50, 60, 70, 80, 90, and 100 % of the samples. In each scenario, 80 % of the samples are used as the training set (TRN) and 20 % as the validation set (TST), ensuring that each breed is represented in both sets. The procedure is repeated 10 times. Analysis is conducted separately for each of the four sample types: intact and minced meat, and intact and minced fat. The correct assignment of samples to their breed of origin is reported as the mean of 10 replicates.

Secondly, an external validation (unsupervised learning) is performed, with a separate TST analysis for each breed. In this scenario, the TST set consists of pigs from only one breed at a time, and no pigs of the TST breed are present in the TRN set. Therefore, the TST samples have to be classified in the remaining 10 breeds of the TRN set.

In both approaches, the optimal number of principal components (PCs) to use in the DAPC is selected via 10-fold cross-validation (30 replicates) in the TRN set, targeting the lowest mean squared error and allowing up to 300 PCs in the model. Clustering within the DAPC is carried out using the K-means algorithm, to perform a range of values for the number of clusters (K). For each K, the BIC (Bayesian Information Criterion) is calculated to identify the optimal number of clusters. Once the clusters are defined, discriminant analysis of principal components (DAPC) is performed by applying linear discriminant analysis (LDA) to the retained PCs to maximise discrimination between groups. Heatmaps of breed assignments are constructed using the R package ComplexHeatmap (Gu et al., 2016).

3. Results

3.1. Descriptive statistics and principal component analysis

The projections of the spectra of meat and fat samples using PCA, as well as the scree plots of the total explained variance of the spectra data, are shown in Figs. 1 and 2 for intact samples (meat and fat, respectively), and in Figs. 3 and 4 for minced samples (meat and fat, respectively). The first PC explained 66.9 % and 74.9 % of the original variability for intact meat and fat, respectively, and 70.7 % and 88.5 % for minced meat and

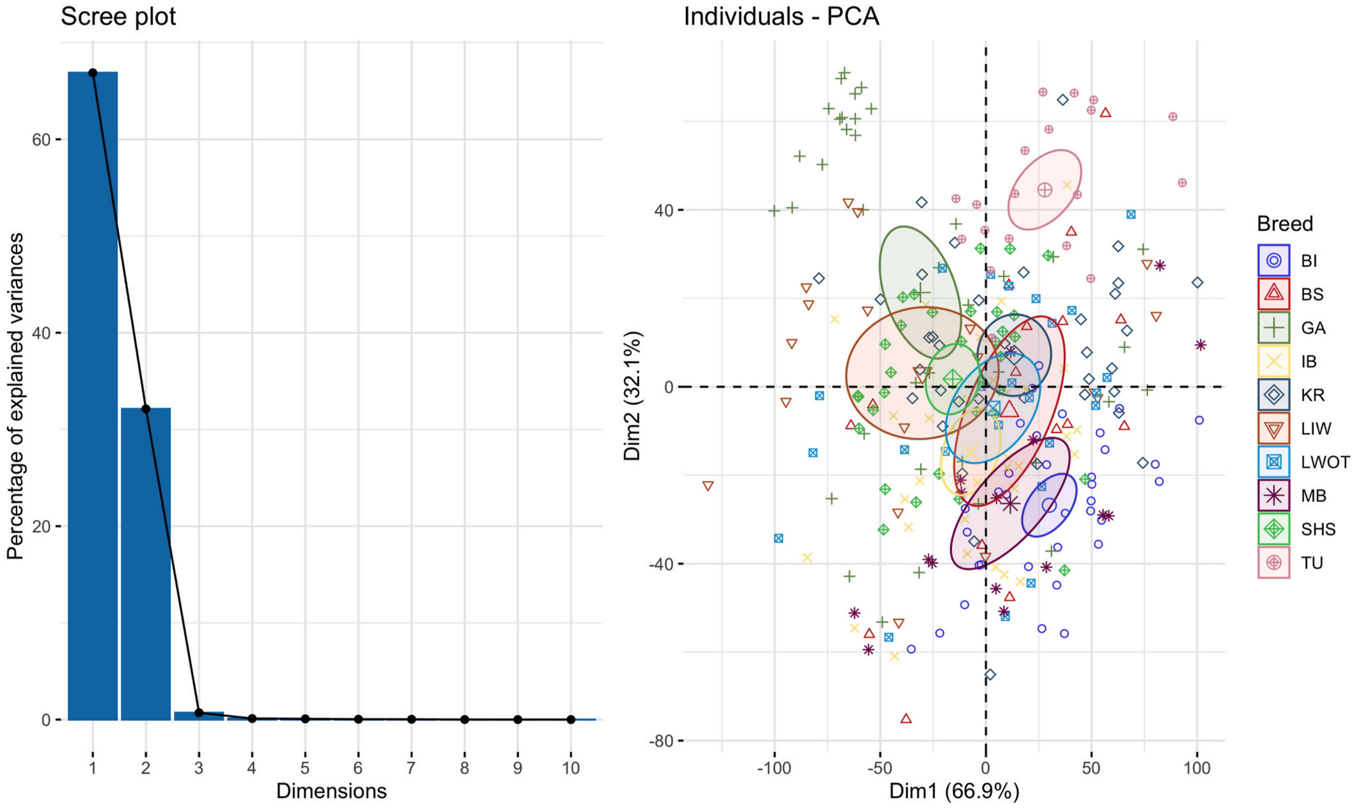


Fig. 1. Scree plot and scatterplot of two principal components (PCs) from the spectra of intact meat.

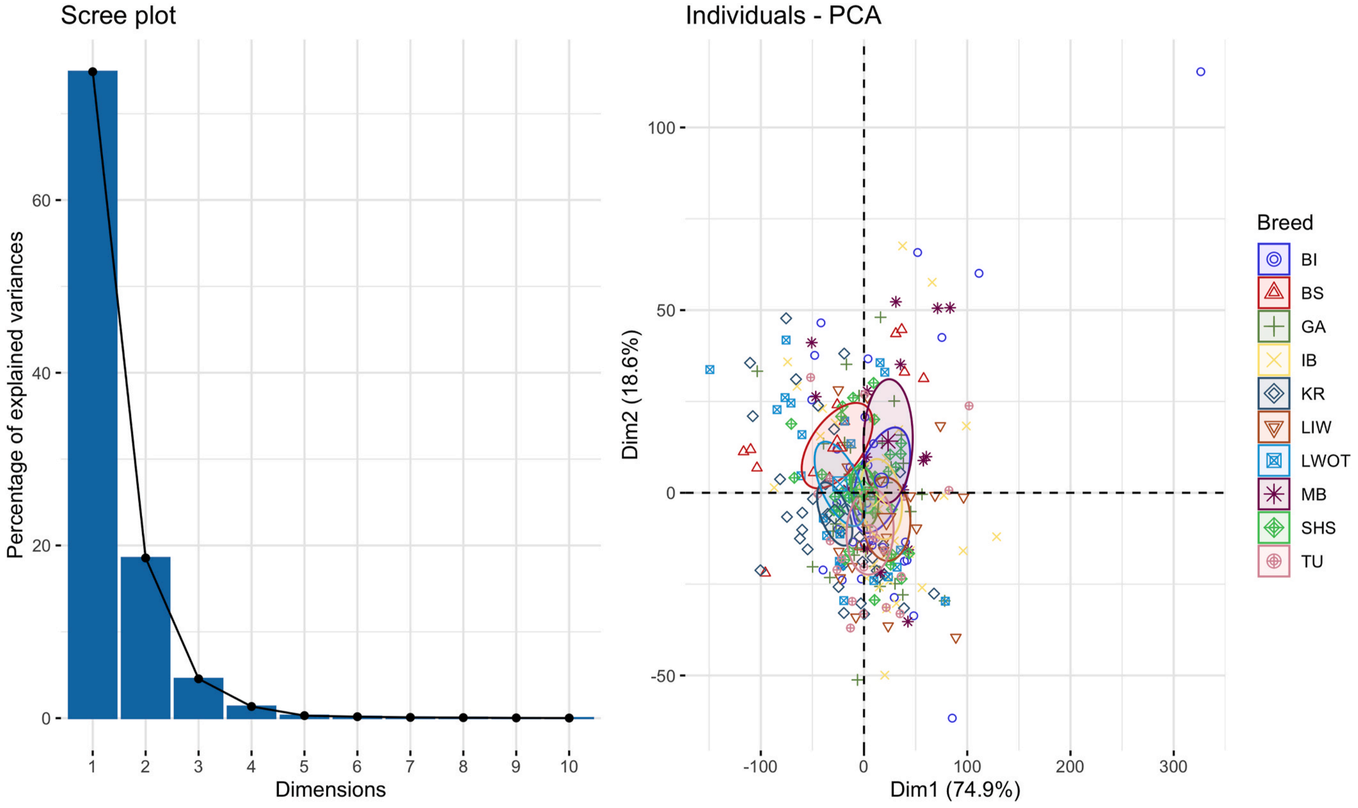


Fig. 2. Scree plot and scatterplot of two principal components (PCs) from the spectra of intact fat.

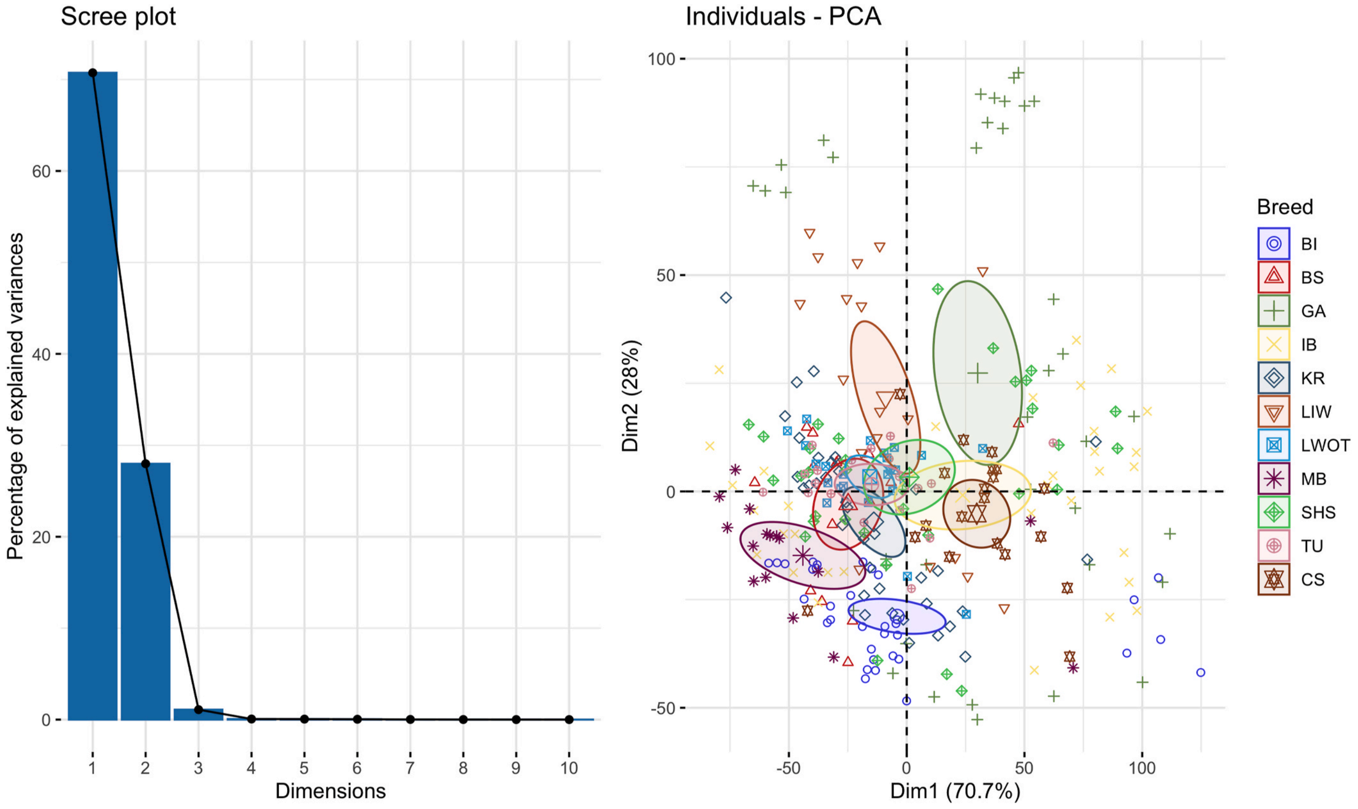


Fig. 3. Scree plot and scatterplot of two principal components (PCs) from the spectra of minced meat.

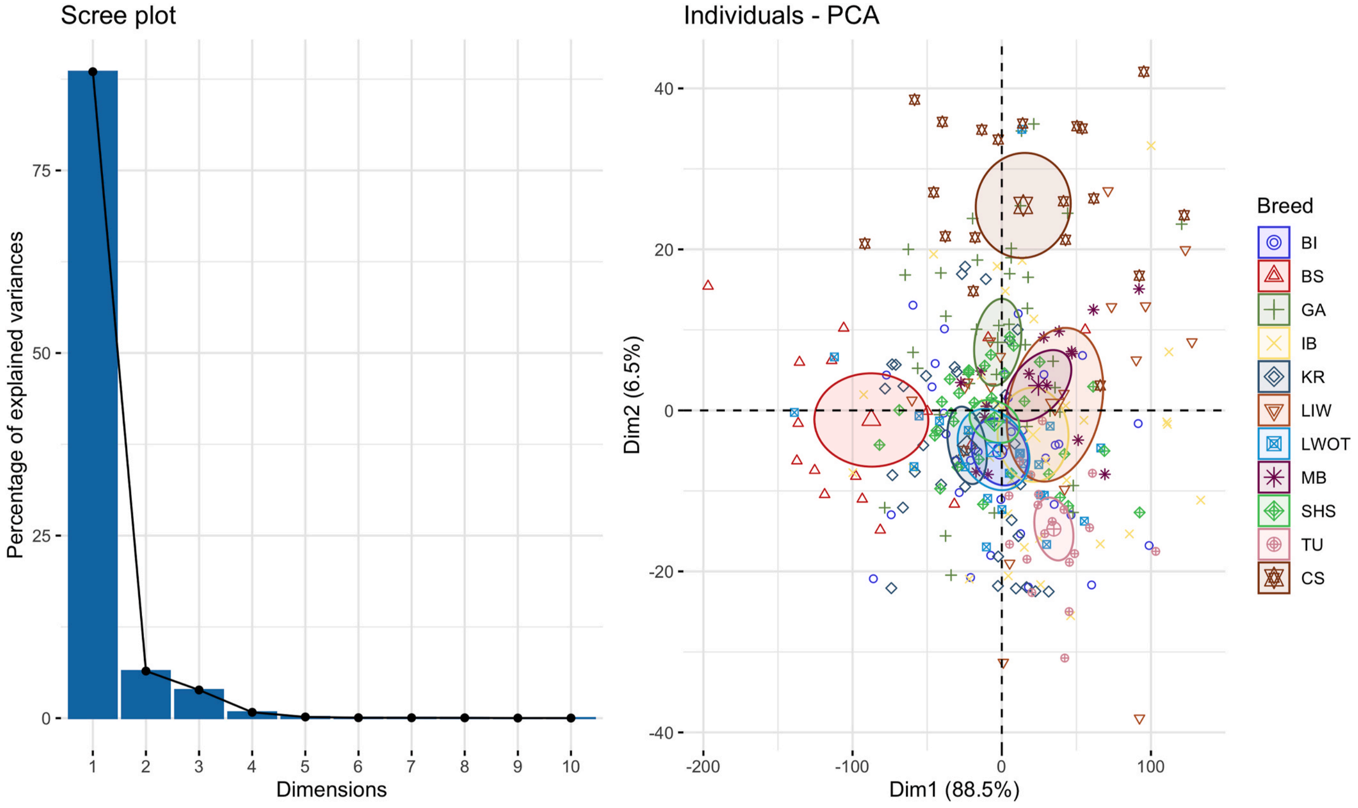


Fig. 4. Scree plot and scatterplot of two principal components (PCs) from the spectra of minced fat.

fat, respectively. In all cases, the first two PCs captured more than 93 % of the total variability, with the highest being observed for intact and minced meat (~99 %) and the lowest for intact fat (93.8 %), with minced fat at 94.9 %.

As shown in the scree plots (Figs. 1–4), which present 10 dimensions per newly defined PC on the x-axis, the first two PCs generally explained most of the variability in both meat and fat samples. However, in the case of intact fat spectra, the third PC contributed some additional information (6.4 %).

3.2. *Success of breed assignment in semi-supervised learning (semi-external cross-validation)*

Figs. 5 and 6 summarise model performance in terms of assignment success for intact and minced meat/fat tissues, respectively. Overall, intact fat performed better than intact meat (Fig. 5), although this tissue difference was less evident for minced samples (Fig. 6). Among the sample types, the assignment success was higher for the minced than the intact samples (on average 76.2 vs. 66.8, respectively). The final number of PCs retained for DAPC (Table 2) varied by sampling and sample type and ranged from 15 to 50. Nevertheless, the number of PCs selected had only a marginal effect on assignment success.

Using the complete dataset as a reference (100 % sampling) and considering intact meat (Fig. 5a), BI achieved the highest correct breed assignment rate (83.8 %), followed by SHS (81.4 %). Intermediate values (between 60 and 70 %) were obtained for TU, GA, LWOT, KR, and IB. Lower classification accuracy was observed for LIW, MB, and BS, with correct assignments ranging from 50.8 % for LIW to 34.6 % for BS.

Regarding intact fat (Fig. 5b), using the whole dataset (100 % sampling), both SHS and KR achieved ~93 % identification success, followed by GA, LIW, and BI (ranging between 83 and 88 %). Other breeds, such as MB and BS, also had medium to high classification accuracy (73.2 and 79.3 %, respectively). An intermediate classification accuracy was observed for TU and IB (~68 %), while the lowest successful identification rate was observed for LWOT (55.6 %).

In the case of minced meat (Fig. 6a), considering the whole dataset, the best assignment was obtained by CS (100 %), followed by SHS (89 %), BI (84.6 %), and ~80 % for KR and LWOT. Intermediate results were achieved by MB (72.8 %), IB (66.9 %), and LIW (64.7 %), while the lowest values were observed for BS, TU, and GA (52–57 %). For minced fat (Fig. 6b), SHS and BI had the highest percentage of correct breed assignment (97.2 and 98.8 %, respectively), followed by KR, MB, and CS (91.3, 88.7, and 80.8 %, respectively). The lowest values (between 60 % and 73 %) were found for TU, LWOT, BS, IB, LIW, and GA.

The DAPC scatterplots of the first two discriminant functions are provided in Supplementary materials (S5) for intact samples (meat and fat; panels a and b, respectively), and minced samples (meat and fat; panels c and d, respectively). These scatterplots show only the first two discriminant functions (LD1 and LD2), which do not capture the full multidimensional variance used by the classification model. In some cases, particularly for intact fat, the higher assignments for specific breeds of the heatmap were visually confirmed by the clustering pattern. However, breeds with high classification accuracy may still appear partially overlapping in the two-dimensional projections, and clear breed-specific separations were not consistently observed across sample types. Furthermore, the loading plots for LD1 and LD2 are provided in the supplementary materials (Figs. S6–S7), showing the relevant spectral regions (wavenumbers) for each tissue and physical structure (panels: a. intact meat, b. intact fat, c. minced meat, d. minced fat). In particular, for LD2, the top-loading spectral regions that contributed most to the discrimination are clearly highlighted.

Relatively to the assignment errors still evident in the heatmaps for 100 % sampling (Fig. 5a), in intact meat the breeds showing misclassification were always the same: BS assigned to LWOT (12.7–23.3 %), to MB (3.4–17.1 %) and to IB (5.8–11.9 %); IB was misclassified as SHS (from 9.5 to 22.1) and BI (8.2–15.5 %); KR was mainly assigned to SHS (from 14.7 to 28.4 %); MB samples were assigned to IB (13.4–29.4 %); and LIW samples to LWOT (6.8–28.0 %). In intact fat samples (Fig. 5b), misclassifications were less evident, with minor error rates in individual cases and low spread misclassification. However, in intact fat, there were also evident cases of breed overlap: IB was assigned to TU from 8.4 to 13.5 %, LWOT was assigned to GA (from 8.4 to 14.9 %), to SHS (6.3–10.7 %) and IB (5.9–17.7 %), MB was misclassified to GA from 11.9 to 20.2 %, and TU was assigned to KR between 16.5 and 26.0 %. Relatively to minced meat (Fig. 6a), BS was misclassified as TU (9.9–26.3 %), LWOT (6.7–12.1 %) and IB (5.2–17.8 %); GA was assigned to KR (6.4–23.2 %) and SHS (8.3–12.2 %); LIW samples were assigned to LWOT (9.6–17.4 %, excluding 50 % sampling); MB was misclassified as IB (9.0–20.6 %); TU as LWOT (10.9–24.3 %) and BS (7.8–18.2 %). In minced fat (Fig. 6b), the main misclassifications were linked to IB assigned to SHS in 13.4–29.7 % of the cases; BS was classified as GA (8–15.6 %), while GA was assigned to MB (6.4–16.2 %); LIW was misclassified to SHS (6.2–16.7 %); LWOT samples were wrongly assigned as SHS (10–15.4 %) and IB (11.4–17.5 %); and TU was mainly assigned to KR (14.1–20.9 %).

The DAPC heatmaps for intact (Fig. 5) and minced (Fig. 6) meat/fat tissues show the effect of sampling size on assignment success in the semi-external cross-validation scenario (Fig. 7). Considering intact meat, classification performed better for BI and SHS, considering an increased sample size from 80 to 100 %. A highly successful assignment was achieved, even with half of the data for BI, GA, LIW, SHS, and TU. However, an inconsistent trend was observed across breeds, with varying sample sizes for GA, IB, LIW, LOWT, KR, MB, and TU.

In intact fat, most breeds (BI, LIW, KR, and SHS) achieved the highest assignment success with a dataset ranging from 80 to 100 %. Even when using 50 % of the available dataset, BI, SHS, and KR achieved similar assignment success rates. Also, in the case of intact fat (as for intact meat) for most breeds (BS, IB, LIW, LWOT, and TU), an increasing sample size didn't suggest an increased trend in the assignment's success.

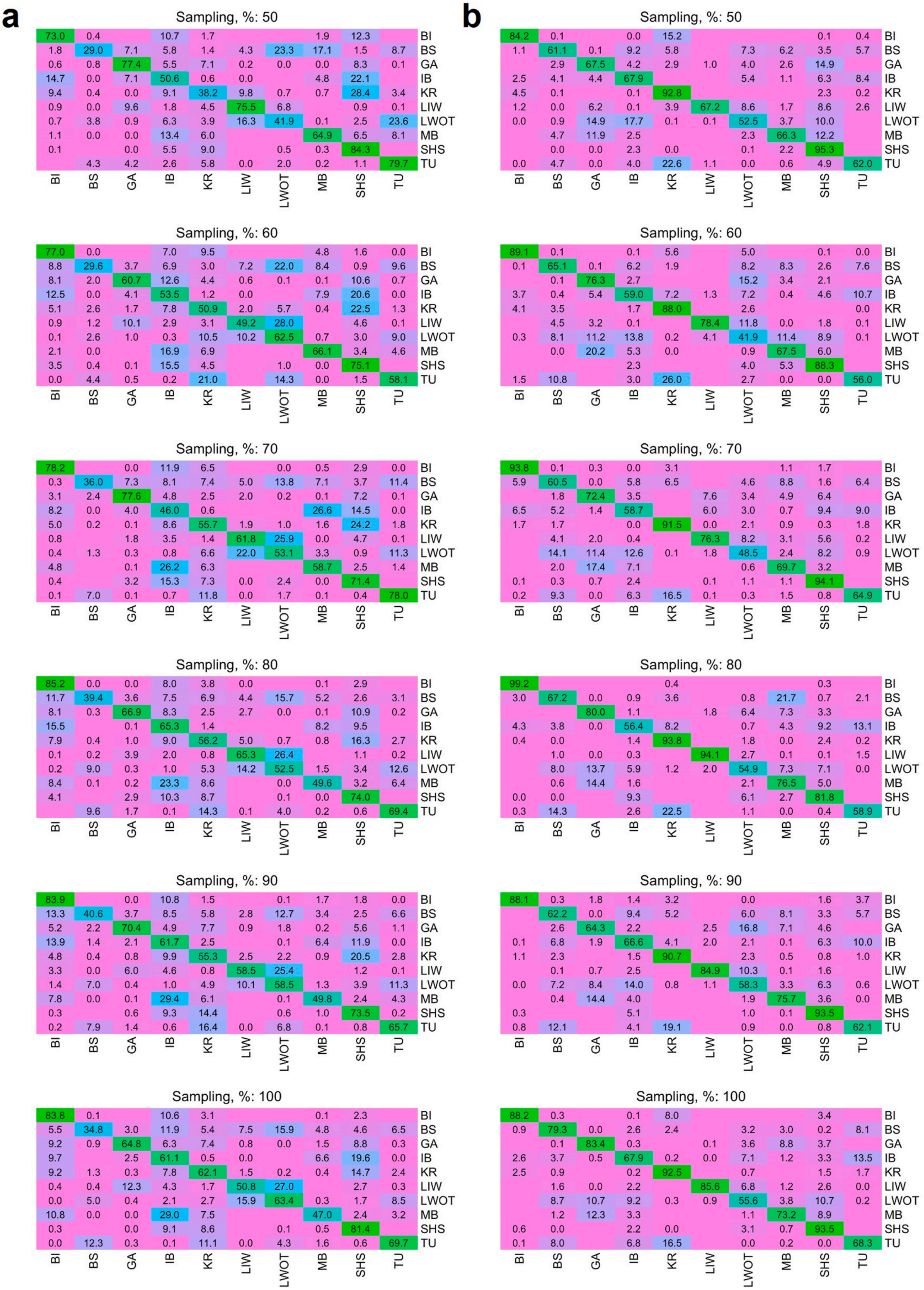
In minced meat, most of the breeds (BI, CS, LWOT, KR, MB, and SHS) had a successful classification (>70 %) with a sample size ranging from 80 to 100 %; even if BI, CS, LIW and had robust results with the use of 50 % of the data, while the intermediate sampling (60–70 %) had a lower success of assignment for most of the breeds. For minced fat, the best assignment was obtained by BI, KR, and MB using the dataset between 90 and 100 %. Some breeds, such as BI, KR, and MB, obtained robust assignments even with a small sample size (50 % of the data analysed), while three breeds (TU, GA, and BS) achieved the best assignments with 60 or 70 % of the dataset used. Also, in this case, some breeds, such as CS, IB, LIW, and LWT, showed inconsistent trends with increasing sample size. Overall, on tissue (meat and fat) and sampling presentation (intact and minced), the results from an increased sample size were not linear, and successful assignment was marginally affected.

3.3. *Success of breed assignment in unsupervised learning methods (external validation)*

The results of the external validation scenario are reported in Table 3 (intact meat and fat tissues) and Table 4 (minced meat and fat tissues). In all cases, irrespective of tissues and preparations, no breed was 100 % assigned to only one breed.

In intact samples (Table 2), results could be summarised as follows:

i) In meat tissues, MB and LIW showed the lowest dispersion: MB was assigned to IB in 75.0 % of the cases and to BI and TU in 18.8 % and 6.2 % of the cases. LIW was assigned to LWOT, GA, and IB at 64.3, 28.6, and 7.1 %, respectively. The BI samples were classified as IB in 58.8 % of the assignments. The remaining breeds were assigned to several breeds, with the highest percentage of assignment to a specific breed ranging from 48.5 % to 37.5 %.



a

Sampling, %: 50

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
73.0	0.4		10.7	1.7			1.9	12.3		BI
1.8	29.0	7.1	5.8	1.4	4.3	23.3	17.1	1.5	8.7	BS
0.6	0.8	77.4	5.5	7.1	0.2	0.0	0.0	8.3	0.1	GA
14.7	0.0	7.1	50.6	0.6	0.0		4.8	22.1		IB
9.4	0.4	0.0	9.1	38.2	9.8	0.7	0.7	28.4	3.4	KR
0.9	0.0	9.6	1.8	4.5	75.5	6.8		0.9	0.1	LIW
0.7	3.8	0.9	6.3	3.9	16.3	41.9	0.1	2.5	23.6	LWOT
1.1	0.0	0.0	13.4	6.0			64.9	6.5	8.1	MB
0.1		0.0	5.5	9.0		0.5	0.3	84.3		SHS
	4.3	4.2	2.6	5.8	0.0	2.0	0.2	1.1	79.7	TU

Sampling, %: 60

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
77.0	0.0		7.0	9.5			4.8	1.6	0.0	BI
8.8	29.6	3.7	6.9	3.0	7.2	22.0	8.4	0.9	9.6	BS
8.1	2.0	60.7	12.6	4.4	0.6	0.1	0.1	10.6	0.7	GA
12.5	0.0	4.1	53.5	1.2	0.0		7.9	20.6	0.0	IB
5.1	2.6	1.7	7.8	50.9	2.0	5.7	0.4	22.5	1.3	KR
0.9	1.2	10.1	2.9	3.1	49.2	28.0		4.6	0.1	LIW
0.1	2.6	1.0	0.3	10.5	10.2	62.5	0.7	3.0	9.0	LWOT
2.1	0.0		16.9	6.9			66.1	3.4	4.6	MB
3.5	0.4	0.1	15.5	4.5		1.0	0.0	75.1		SHS
0.0	4.4	0.5	0.2	21.0		14.3	0.0	1.5	58.1	TU

Sampling, %: 70

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
78.2		0.0	11.9	6.5		0.0	0.5	2.9	0.0	BI
0.3	36.0	7.3	8.1	7.4	5.0	13.8	7.1	3.7	11.4	BS
3.1	2.4	77.6	4.8	2.5	2.0	0.2	0.1	7.2	0.1	GA
8.2	0.0	4.0	46.0	0.6			26.6	14.5	0.0	IB
5.0	0.2	0.1	8.6	55.7	1.9	1.0	1.6	24.2	1.8	KR
0.8		1.8	3.5	1.4	61.8	25.9	0.0	4.7	0.1	LIW
0.4	1.3	0.3	0.8	6.6	22.0	53.1	3.3	0.9	11.3	LWOT
4.8		0.1	26.2	6.3			58.7	2.5	1.4	MB
0.4		3.2	15.3	7.3	0.0	2.4	0.0	71.4		SHS
0.1	7.0	0.1	0.7	11.8	0.0	1.7	0.1	0.4	78.0	TU

Sampling, %: 80

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
85.2	0.0	0.0	8.0	3.8	0.0		0.1	2.9		BI
11.7	39.4	3.6	7.5	6.9	4.4	15.7	5.2	2.6	3.1	BS
8.1	0.3	66.9	8.3	2.5	2.7	0.0	0.1	10.9	0.2	GA
15.5		0.1	65.3	1.4			8.2	9.5		IB
7.9	0.4	1.0	9.0	56.2	5.0	0.7	0.8	16.3	2.7	KR
0.1	0.2	3.9	2.0	0.8	65.3	26.4		1.1	0.2	LIW
0.2	9.0	0.3	1.0	5.3	14.2	52.5	1.5	3.4	12.6	LWOT
8.4	0.1	0.2	23.3	8.6		0.0	49.6	3.2	6.4	MB
4.1		2.9	10.3	8.7		0.1	0.0	74.0		SHS
	9.6	1.7	0.1	14.3	0.1	4.0	0.2	0.6	69.4	TU

Sampling, %: 90

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
83.9		0.0	10.8	1.5		0.1	1.7	1.8	0.0	BI
13.3	40.6	3.7	8.5	5.8	2.8	12.7	3.4	2.5	6.6	BS
5.2	2.2	70.4	4.9	7.7	0.9	1.8	0.2	5.6	1.1	GA
13.9	1.4	2.1	61.7	2.5		0.1	6.4	11.9	0.0	IB
4.8	0.4	0.8	9.9	55.3	2.5	2.2	0.9	20.5	2.8	KR
3.3	0.0	6.0	4.6	0.8	58.5	25.4		1.2	0.1	LIW
1.4	7.0	0.4	1.0	4.9	10.1	58.5	1.3	3.9	11.3	LWOT
7.8	0.0	0.1	29.4	6.1		0.1	49.8	2.4	4.3	MB
0.3		0.6	9.3	14.4		0.6	1.0	73.5	0.2	SHS
0.2	7.9	1.4	0.6	16.4	0.0	6.8	0.1	0.8	65.7	TU

Sampling, %: 100

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
83.8	0.1		10.6	3.1			0.1	2.3		BI
5.5	34.8	3.0	11.9	5.4	7.5	15.9	4.8	4.6	6.5	BS
9.2	0.9	64.8	6.3	7.4	0.8	0.0	1.5	8.8	0.3	GA
9.7		2.5	61.1	0.5	0.0		6.6	19.6	0.0	IB
9.2	1.3	0.3	7.8	62.1	1.5	0.2	0.4	14.7	2.4	KR
0.4	0.4	12.3	4.3	1.7	50.8	27.0		2.7	0.3	LIW
0.0	5.0	0.4	2.1	2.7	15.9	63.4	0.3	1.7	8.5	LWOT
10.8	0.0	0.0	29.0	7.5			47.0	2.4	3.2	MB
0.3		0.0	9.1	8.6		0.1	0.5	81.4		SHS
0.0	12.3	0.3	0.1	11.1	0.0	4.3	1.6	0.6	69.7	TU

b

Sampling, %: 50

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
84.2	0.1		0.0	15.2				0.1	0.4	BI
1.1	61.1	0.1	9.2	5.8		7.3	6.2	3.5	5.7	BS
	2.9	67.5	4.2	2.9	1.0	4.0	2.6	14.9		GA
2.5	4.1	4.4	67.9			5.4	1.1	6.3	8.4	IB
4.5	0.1		0.1	92.8				2.3	0.2	KR
1.2	0.0	6.2	0.1	3.9	67.2	8.6	1.7	8.6	2.6	LIW
0.0	0.9	14.9	17.7	0.1	0.1	52.5	3.7	10.0		LWOT
	4.7	11.9	2.5			2.3	66.3	12.2		MB
	0.0	0.0	2.3	0.0		0.1	2.2	95.3		SHS
0.0	4.7	0.0	4.0	22.6	1.1	0.0	0.6	4.9	62.0	TU

Sampling, %: 60

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
89.1	0.1		0.1	5.6		5.0		0.1	0.0	BI
0.1	65.1	0.1	6.2	1.9		8.2	8.3	2.6	7.6	BS
	0.1	76.3	2.7			15.2	3.4	2.1		GA
3.7	0.4	5.4	59.0	7.2	1.3	7.2	0.4	4.6	10.7	IB
4.1	3.5		1.7	88.0		2.6			0.0	KR
	4.5	3.2	0.1		78.4	11.8	0.0	1.8	0.1	LIW
0.3	8.1	11.2	13.8	0.2	4.1	41.9	11.4	8.9	0.1	LWOT
	0.0	20.2	5.3	0.0		0.9	67.5	6.0		MB
			2.3			4.0	5.3	88.3		SHS
1.5	10.8		3.0	26.0		2.7	0.0	0.0	56.0	TU

Sampling, %: 70

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
93.8	0.1	0.3	0.0	3.1			1.1	1.7		BI
5.9	60.5	0.0	5.8	6.5		4.6	8.8	1.6	6.4	BS
	1.8	72.4	3.5		7.6	3.4	4.9	6.4		GA
6.5	5.2	1.4	58.7		6.0	3.0	0.7	9.4	9.0	IB
1.7	1.7		0.0	91.5	0.0	2.1	0.9	0.3	1.8	KR
	4.1	2.0	0.4		76.3	8.2	3.1	5.6	0.2	LIW
	14.1	11.4	12.6	0.1	1.8	48.5	2.4	8.2	0.9	LWOT
	2.0	17.4	7.1			0.6	69.7	3.2		MB
0.1	0.3	0.7	2.4		0.1	1.1	1.1	94.1		SHS
0.2	9.3	0.0	6.3	16.5	0.1	0.3	1.5	0.8	64.9	TU

Sampling, %: 80

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
99.2				0.4				0.3		BI
3.0	67.2	0.0	0.9	3.6		0.8	21.7	0.7	2.1	BS
		80.0	1.1		1.8	6.4	7.3	3.3		GA
4.3	3.8	0.0	56.4	8.2		0.7	4.3	9.2	13.1	IB
0.4	0.0		1.4	93.8		1.8	0.0	2.4	0.2	KR
	1.0	0.0	0.3		94.1	2.7	0.1	0.1	1.5	LIW
	8.0	13.7	5.9	1.2	2.0	54.9	7.3	7.1	0.0	LWOT
	0.6	14.4	1.6			2.1	76.5	5.0		MB
0.0	0.0		9.3			6.1	2.7	81.8		SHS
0.3	14.3		2.6	22.5		1.1	0.0	0.4	58.9	TU

Sampling, %: 90

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
88.1	0.3	1.8	1.4	3.2		0.0		1.6	3.7	BI
	62.2	0.0	9.4	5.2		6.0	8.1	3.3	5.7	BS
	2.6	64.3	2.2		2.5	16.8	7.1	4.6		GA
0.1	6.8	1.9	66.6	4.1	2.0	2.1	0.1	6.3	10.0	IB
1.1	2.3		1.5	90.7		2.3	0.5	0.8	1.0	KR
	0.1	0.7	2.5		84.9	10.3	0.1	1.6		LIW
0.0	7.2	8.4	14.0	0.8	1.1	58.3	3.3	6.3	0.6	LWOT
	0.4	14.4	4.0			1.9	75.7	3.6	0.0	MB
0.3			5.1			1.0	0.1	93.5		SHS
0.8	12.1		4.1	19.1		0.9	0.0	0.8	62.1	TU

Sampling, %: 100

BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	
88.2	0.3		0.1	8.0				3.4		BI
0.9	79.3	0.0	2.6	2.4		3.2	3.0	0.2	8.1	BS
	0.1	83.4	0.3		0.1	3.6	8.8	3.7		GA
2.6	3.7	0.5	67.9	0.2	0.0	7.1	1.2	3.3	13.5	IB
2.5	0.9		0.2	92.5		0.7		1.5	1.7	KR
	1.6	0.0	2.2		85.6	6.8	1.2	2.6	0.0	LIW
	8.7	10.7	9.2	0.3	0.9	55.6	3.8	10.7	0.2	LWOT
	1.2	12.3	3.3			1.1	73.2	8.9		MB
0.6	0.0		2.2	0.0		3.1	0.7	93.5		SHS
0.1	8.0		6.8	16.5		0.0	0.2	0.0	68.3	TU

Fig. 5. Heatmap of the DAPC with semi-external cross-validation assignment (%) per breed considering intact meat (a) and fat (b); y-axes show the observed and x-axes the predicted breed.

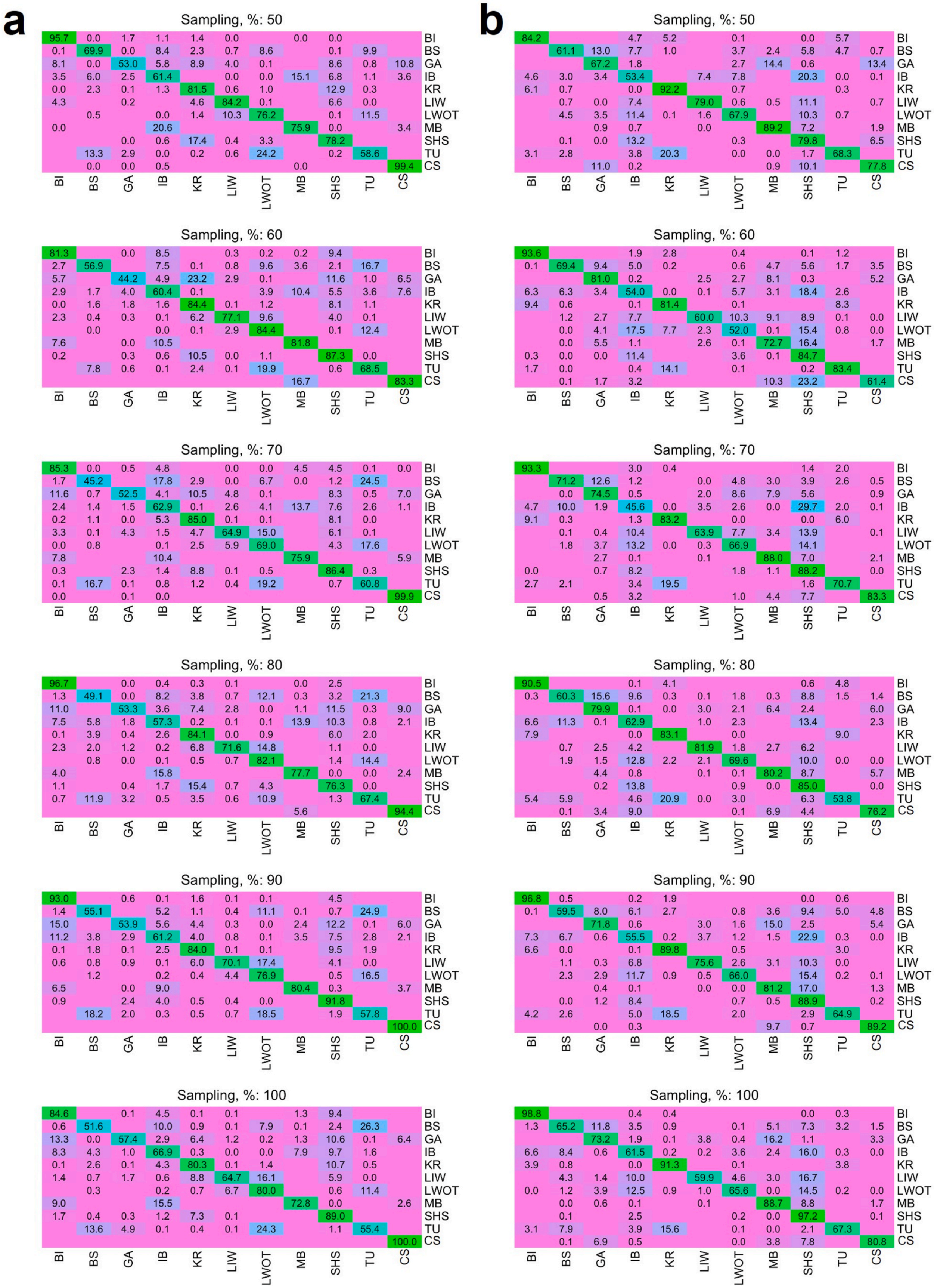


Fig. 6. Heatmap of the DAPC with semi-external cross-validation assignment (%) per breed considering minced meat (a) and fat (b); y-axes show the observed and x-axes the predicted breed.

Table 2
Average number of PCs used by DAPC for different tissues (meat and fat), physical structure (intact and minced), and sampling (from 50 to 100 %).

Tissue	Physical structure	Sampling (%)					
		50	60	70	80	90	100
meat	intact	24	26	28	28	29	30
fat		33	36	34	36	36	36
meat	minced	26	28	26	27	26	25
fat		30	30	30	31	30	29

ii) In fat tissues, only BI reported a high percentage of assignment (about 60 %) with one breed (SHS), followed by KR, which was classified as TU in 50.0 % of the cases. Other breeds were assigned to one breed, with the highest values ranging from 30.4 % to 44.7 %.

The external success of the assignment for minced tissues (Table 3) was as follows:

i) In meat tissues, LWOT samples were classified as TU in 70.8 % of the cases, while other breeds (BS, KR, MB, CS, and TU) were assigned to one breed in 50–56.3 %. Among those, MB was assigned to three breeds (BI, IB, and CS), while BI and LWOT were assigned to four breeds. Among the remaining breeds, IB, LIW, and SHS were assigned to other breeds at lower percentages (~27.3 %), while GA was assigned as LIW in 42.1 % of the cases.

ii) In fat tissues, four breeds (BI, GA, SHS, TU, and CS) were assigned to other breeds, with the maximum obtained values ranging from 60 % to 70 %. The MB and KR were similar to SHS (56.3 %) and TU (51.4 %). Other breeds (BS, IB, LIW, and LWOT) were assigned to one breed, with the highest values ranging from 33 % to 41.7 %.

Interestingly, for many breeds, their assignment to other breeds differed according to the sample tissue, meat or fat. This was observed for intact and minced samples (Tables 3 and 4). For example, using intact meat, MB was assigned to IB at 75 %, BI to IB at 58.8 %, LIW to LWOT at 64.3 %. However, MB was assigned to SHS and GA in intact fat at 37.5 % and 25 %, respectively, BI to SHS at 60.6 %, and LIW to IB at 27.8 %. Using minced meat, CS was assigned to MB by 5 %, TU to LWOT by 53 %, and none to KR. In contrast, 70 % of the CS was assigned to MB in minced fat, and 68 % of the TU was assigned to KR without any assignment to LWOT.

4. Discussion

The ability of FT-NIRS to identify or discriminate biological samples (breeds in this case) is based both on direct spectral information, that is, the vibrational responses of the frequencies produced by the functional

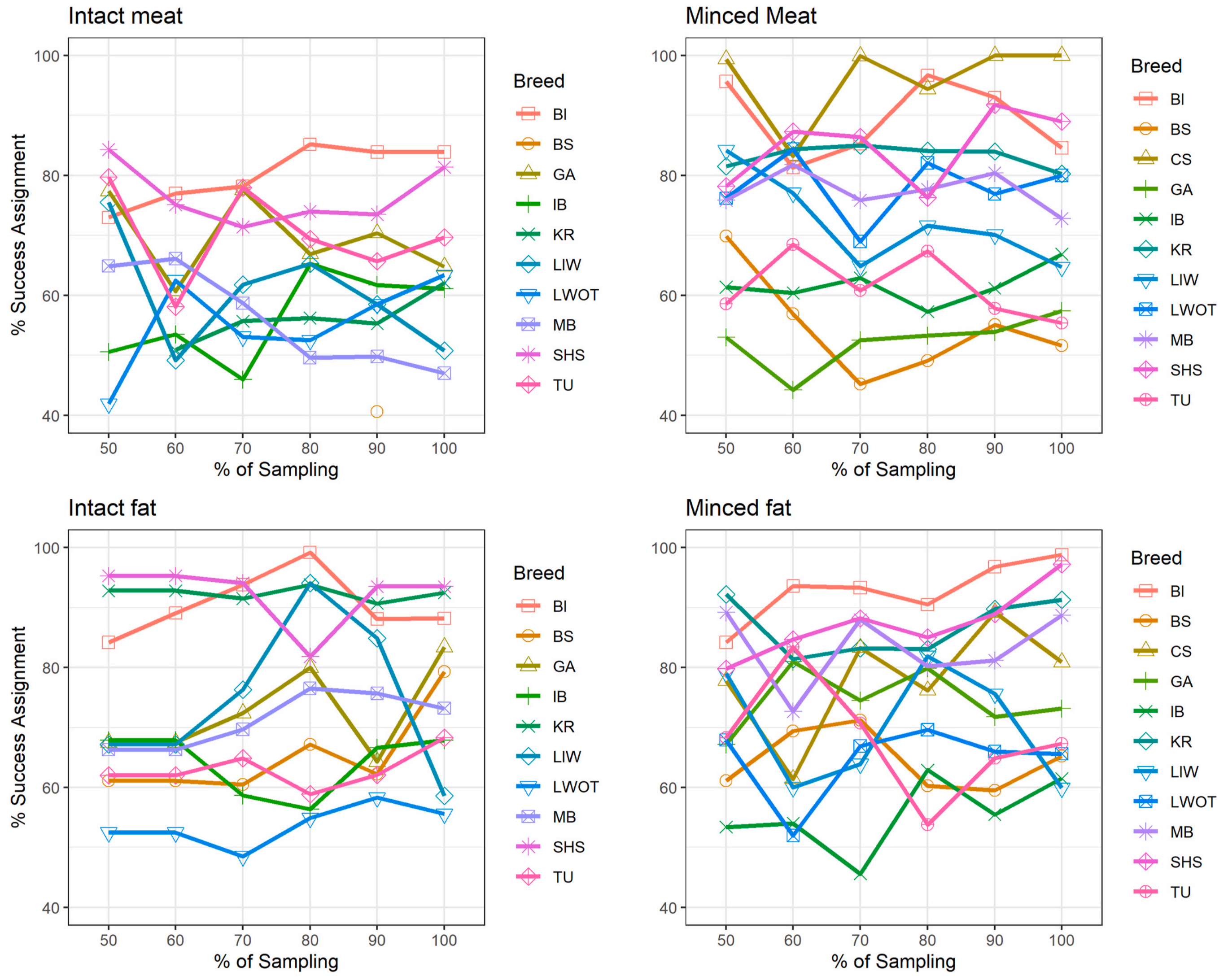


Fig. 7. Sample size effect in semi-external cross-validation assignment (%).

Table 3
Breed assignment success (%) in intact meat and fat using external validation.

Meat	BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	Total %
BI	–	0	0	58.8	17.7	0	0	14.7	8.8	0	100
BS	6.3	–	6.3	12.5	6.3	6.3	37.5	6.3	0	18.8	100
GA	2.6	5.3	–	18.4	15.8	42.1	0	2.6	13.2	0	100
IB	21.2	0	3.0	–	3.0	0	0	33.3	39.4	0	100
KR	11.1	5.6	0	11.1	–	8.3	0	5.6	41.7	16.7	100
LIW	0	0	28.6	7.1	0	–	64.3	0	0	0	100
LWOT	0	12.5	0	0	4.2	54.2	–	0	4.2	25.0	100
MB	18.8	0	0	75.0	0	0	0	–	0	6.3	100
SHS	3.0	0	3.0	48.5	45.5	0	0	0	–	0	100
TU	0	15.8	5.3	0	47.4	0	26.3	5.3	0	–	100
Fat	**BI**	**BS**	**GA**	**IB**	**KR**	**LIW**	**LWOT**	**MB**	**SHS**	**TU**	**Total %**
BI	–	0	0	6.1	27.3	0	0	0	60.6	6.1	100
BS	0	–	0	18.8	6.3	0	18.8	25.0	0	31.3	100
GA	0	2.6	–	21.1	0	2.6	44.7	26.3	2.6	0	100
IB	9.4	12.5	6.3	–	3.1	3.1	37.5	3.1	25.0	0	100
KR	33.3	11.1	0	0	–	0	2.8	2.8	0	50.0	100
LIW	0	0	5.6	27.8	0	–	11.1	0	38.9	16.7	100
LWOT	0	8.7	13.0	26.1	4.4	4.4	–	13.0	30.4	0	100
MB	0	12.5	25.0	6.3	0	0	18.8	–	37.5	0	100
SHS	21.2	0	0	15.2	0	3.0	24.2	36.4	–	0	100
TU	0	36.8	0	15.8	42.1	0	0	0	5.3	–	100

Legend: Predicted breeds are reported on the rows.

Table 4
Breed assignment success (%) in minced meat and fat using external validation.

Meat	BI	BS	GA	IB	KR	LIW	LWOT	MB	SHS	TU	CS	Total %
BI	–	0	17.7	47.1	0	0	0	14.7	20.6	0	0	100
BS	0	–	0	6.3	6.3	0	18.8	0	12.5	56.3	0	100
GA	13.2	0	–	10.5	13.2	42.1	0	0	18.4	0	2.6	100
IB	27.3	3.0	15.2	–	3.0	0	6.1	27.3	18.2	0	0	100
KR	0	2.8	2.8	2.8	–	25.0	0	0	55.6	11.1	0	100
LIW	5.6	0	27.8	0	22.2	–	5.6	0	22.2	16.7	0	100
LWOT	0	8.3	0	0	4.2	16.7	–	0	0	70.8	0	100
MB	50	0	0	37.5	0	0	0	–	0	0	12.5	100
SHS	6.1	0	24.2	9.1	18.2	15.2	27.3	0	–	0	0	100
TU	0	26.3	10.5	0	0	5.3	52.6	0	5.3	–	0	100
CS	0	10.0	50.0	25.0	0	0	0	5.0	5.0	5.0	–	100
Fat	**BI**	**BS**	**GA**	**IB**	**KR**	**LIW**	**LWOT**	**MB**	**SHS**	**TU**	**CS**	**Total %**
BI	–	2.9	0	2.9	67.7	0	0	0	11.8	14.7	0	100
BS	0	–	13.3	33.3	0	0	6.7	13.3	13.3	13.3	6.7	100
GA		0	–	29.7	0	2.7	59.5	5.4	0	0	2.7	100
IB	9.4	9.4	3.1	–	0	0	37.5	3.1	31.3	3.1	3.1	100
KR	22.9	5.7	0	2.9	–	0	17.1	0	0	51.4	0	100
LIW	0	0	29.4	5.9	0	–	35.3	11.8	17.7	0	0	100
LWOT	0	0	41.7	25.0	4.2	8.3	–	0	20.8	0	0	100
MB	0	0	18.8	0	0	0	0	–	56.3	0	25.0	100
SHS	0	3.0	3.0	66.7	0	0	6.1	9.1	–	0	12.1	100
TU	10.5	10.5	0	5.3	68.4	0	0	0	5.3	–	0	100
CS	0	0	20.0	0	0	0	0	70.0	10.0	0	–	100

Legend: Predicted breeds are reported on the rows.

groups, and on the indirect interaction of sample properties distributed over a large wavenumber range that allows spectral information to pass (Pasquini, 2018). In this study, the information and structure of spectral sources and their combinations were exploited by summarising informative spectral variability into a small number of PCs, which were further analysed for pattern recognition using discriminant analysis. From the results, a high variability was mostly captured by two PCs for different tissues and physical structures (intact and minced). However, this variability within and among breeds, combined with a large amount of data, is considered a positive feature in chemometric applications because it improves the accuracy of the prediction models (Alomar et al., 2003).

The comparison between tissues showed that fat performed better than meat in the assignment's success, but this was not universal across breeds. Similar results were reported by Horcada et al. (2020), with better classification of fat spectra (about 75 % of samples) than of meat spectra (around 65 % of samples). The different complexities and constitutive homogeneities characterising the two tissues could explain their differences in the optical response. Furthermore, the higher discrimination power of fat over meat could be linked to the content of myofibrillar proteins of meat samples, especially in intact muscles, which probably interfere with the NIR absorption of light along the fibres and increase the noise in the spectra. Indeed, Prieto et al. (2009) suggested that myofibrils, or muscle fibres, can behave as optical fibres, conducting light along their fibrils and causing reflections.

Differences in sample presentation (intact or minced) affected the results, with lower assignment success for intact samples, especially in meat. It is known that the types of cells and the physical and structural traits of tissues provide spectral data that are more or less informative. Also, the presence of pigments, connective tissue, and intra- or extracellular fluids, as well as moisture, can be considered interfering elements (Dadousis et al., 2022). Previous studies have reported better NIRS performance in minced samples than in intact samples for predicting chemical composition and some technological parameters (Gvozdanović et al., 2019; Muñoz et al., 2018, 2019; Ortiz et al., 2020). This is due to differences in absorbance capacity (macroscopic surface reflectance) and sample homogeneity: intact samples are certainly more heterogeneous than minced ones. Indeed, previous studies have reported that grinding can affect the chemical and physical characteristics of the sample, including myoglobin and protein precipitation in the sarcoplasm, fibre organisation, myofibrillar birefringence, sarcomere length, moisture, and intramuscular fat.

Assignment success with semi-external cross-validation using NIRS data from intact meat samples (the whole dataset sampling) of at least 83 % was achieved for only 2 breeds (BI and SHS). From the comparison with a previous study, it emerged that considering only two genetic types (purebred Iberian and Iberian-Duroc crossbred), Horcada et al. (2020) found a lower correct classification of IB between 61 and 63 % using fresh, intact meat. On the contrary, Prieto et al. (2015) correctly classified intact pig meat with 94 % success according to three pig breeds, not all autochthonous (Lacombe, Duroc, Iberian), based on partial least squares discriminant analysis of the spectra. Regarding intact fat, in this study, the success of breed assignments was higher than 83 % for 5 breeds, surpassing the results reported by Horcada et al. (2020), who found about 75 % of samples classified as IB or crossbred based on subcutaneous fresh intact fat spectra. When considering minced samples, correct assignments with more than 80 % were achieved for five breeds, both in minced meat and in fat. Despite the relatively small sample size from a statistical perspective, encouraging results were obtained for BI, CS, and SHS in both minced meat and tissues, suggesting they may be considered for technical applications. Alomar et al. (2003) reported a correct assignment of 78 % using minced meat samples from two bovine breeds. McDevitt et al. (2005) reported a 96.5 % classification success rate across three chicken genotypes using dry-ground chicken carcasses. As anticipated above, it can be noted that compared to previous studies (Alomar et al., 2003; Horcada et al., 2020; McDevitt et al., 2005) that examined breed classification based either on meat and fat spectral data, these results referred to a greater number of breeds, which means both a greater chance of classification and also a greater risk of misclassification. Considering the DAPC misclassification, lean tissue resulted in higher errors, especially in intact samples. However, some misclassifications can be explained by genomic similarity between some breeds. For example, IB was assigned to TU in intact fat, and this relationship was previously reported by Dadousis et al. (2022) using genomic data. Other misclassifications, such as MB in IB and LIW in LWOT, observed in meat samples could also be attributed to common ancestry. For example, the MB breed was assigned to IB in meat samples (both intact and minced), likely due to its shared origin with both, both being part of the Southern European nucleus population (Muñoz et al., 2018). The closeness between LIW and LWOT was not surprising, given the sharing of parts of the genome linked to phenotypic characteristics and the origin of Lithuanian White pigs (Muñoz et al., 2018), as both breeds were previously raised in the same geographic area. Some relationships among them could be linked to uncontrolled mating and the high probability of sharing common ancestors, as suggested by Gvozdanović et al. (2019) and Muñoz et al. (2018), as well as to similarities among traditional rearing systems. In addition to genetic factors, the feeding system and individual animal variability can influence correct assignment, as reported by Horcada et al. (2020) and McDevitt et al. (2005) for Iberian pigs and chickens, respectively.

Regarding the effect of sample size, this study found no linearity within each breed, indicating that increasing the size of the TRN set did not necessarily improve assignment results. In a study using genomic data, from the same pig breeds as this study (Dadousis et al., 2022), found that an increased sample size yielded higher mean accuracies in correctly classifying pigs to their breed of origin. However, in that genomic study, the dataset consisted of ~1200 individual pigs, potentially capturing more variability and allowing better assessment of the effect of sample size.

Using external validation, this study aimed to assess a real case scenario in which samples from a specific breed, excluded from the training set, were assigned to the remaining 10 breeds in the validation dataset. Meat samples showed a higher percentage of re-assignment with at least one specific breed than fat tissue. According to this scenario, the greater dispersion of the values and the lower similarity between the spectra of breeds suggested that fat fits the models better than meat and could be more useful for breed discrimination using NIRS (Pasquini, 2018). However, as previously reported, the higher difference between intact tissues was probably linked to the higher heterogeneity of the muscle. Furthermore, the same trend observed in the intact samples also emerged in the minced samples: the results for the MB pigs were similar to those for IB, and LIW was similar to LWOT. According to previous genomic research on pigs of the same breeds (Dadousis et al., 2022; Muñoz et al., 2019), the classification seems consistent with the importance of the historical-geographical area and the common ancestors and factors indirectly linked to the rearing system. In this study, all pigs belonging to autochthonous breeds were often reared in a traditional free-range system; the feed composition differed among systems (breeds, thereby contributing to spectral absorbance features). Previous literature reported the potential NIRS application in the classification based on multiple factors: Sun et al. (2012) showed a 100 % correct classification of lamb meat by geographical origin, referring to pastoral and agricultural region samples, including differences in location, feeding condition (pasture or concentrate), breeds, and soil characteristics. Nevertheless, the results of the research report on unique factors, such as feeding regimes, were contradictory. Prieto et al. (2015) reported no success in pork sample classification based on feeding, while Zamora-Rojas et al. (2012) reported that NIRS technology was able to successfully classify Iberian pig carcasses according to feeding regime (>90 %) in an extreme situation (acorn versus feed).

5. Conclusions

In conclusion, coupling FT-NIRS data with DAPC on fat samples yielded more accurate assignments across all breeds and statistical methods. A higher success rate of breed assignment was obtained for minced samples than for intact samples, especially in meat. From a practical point of view, these results could translate into the potential to effectively differentiate local production, with an assignment success rate above 80 % for specific autochthonous pig breeds. The alternative possibility of directly extracting spectral absorbance information from fresh fat, which is easier to obtain from a carcass without affecting muscle tissue or cut integrity, seems feasible for some breeds. The potential to develop tools capable of discriminating local production emerges, thereby increasing consumer trust in small-scale, autochthonous pig breed products.

The results of this study could be implemented by comparing commercial pig breeds, highlighting spectral differences due to tissue characteristics, and providing a comprehensive overview. Further studies could explore the possibility of developing a robust model for discriminating and predicting the origin of pig fat or meat by integrating additional factors, such as breed, diet, age, and geographical area, into NIRS, applying more advanced machine learning methods, and including more samples.

CRediT authorship contribution statement

S. Parrini: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **C. Dadousis:** Visualization, Validation, Software, Methodology, Data curation, Conceptualization, Formal analysis, Writing – original draft. **F. Sirtori:**

Visualization, Investigation, Resources. **M.C. Fabbri:** Formal analysis, Data curation, Methodology. **M. Čandek-Potokar:** Resources, Project administration, Funding acquisition, Writing – review & editing. **J.M. Garcia-Casco:** Resources, Project administration, Funding acquisition, Writing – review & editing. **B. Lebret:** Resources, Project administration, Funding acquisition, Supervision, Writing – review & editing. **R. Nieto:** Visualization, Resources, Project administration, Funding acquisition, Writing – review & editing. **C. Aquilani:** Writing – review & editing, Visualization, Investigation. **R. Bozzi:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Data availability statement

None of the data was deposited in an official repository. The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

Statement on the use of generative AI and AI-assisted technologies in the writing process

No generative AI or AI-assisted technologies were used during the preparation of this work.

Funding sources

This study received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 634476 (project acronym TREASURE). The content of this paper reflects only the author's view, and the European Union Agency is not responsible for any use that may be made of the information it contains.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biosystemseng.2025.104366.

References

Acciaro, M., Decandia, M., Sitzia, M., Manca, C., Giovanetti, V., Cabiddu, A., Addis, M., Rassu, S. P. G., Molle, G., & Dimauro, C. (2020). Discriminant analysis as a tool to identify bovine and ovine meat produced from pasture or stall-fed animals. *Italian Journal of Animal Science, 19*(1), 1065–1070. https://doi.org/10.1080/1828051X.2020.1816507

Alomar, D., Gallo, C., Castañeda, M., & Fuchslocher, R. (2003). Chemical and discriminant analysis of bovine meat by near infrared reflectance spectroscopy (NIRS). *Meat Science, 63*(4), 441–450. https://doi.org/10.1016/S0309-1740(02)00101-8

Alonso, M. E., González-Montaña, J. R., & Lomillos, J. M. (2020). Consumers' concerns and perceptions of farm animal welfare. *Animals: An Open Access Journal from MDPI, 10*(3), 385. https://doi.org/10.3390/ani10030385

Bonneau, M., & Lebret, B. (2010). Production systems and influence on eating quality of pork. *Meat Science, 84*(2), 293–300. https://doi.org/10.1016/j.meatsci.2009.03.013

Čandek-Potokar, M., Lukač, N. B., Tomažin, U., Škrlep, M., Nieto, R., Candek-Potokar, M., & Linan, R. M. N. (2019). Analytical review of productive performance of local pig breeds. In *European local pig breeds—diversity and performance. A study of project TREASURE*. IntechOpen. https://doi.org/10.5772/intechopen.84214.

Dadousis, C., Muñoz, M., Óvilo, C., Fabbri, M. C., Araújo, J. P., Bovo, S., Potokar, M.Č., Charneca, R., Crovetti, A., Gallo, M., García-Casco, J. M., Karolyi, D., Kušec, G., Martins, J. M., Mercat, M.-J., Pugliese, C., Quintanilla, R., Radović, Č., Razmaite, V., & Bozzi, R. (2022). Admixture and breed traceability in European indigenous pig breeds and wild boar using genome-wide SNP data. *Scientific Reports, 12*(1), Article 1. https://doi.org/10.1038/s41598-022-10698-8

Danezis, G. P., Tsagkaris, A. S., Camin, F., Brusic, V., & Georgiou, C. A. (2016). Food authentication: Techniques, trends & emerging approaches. *TrAC, Trends in Analytical Chemistry, 85*, 123–132. https://doi.org/10.1016/j.trac.2016.02.026

Fanelli, V., Mascio, I., Miazzi, M. M., Savoia, M. A., De Giovanni, C., & Montemurro, C. (2021). Molecular approaches to agri-food traceability and authentication: An updated review. *Foods, 10*(7), Article 7. https://doi.org/10.3390/foods10071644

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics, 32*(18), 2847–2849. https://doi.org/10.1093/bioinformatics/btw313

Gvozdanović, K., Margeta, V., Margeta, P., Djurkin Kušec, I., Galović, D., Dovč, P., & Kušec, G. (2019). Genetic diversity of autochthonous pig breeds analyzed by microsatellite markers and mitochondrial DNA D-loop sequence polymorphism. *Animal Biotechnology, 30*(3), 242–251. https://doi.org/10.1080/10495398.2018.1478847

Horcada, A., Valera, M., Juárez, M., & Fernández-Cabanás, V. M. (2020). Authentication of Iberian pork official quality categories using a portable near infrared spectroscopy (NIRS) instrument. *Food Chemistry, 318*, Article 126471. https://doi.org/10.1016/j.foodchem.2020.126471

Jombart, T. (2008a). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics, 24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T. (2008b). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics, 24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics, 11*, 94. https://doi.org/10.1186/1471-2156-11-94

Mabood, F., Boqué, R., Alkindi, A. Y., Al-Harrasi, A., Al Amri, I. S., Boukra, S., Jabeen, F., Hussain, J., Abbas, G., Naureen, Z., Haq, Q. M. I., Shah, H. H., Khan, A., Khalaf, S. K., & Kadim, I. (2020). Fast detection and quantification of pork meat in other meats by reflectance FT-NIR spectroscopy and multivariate analysis. *Meat Science, 163*, Article 108084. https://doi.org/10.1016/j.meatsci.2020.108084

McDevitt, R. M., Gavin, A. J., Andrés, S., & Murray, I. (2005). The ability of visible and near infrared reflectance spectroscopy to predict the chemical composition of ground chicken carcasses and to discriminate between carcasses from different genotypes. *Journal of Near Infrared Spectroscopy, 13*(3), 109–117. https://doi.org/10.1255/jnirs.463

Meza-Márquez, O. G., Gallardo-Velázquez, T., & Osorio-Revilla, G. (2010). Application of mid-infrared spectroscopy with multivariate analysis and soft independent modeling of class analogies (SIMCA) for the detection of adulterants in minced beef. *Meat Science, 86*(2), 511–519. https://doi.org/10.1016/j.meatsci.2010.05.044

Muñoz, M., Bozzi, R., García, F., Núñez, Y., Geraci, C., Crovetti, A., García-Casco, J., Alves, E., Škrlep, M., Charneca, R., Martins, J. M., Quintanilla, R., Tibau, J., Kušec, G., Djurkin-Kušec, I., Mercat, M. J., Riquet, J., Estellé, J., Zimmer, C., … Óvilo, C. (2018). Diversity across major and candidate genes in European local pig breeds. *PLoS One, 13*(11), Article e0207475. https://doi.org/10.1371/journal.pone.0207475

Muñoz, M., Bozzi, R., García-Casco, J., Núñez, Y., Ribani, A., Franci, O., García, F., Škrlep, M., Schiavo, G., Bovo, S., Utzeri, V. J., Charneca, R., Martins, J. M., Quintanilla, R., Tibau, J., Margeta, V., Djurkin-Kušec, I., Mercat, M. J., Riquet, J., … Óvilo, C. (2019). Genomic diversity, linkage disequilibrium and selection signatures in European local pig breeds assessed with a high density SNP chip. *Scientific Reports, 9*(1), Article 13546. https://doi.org/10.1038/s41598-019-49830-6

Muñoz, M., García-Casco, J. M., Alves, E., Benítez, R., Barragán, C., Caraballo, C., Fernández, A. I., García, F., Núñez, Y., Óvilo, C., Fernández, A., Rodríguez, C., & Silió, L. (2020). Development of a 64 SNV panel for breed authentication in Iberian pigs and their derived meat products. *Meat Science, 167*, Article 108152. https://doi.org/10.1016/j.meatsci.2020.108152

Oliveri, P., Malegori, C., Mustorgi, E., & Casale, M. (2021). Qualitative pattern recognition in chemistry: Theoretical background and practical guidelines. *Microchemical Journal, 162*, Article 105725. https://doi.org/10.1016/j.microc.2020.105725

Ortiz, A., Parrini, S., Tejerina, D., Pinto de Araújo, J. P., Čandek-Potokar, M., Crovetti, A., Garcia-Casco, J. M., González, J., Hernández-García, F. I., Karolyi, D., Margeta, V., Martins, J. M., Nieto, R., Petig, M., Razmaite, V., Sirtori, F., Lebret, B., & Bozzi, R. (2020). Potential use of near-infrared spectroscopy to predict fatty acid profile of meat from different European autochthonous pig breeds. *Applied Sciences, 10*(17), Article 17. https://doi.org/10.3390/app10175801

Parrini, S., Sirtori, F., Čandek-Potokar, M., Charneca, R., Crovetti, A., Kušec, I. D., Sanchez, E. G., Cebrian, M. M. I., Garcia, A. H., Karolyi, D., Lebret, B., Ortiz, A., Panella-Riera, N., Petig, M., Jesus da Costa Pires, P., Tejerina, D., Razmaite, V., Aquilani, C., & Bozzi, R. (2023). Prediction of fatty acid composition in intact and minced fat of European autochthonous pigs breeds by near infrared spectroscopy. *Scientific Reports, 13*(1), Article 1. https://doi.org/10.1038/s41598-023-34996-x

Pasquini, C. (2018). Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. https://www.sciencedirect.com/science/article/pii/S0003267018304793.

Prieto, N., Juárez, M., Larsen, I. L., López-Campos, Ó., Zijlstra, R. T., & Aalhus, J. L. (2015). Rapid discrimination of enhanced quality pork by visible and near infrared spectroscopy. *Meat Science, 110*, 76–84. https://doi.org/10.1016/j.meatsci.2015.07.006

Prieto, N., Pawluczyk, O., Dugan, M. E. R., & Aalhus, J. L. (2017). A review of the principles and applications of near-infrared spectroscopy to characterize meat, fat, and meat products. *Applied Spectroscopy, 71*(7), 1403–1426. https://doi.org/10.1177/0003702817709299

Prieto, N., Roehe, R., Lavín, P., Batten, G., & Andrés, S. (2009). Application of near infrared reflectance spectroscopy to predict meat and meat products quality: A review. *Meat Science, 83*(2), 175–186. https://doi.org/10.1016/j.meatsci.2009.04.016

R Foundation for Statistical Computing. (2021). R core team. R: A language and environment for statistical computing. *Software*. https://www.R-project.org/.

Sun, S., Guo, B., Wei, Y., & Fan, M. (2012). Classification of geographical origins and prediction of δ13C and δ15N values of lamb meat by near infrared reflectance spectroscopy. *Food Chemistry, 135*(2), 508–514. https://doi.org/10.1016/j.foodchem.2012.05.004

Thia, J. A. (2023). Guidelines for standardizing the application of discriminant analysis of principal components to genotype data. *Molecular Ecology Resources, 23*(3), 523–538. https://doi.org/10.1111/1755-0998.13706

Xu, L., Deng, D.-H., Cai, C.-B., & Yang, H.-W. (2011). Automatic discrimination of the geographical origins of milks by excitation-emission fluorescence spectrometry and chemometrics. *Journal of Automated Methods & Management in Chemistry*, 1–6. https://doi.org/10.1155/2011/323196, 2011.

Zamora-Rojas, E., Pérez-Marín, D., De Pedro-Sanz, E., Guerrero-Ginel, J. E., & Garrido-Varo, A. (2012). In-situ Iberian pig carcass classification using a micro-electro-mechanical system (MEMS)-based near infrared (NIR) spectrometer. *Meat Science, 90*(3), 636–642. https://doi.org/10.1016/j.meatsci.2011.10.006