PORK FAT FATTY ACID PROFILING WITH RAMAN SPECTROSCOPY AND PRINCIPAL COMPONENT ANALYSIS

Marina A. Nikitina, Irina M. Chernukha*, Viktoriya A. Pchelkina, Nikolai A. Ilin, Andrey B. Lisitsyn

V. M. Gorbatov Federal Research Center for Food Systems, Moscow, Russia

ARTICLE INFO

ABSTRACT

Pork quality is largely influenced by the fat composition. Getting reliable data is critical for any successful analysis. Using multidimensional statistical analysis methods opens up the possibility to visualize Raman spectroscopy results of fatty acid (FA) profiles and identify animals by breed. Back fat of the Altai breed (sample 1), Duroc (sample 2) and Livenskaya (sample 3) was analyzed. At least 36 spectra were taken from each sample by the Renishaw inVia Reflex confocal Raman spectrometer and analyzed using the principal component method. The Cattell's scree test was applied to determine the number of components "significant main components" to retain, namely – PC1, PC2 and PC3 (85%). It is shown samples 1 and 3 FA form clusters in all graphs of the "significant" main components. Sample 2 forms an area on the PC1 / PC2 graph and locates in the I, III and IV quadrants. Sample 1 - in the II quadrant, and Sample 3 - in the IV quadrant. For PC1 the most prominent variable is 1650 cm^{-1}, responsible for the C = C molecular bond, for the saturated FA and conjugated linoleic acid. Spectrum 1650 cm^{-1} is important in the intraspecific classification of pork. For PC2 – main contribution of 868 cm^{-1} was marked, and 1368 cm^{-1} for PC3. Each spectrum characterized group of the pork backfat FA. PCA makes it possible to: (1) to evaluate pork fat lipid profile by groups - saturated, mono-, polyunsaturated, with a long carbon chain, etc.; (2) obtain reliable differences between breeds; (3) identify individual FA, via Raman spectra patterns.

Keywords: Raman spectra, hidden patterns, "significant" main component, chemometrics

INTRODUCTION

Principal component analysis (PCA) is used to reduce (lower) the dimensionality of initial data and to study/identify similarities and hidden conformities between samples where the relationship between data and grouping are not yet clear or have not been identified.

Karl Pearson is considered the founder of the PCA method. In 1901, he published his research paper "On lines and planes of closest fit to systems of points in space" (Pearson, 1901). However, back in 1889, an English mathematician James Joseph Sylvester published a work on the same topic (Sylvester, 1889).

Principal component analysis refers to multivariate methods of statistical analysis (see figure 1). Sometimes this method is also called the Karhunen-Loève transform or the Hotelling transform.

Currently, the PCA method is very popular and is used in many areas, including the food industry: in studying the taste of wines (Vilanova et al. 2010) and marmalades based on agar-agar, gelatin, and pectin (Zhilinskaya et al. 2018); the texture of extruded snacks (Paula and Conti-Silva, 2014); cheeses (Deegan et al. 2014); and the influence of

the plant component on the overall acceptability of plant-based burgers (Dhanapal and Erkinbaev, 2024). A number of authors have used PCA to assess the impact of individual animal and poultry characteristics on raw material quality. Thus, the work (Li et al. 2024) studied the impact of the genetic characteristic of growth rate on the duck meat quality.

Chinese researchers Xiao (2024) successfully used PCA to identify correlations between the meat characteristics of 434 broilers of 12 different breeds from Southern China. The method was successfully used to identify yak meat depending on fattening and growing conditions – pasture- or grain-fed ones (Liu et al. 2024).
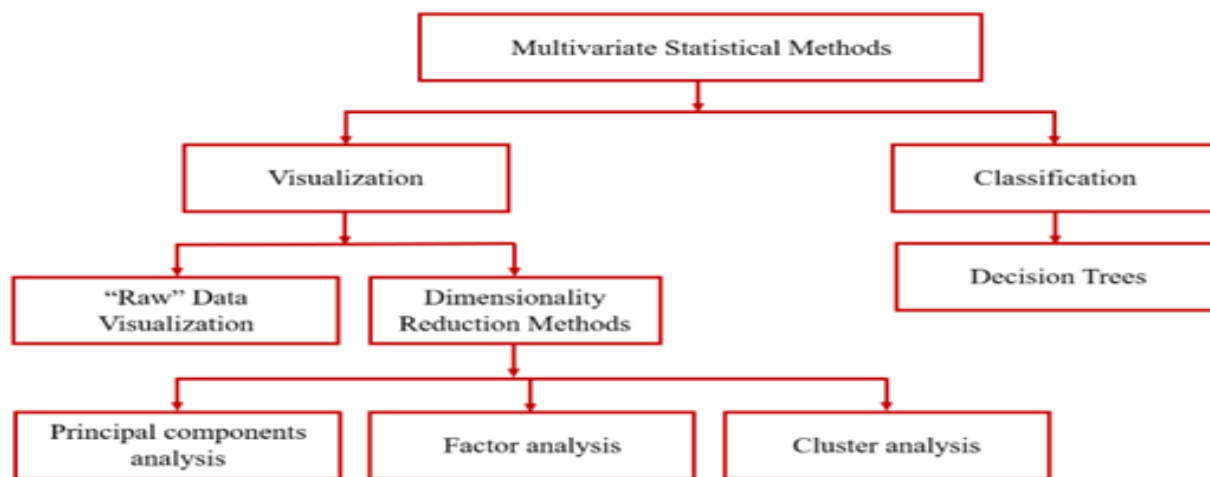

Fig. 1: Classification of multivariate statistical methods

Many works are devoted to the use of chemometrics - a set of mathematical tools for studying (bio)chemical processes in living tissues. In particular, when analyzing the characteristics of agricultural raw materials and products.

In the study by Szykuła (2023), Raman spectroscopy in combination with chemometric methods (PCA and PLS-DA) was used to study the differences in fat tissue profiles for inter- and intra-species classification of meat. A clear separation was shown between pork, lamb, and chicken samples. The best separation was observed for the first principal component (explaining 78.4% of the total variance) and the second principal component (explaining 12.1% of the total variance).

In the work (Saleem et al. 2021), the capabilities of Raman spectroscopy and PCA were used to study the characteristics of goat, cow, and buffalo fat. The scientists studied the relative concentration of beta-carotene, fatty acids, lipids, conjugated linoleic acid (CLA), and vitamin D in them. The score plot shows that 51% of the variance is contributed by PC1 and 10% - by PC2.

Berhe (2016), in their study, investigated models of partial least squares (PLS) for predicting individual fatty acids based on Raman spectra related to iodine value and other common characteristics of fatty acids in pork fat. Principal component analysis was performed to extract information about variations in the dataset. PCA proves that there are differences between the outer and inner layers of animal fat. As a result, PLS models were obtained with good correlation (from 0.78 to 0.90) between Raman spectra and the following parameters: iodine value (IV), saturated fatty acids (SFA), monounsaturated fatty acids (MUFA), and polyunsaturated fatty acids (PUFA).

Logan (2020) conducted a study on using Raman spectroscopy to differentiate between grass-fed and grain-fed cattle carcasses in order to develop a method for identifying feeding systems. The experiment involved 300 beef carcasses (150 grass-fed and 150 grain-fed animals). The clustering in the scatterplot obtained by PCA shows that the variation in the spectra is related to the grain-fed and grass-fed systems. As a result, the scientists stated that it is possible to differentiate between grass-fed and grain-fed cattle carcasses using Raman spectra.

Robert (2020) examined 90 red meat samples: 1) beef (Bos Taurus); 2) lamb (Ovis aries); 3) venison (Cervus elaphus scoticus, hippelaphus и pannonensis). For each type of meat, 30 samples were selected. Red meat samples were measured using Raman spectroscopy and analyzed by PCA as an unsupervised multidimensional analysis tool, while support vector machine (SVM) classification and partial least squares discriminant analysis (PLSDA) were used as a supervised multidimensional analysis tool. The first two principal components explained 59% of the variance in meat samples. The PCA score plot shows that beef, venison, and lamb are clearly separated from each other. Moreover, marker signals are identified: 911 cm^{-1} for lamb and 1267-1316 cm^{-1} for beef. Venison samples are grouped in the positive space of PC2, while beef and lamb samples are grouped in the negative space of PC2.

The work (Ostovar pour et al. 2019) presents a study using chemometric analysis to determine chemically specific spectral characteristics of meat suitable for chemical identification. The chemicals studied were glycogen, glucose,

lactate, and cortisol, which are predictors of meat quality. The sample was beef rump, consisting of the muscles *M. biceps femoris* and *M. gluteus medius*. The authors conclude that the chemometric analysis clearly shows the separation of metabolites into four distinct groups, even for such chemically similar compounds as glucose, glycogen, and lactate.

Abbas (2009) compared the Raman spectra of poultry, pork, beef, lamb, and fish fats using PCA. The method was applied to the collected spectra in the range between 3100 and 2650 cm⁻¹, and 1800-1200 cm⁻¹. The overall variation of the dispersion was 91% (PC1=67%, PC2=24%). The visual representation of the method showed the differences between the fats. Fish, poultry, pork, and beef fats were clearly differentiated. Thus, the Raman spectra reflect the composition of these samples well.

In the work (Boyaci et al. 2014), seven types of meat and their salami were successfully differentiated from each other according to their origin using Raman spectroscopy and PCA. The results of this study showed that Raman spectroscopy with a chemometric method can be used to determine the origin of meat types.

Lyndgaard (2011), in their study, applied PCA to Raman spectra of porcine adipose tissue to differentiate fat layers (external and internal) and to evaluate the variation in fatty acid composition with fat depth and fat layer. The samples for this study were 16 pig carcasses. Variability in fatty acid composition was ensured by selecting carcasses from the two extreme fattening groups and the normal central group. To relieve proper depth characterization, the minimum thickness of the loin fat should be 16 mm. The PCA plot shows almost complete separation of the fat layers along the first principal component, which describes 79% of the total spectral variation. As a result, the authors conclude that Raman spectroscopy is a potential measurement method for grading pork carcasses on the conveyor belt.

Gao (2020) compared the capabilities of near-infrared Raman (FT-Raman, λex~1064 nm) and visible Raman (vis-Raman, λex~532 nm) for animal fats. PCA plots showed that lard and chicken samples were on the negative semi-axis of PC1, while beef and lamb samples fell in the positive direction of PC1. Besides, PC2 separated lard and chicken, as well as beef and lamb. The complete separation of animal fats from different species indicates that fat characteristics of different animal fats are significantly different and have significant potential for species identification analysis. Chemometric analysis showed that vis-Raman spectroscopy has better discrimination ability for animal fats compared with FT-Raman spectroscopy.

Our study is related to the possibility of using PCA in the analysis of Raman spectra of meat raw material (fat) samples in order to visualize the results and obtain data for assessing and comparing fatty acid (FA) profiles of pigs of different breeds.

OBJECTS AND METHODS

Chilled (4±2 °C) samples of adipose tissue (back fat) of pigs of three different breeds Altai (Sample 1, n=3), Duroc (Sample 2, n=3) and Livni (Sample 3, n=3) obtained on pig slaughter lines at enterprises in Moscow, Barnaul, and Livny were the objects.

All pigs were raised using standardized Russian technologies. Samples were obtained from animals weighing 110±10 kg 24 hours after slaughter. Sampling of 5x5 cm in size and depth from the surface of subcutaneous fat to the muscle layer was carried out between the 10-th and 11-th ribs. At least three copies of each sample were taken from one animal, and the average value was used for further data processing.

The samples were packed in plastic containers and transported at 6±2ºC to the laboratory. The samples were analyzed within 24 hours from the moment of sampling. Six pieces of no more than 10×10×5 mm in size were selected from each sample for further analysis. In accordance with Olsen's data (2020), we did not perform preliminary sample preparation, since they showed the possibility of predicting fatty acid regions based on spectra measured directly in adipose tissue.

Spectra were made on the Renishaw inVia Reflex confocal Raman dispersion spectrometer (Renishaw plc, Wotton-under-Edge, UK) using a 785 nm laser.

The spectrometer was calibrated before each study by recording the Raman spectrum of a silicon crystal wafer at 520 cm⁻¹ (exposure time 1 s, laser power 10 mW, 1 scan). A lens with a magnification of L50× power was used to focus the laser on the surface of the pieces. Spectra were collected at 100 mW laser power, exposure time of 10 s, 3 accumulations. At least 6 spectra were measured from different points on each piece of the sample. Measurements were recorded in the range of 800-1800 cm⁻¹. Laser power and integration time were optimized to avoid photodegradation. All spectral analysis and pre-processing (cosmic ray removal, baseline correction using intelligent polynomial algorithms, smoothing using the Savitzky-Golay algorithm, normalization (up to 1000)) were performed using Renishaw WiRE 5.2 (Renishaw plc, Wotton-under-Edge, UK) software. The data are a m×n matrix, where m is the number of samples, n is the number of points in the spectrum.

The principal component method represents the data (X matrix) as a set of linearly independent vectors in a new space. These vectors are called principal components. According to (Pomeranzev, 2008), X matrix can be decomposed into two matrices of significantly smaller dimensions.

$$X = TP^t \qquad (1)$$

where   is projection matrix of points onto principal components (score matrix);

 is transpose of a matrix consisting of principal components in old coordinates (loading matrix).

In this case, the score matrix () has almost all the information about the spectral data despite the fact that it has a smaller dimension relative to X matrix.

Each of the principal components is responsible for some hidden parameter characterizing the data, and the older the principal component, the greater the percentage of information about the original system it is responsible for.

To determine the practicability of reducing the dimensionality of the data (i.e., compressing them) with significant results without losing the original information and assessing the equality of the variances of several samples, the Bartlett's test of sphericity was calculated. If the criterion value is less than 0.05, factor analysis is acceptable. The criterion was calculated using the formula (Bartlett, 1937):

$$\chi^2 = N' \ln \frac{|AA^T|}{|R|},$$ (2)

where $N' = N - \frac{1}{6}(2n + 5) - \frac{2}{3}m;$

|AA| is determinant of the reproduced correlation matrix;
|R| is determinant of the original correlation matrix;
n is number of variables;
m is number of identified common factors;
N is number of objects of study.

The number of significant principal components was determined using the Kaiser (Kaiser, 1960) and Cattell's scree test (Cattell, 1966) (see figure 2).

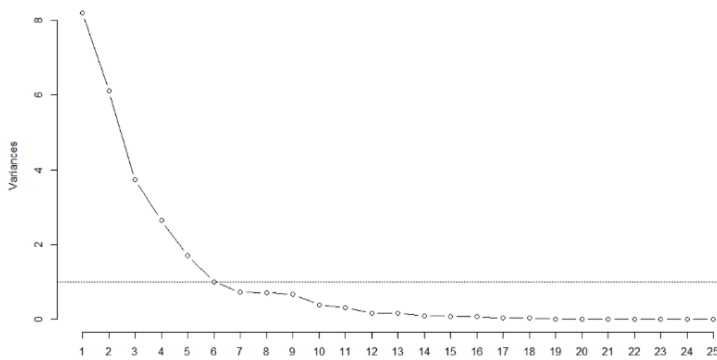Values greater than one (the dotted line on the graph) are considered "significant" in PCA.

Fig. 2: Cattell's scree test diagram

The Kaiser–Meyer–Olkin criterion shows how suitable the data is for factor analysis. If the criterion values are in the range from 0.8 to 1, then the data fits best.

$$KMO = \frac{\sum_{i=1}^{n}\sum_{\substack{j=1 \\ i \neq j}}^{n} r_{ij}^2}{\sum_{i=1}^{n}\sum_{\substack{j=1 \\ i \neq j}}^{n} p_{ij}^2 + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i \neq j}}^{n} r_{ij}^2}$$ (3)

where $p_{ij} = \frac{R_{ij}}{\sqrt{R_{ii} \cdot R_{jj}}}$ is partial correlation coefficient;
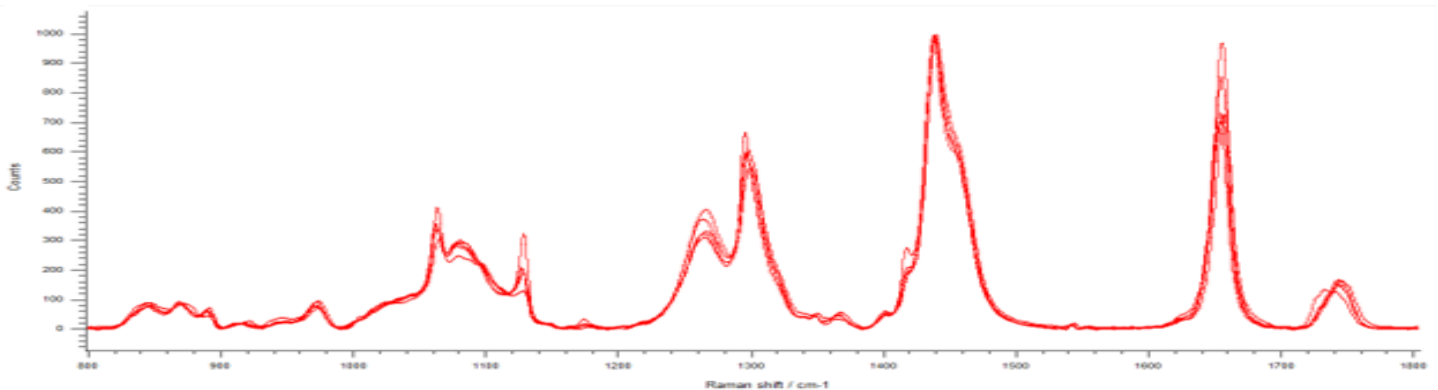
ij=R(X_i,X_j) is Pearson correlation.

Data processing by the PCA method (Pomeranzev, 2008; Rodionova et al. 2021) was carried out in the R Studio (developer R-Tools Technology) environment in the R programming language with the inclusion of libraries *tidyverse, ggrepel, ggplot2, factoextra, pls, nipals* (Mastickij and Shitikov, 2015; Kabacoff, 2022). R software is a freely distributed cross-platform software tool used for statistical calculations and data visualization. R distributions are available on websites The Comprehensive R Archive Network. Statistical processing of the obtained results of Raman spectra intensity was carried out in the MS Excel spreadsheet processor at a significance level of 0.05. The data are presented as the mean value and standard deviation.

To determine the reliability of differences in mean values, non-parametric test Kruskal-Wallis was used using the Dunn's test. The probability of 0.05 was chosen as a significant level.

RESULTS AND DISCUSSION

Fatty acid proportions of 1) Altai backfat: ΣUFA/ΣMUFA = 1.83±0.23; ΣUFA/ΣPUFA = 2.21±0.63; ΣMUFA/ΣPUFA = 1.21±0.21; ΣSFA/ΣUFA = 0.59±0.04; 2) Duroc backfat ΣUFA/ΣMUFA = 1.31±0.19; ΣUFA/ΣPUFA = 4.26±0.53; ΣMUFA/ΣPUFA = 3.26±0.42; ΣSFA/ΣUFA = 0.78±0.08; 3) Livni backfat ΣUFA/ΣMUFA = 1.22±0.18; ΣUFA/ΣPUFA = 5.49±0.67; ΣMUFA/ΣPUFA = 4.48±0.59; ΣSFA/ΣUFA = 0.66±0.09.

Fatty acid profiles were obtained in the range of 800-1800 cm^{-1} and contained distinct spectral signatures. The obtained Raman spectra of the samples are shown in figure 3
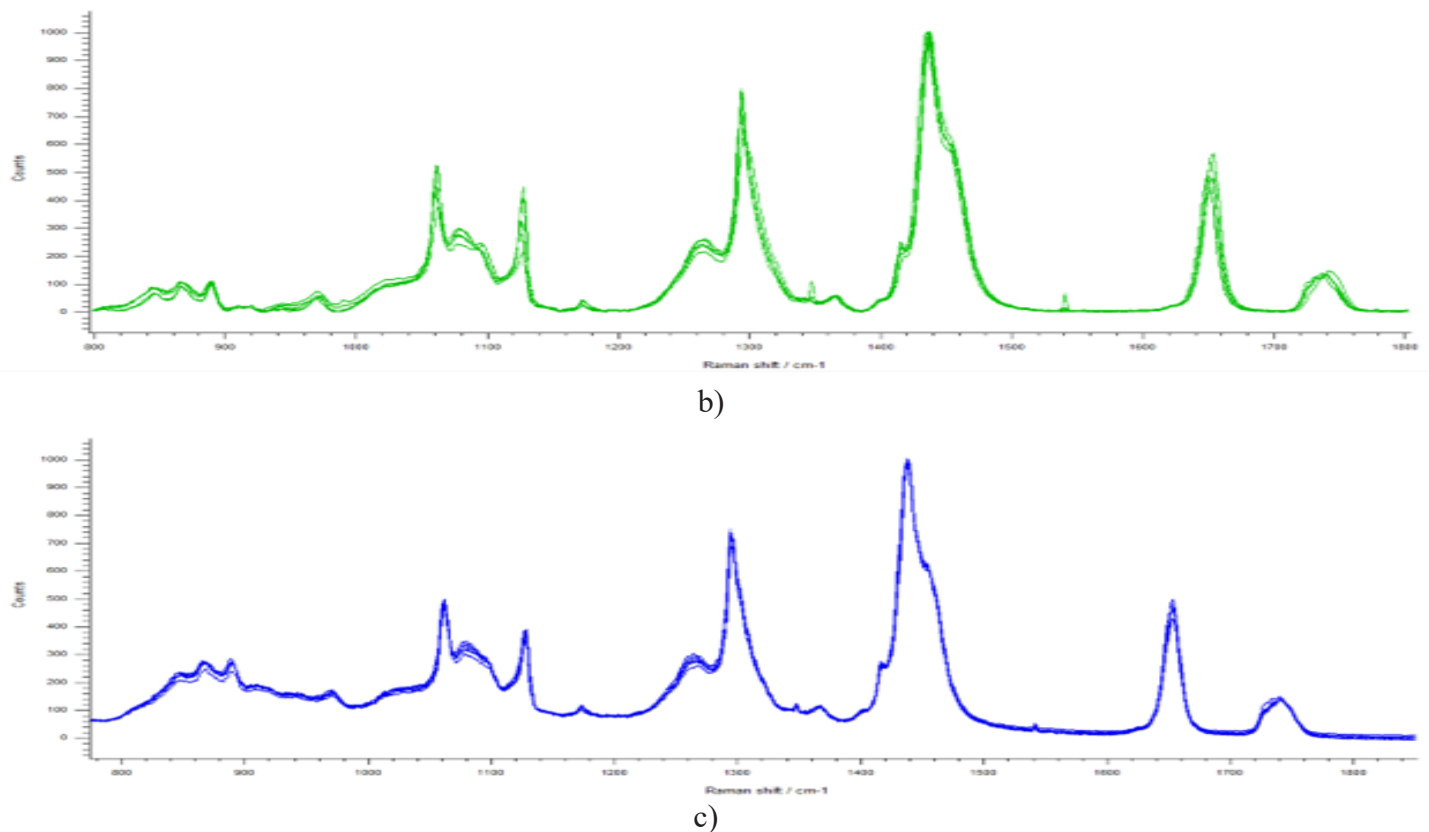

a)

b)

c)

Fig. 3: Raman spectra of the samples in the range 800-1800 cm^{-1}: a) Sample 1; b) Sample 2; c) Sample 3
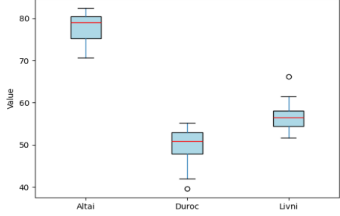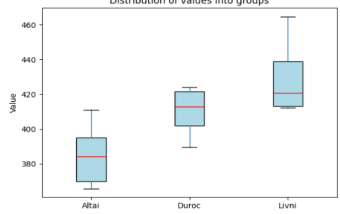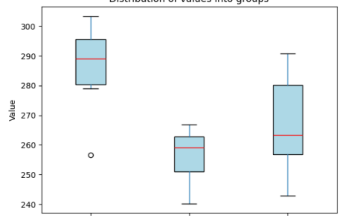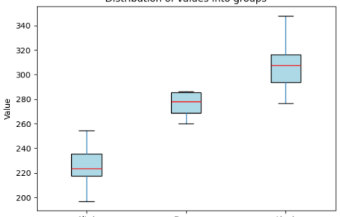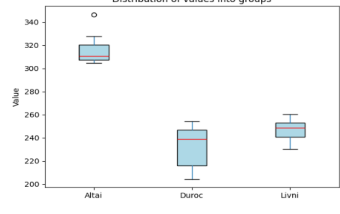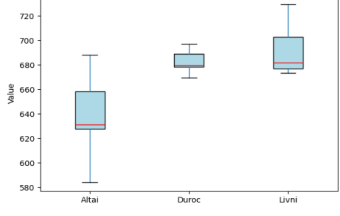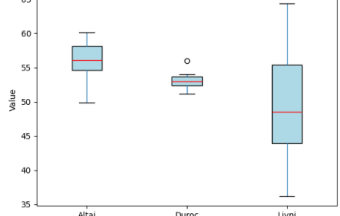
Some blur of the spectra is explained by the fact that each of figures 3 shows 5 spectra from one sample.
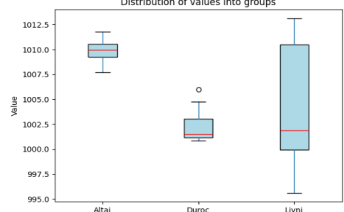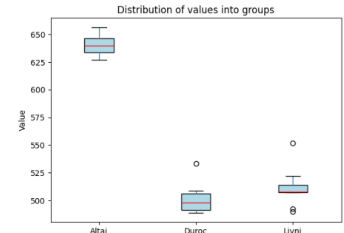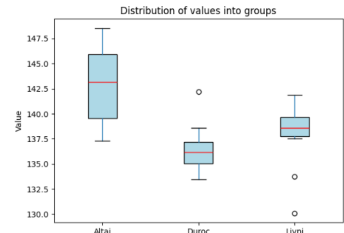
Based on the analyzed variables, the data in the range between 700 cm^{-1} and 1800 cm^{-1} were selected, where peaks associated with lipids are contained. All the obtained spectra showed identical signals, the characteristics of which are given in the works (Saleem et al. 2021; Berhe et al. 2016; Robert et al. 2020; Pchelkina et al. 2022; Czamara et al. 2014), differing only in their intensity (Table 1). Samples 1 and 3 were characterized by a more intense signal 868 cm^{-1} (**C-C** stretching) compared to Sample 3, signal 970 cm^{-1} (=**C-H** out-of-plane b**a**nd *cis* isomer) in Samples 2 and 3 did not statistically differ in its intensity (P<0.05), and in Sample 1 it was characterized by a value 1.5 times greater. The peak at 1061 cm^{-1} (**C-C** aliphatic out-of-phase stretching) was significantly different between Samples 1 and 3, and 1080

cm^{-1} (**C-C** aliphatic stretching) showed no statistically significant differences between all three samples (P<0.05). The signal at 1127 cm^{-1} (**C-C** aliphatic in-phase stretching) was the least intense in Sample 1, and was the most intense in Sample 3. The signal at 1266 cm^{-1} (=**C-H** symmetric rock *cis* isomer) was 1.4 times more intense in Sample 1 than in Samples 2 and 3 (P<0.05), and 1300 cm^{-1} (CH2 twisting) was statistically significantly different between Samples 1 and 3 (P<0.05). The signals at 1368 cm^{-1} (CH3 symmetric deformation (umbrella)) and 1438 cm^{-1} (CH2 symmetric deformation (scissoring)) and 1740 cm^{-1} (**C=O** stretching) did not show statistically significant differences in intensity between all three samples (P<0.05). The signal at 1650 cm^{-1} (**C=C** stretching) was 1.2 times more intense in Sample 1 than in Samples 2 and 3.

Table 1. Raman spectra average intensity of samples tested

Band number	Band position (cm^{-1})	Sample 1 Altai	Sample 2 Duroc	Sample 3 Livni	Box plots
1	868	100.25±2.78[a-b] p_{a-b}=0.001 (yes) p_{a-c}=1 (no) p_{b-c}=0.001 (yes)	91.36±3.56[a-b,b-c]	99.23±5.77[b-c]	

Band number	Band position (cm^{-1})	Sample 1 Altai	Sample 2 Duroc	Sample 3 Livni	Box plots
2	970	$78.00\pm6.52^{a-b,a-c}$ $p_{a-b}=0$ (yes) $p_{a-c}=0.013$ (yes) $p_{b-c}=0.16$ (no)	$46.54\pm7.29^{a-b}$	$51.67\pm6.12^{a-c}$	
3	1061	$389.43\pm14.10^{a-b,a-c}$ $p_{a-b}=0.014$ (yes) $p_{a-c}=0$ (yes) $p_{b-c}=0.44$ (no)	$409.79\pm8.64^{a-b}$	$434.84\pm19.44^{a-c}$	
4	1080	$288.77\pm13.33^{a-b}$ $p_{a-b}=0.001$ (yes) $p_{a-c}=0.081$ (no) $p_{b-c}=0.59$ (no)	$254.53\pm12.74^{a-b}$	265.04 ± 15.74	
5	1127	$230.46\pm18.96^{a-b,a-c}$ $p_{a-b}=0.022$ (yes) $p_{a-c}=0$ (yes) $p_{b-c}=0.06$ (no)	$274.46\pm11.86^{a-b}$	$310.54\pm17.74^{a-b,a-c}$	
6	1263-1266	$322.23\pm23.02^{a-b,a-c}$ $p_{a-b}=0$ (yes) $p_{a-c}=0.003$ (yes) $p_{b-c}=0.86$ (no)	$232.71\pm19.40^{a-b}$	$241.12\pm21.02^{a-c}$	
7	1300-1306	$652.74\pm20.70^{a-b,a-c}$ $p_{a-b}=0.008$ (yes) $p_{a-c}=0.001$ (yes) $p_{b-c}=1$ (no)	$681.19\pm9.12^{a-b}$	$699.53\pm19.04^{a-c}$	
8	1368	$54.58\pm3.52^{a-c}$ $p_{a-b}=0.28$ (no) $p_{a-c}=0.044$ (yes) $p_{b-c}=1$ (no)	53.79 ± 1.77	$51.49\pm9.28^{a-c}$	

Band number	Band position (cm⁻¹)	Sample 1 Altai	Sample 2 Duroc	Sample 3 Livni	Box plots
9	1438	$1009.48\pm1.78^{a\text{-}b}$ $p_{a\text{-}b}=0$ (yes) $p_{a\text{-}c}=0.5$ (no) $p_{b\text{-}c}=1$ (no)	$1001.81\pm2.79^{a\text{-}b}$	1007.73 ± 6.95	
10	1650-1655-1657	$645.71\pm11.11^{a\text{-}b,a\text{-}c}$ $p_{a\text{-}b}=0$(yes) $p_{a\text{-}c}=0.003$ (yes) $p_{b\text{-}c}=0.79$ (no)	$503.61\pm15,03^{a\text{-}b}$	$511.78\pm11.72^{a\text{-}c}$	
11	1740	$141.11\pm4.96^{a\text{-}b}$ $p_{a\text{-}b}=0.002$ (yes) $p_{a\text{-}c}=0.17$ (no) $p_{b\text{-}c}=0.44$ (no)	$136.03\pm2,64^{a\text{-}b}$	137.70 ± 4.24	

a-b, b-c, a-c are significant differences between the samples of Altai, Duroc, Livni (within one line), non-parametric test Kruskal-Wallis, Dunn's test -(P<0.05)

Analysis of data in Table 1 and figure 1 showed the following. The greatest interbreed differences are observed in the intensity of signals 970, 1061 and 1127 cm⁻¹. In all spectra, the fatty acid profile of the back fat of Altai pigs significantly differs from the back fat of Livni and Duroc pigs. The exception is the intensity peak 868 cm⁻¹, which characterizes the structure of the carbon chain of fatty acids with a carboxyl group. The significance of these parameters is confirmed by the data in Table 3. Such differences in the FA spectra are expected due to the differences between these breeds: Altai is a meat breed, and Livni is a meat-lard breed. At the same time, Duroc is also positioned as a meat breed, but reliable differences in the FA profiles obtained by Raman spectroscopy were found between Altai and Duroc. It is worth noting that there are no significant differences in the peak intensities for the FA profiles of Duroc and Livni pig lard, characteristic of the sums of saturated and unsaturated FA. The data presented are consistent with the results we obtained earlier (Chernukha et al. 2023).

At that, the greatest differences were observed in the intensity of peaks 1263-1266 and 1650-1655-1657 cm⁻¹.

Animal fats from Altai pigs gave obvious differences in the ratios of peak intensities at 1653/1745 cm⁻¹, while for Duroc and Livni the differences were unreliable.

The obtained values of the Bartlett's test of sphericity (6.12e-6 or 0.00000612, see formula (2) allow performing principal component analysis (p-value<0.05), as well as the Kaiser–Meyer–Olkin criterion (0.809, see formula (3). As a result of original matrix processing (17×12), 12 principal components were formed (Table 2).

Table 2. Characteristics of importance of the principal components

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC1**2**
Standard deviation*	2.59	1.35	1.31	0.79	0.72	0.61	0.40	0.23	0.09	0.08	0.47	0.02
Proportion of variance**	0.56	0.15	0.14	0.05	0.04	0.03	0.01	0.0043	0.0007	0.0006	0.0002	0.0000
Cumulative Proportion***	0.56	0.71	0.85	0.91	0.95	0.98	0.9942	0.9985	0.9992	0.9998	0.9999	1.0000

* mean-square deviation of the principal components;
** proportion of deviations of the principal components;
*** proportion of the total variance in the data that explains a certain number of principal components

With a larger number of components, the variance decreases, which is natural. PC1 describes 56% of the variance of the original data set, and PC1 and PC2 explain 71% of the variance of the original data. For example, seven principal components PC1-PC7 describe about 99% of the variance from the original data set, etc.

The number of "significant" principal components was determined using the Cattell's scree test - these are PC1, PC2 and PC3.

Table 3 presents the weights of the "significant" principal components with which the "old" variables are included in the "new" ones. New variables are calculated as the sum of the products of the weight values by the corresponding criterion (see formula (1).

Table 3. Weight coefficients of new variables

Spectrum signal (cm⁻¹)	PC1	PC2	PC3
868	-0.1066	**0.5895**	-0.1787
970	-0.3258	0.2419	-0.2049
1061	0.3191	0.3405	-0.2004
1080	-0.3237	0.0546	-0.2236
1127	0.3396	0.2899	-0.1664
1266	-0.3580	0.1351	-0.1856
1300	0.3228	0.2258	-0.3239
1418	0.1845	0.3610	0.3393
1368	0.0514	-0.2345	**-0.6631**
1438	-0.2492	0.3604	0.3152
1650	**-0.3721**	0.0764	-0.1199
1740	-0.3017	0.0284	-0.0193

Component stability testing was done using Cosine Similarity. As a result, average cosine similarity value/stability coefficient is 0.92, closer to 1 (one), which means the model is stable.

In PC2, signal 868 cm⁻¹ (weight coefficient 0.5895) significantly predominates, which characterizes the *C-C* interaction, as in PC3, signal 1368 cm⁻¹ (weight coefficient 0.6631), which shows the number of functional groups CH_3. Considering that the structural formula of all fatty acids includes carbon-carbon bonds and methyl groups (CH_3), it can be argued that the main functional difference is carried by the first principal component.

The values obtained from the spectrum signals 1650 cm⁻¹ (-0.3721), 1266 cm⁻¹ (-0.3580), showing unsaturated bonds $C=C$ and $=C-H$, are included in PC1 with the highest weighting coefficients. Olsen (2007) attributes the peaks in 1263/1266, 1650/1655 cm⁻¹ regions to double bonds. These types of bonds are present in off-peak fatty acids. In combination with the second or third principal components, it determines the specificity of the FA profile. In our example, the proportion of unsaturated FAs predominates in Sample 1 – Altai (1298.45), followed by Sample 3 – Livni (1211.31) and Sample 2 – Duroc (1184.8). The intensity of peaks at 970 cm⁻¹ was recorded to be 1.7 and 1.5 times higher in Altai lard, compared to the lard of Duroc and Livni pigs, respectively. This trend clearly correlated with the previously obtained UFA distribution in the lard of the pigs of the studied breeds (Chernukha et al. 2023), calculated from the FA profile obtained by the GLC method 56.07 ± 2.93%; 60.18 ± 3.91% and 62.82 ± 1.68% for the Duroc, Livni, and Altai breeds, respectively.

The value of 1650/1655 cm⁻¹ signal, according to Gómez-Mascaraque (2020), is related to the concentration of alpha-linolenic acid. This value, according to the Table 1 data, is maximum in the samples of lard of the Altai breed of pigs and is approximately equal in the lard of the Duroc and Livni breeds of pigs. This is consistent with the GC data showing the content of alpha-linolenic FA 1.01 ± 0.07, 0.60 ± 0.32 and 0.60 ± 0.15%, which we obtained earlier (Chernukha et al. 2023).

These differences can also be explained not only by differences in the ratio of saturated/unsaturated FAs, but also by the prevalence of individual ones, differing in spatial structure. This assumption is consistent with the data (Li-Chan et al. 1994). These researchers, using continuous gradient temperature Raman spectroscopy (GTRS), showed the possibility of identifying and differentiating certain regions of the carbon chain. Previously, it was shown that Raman spectroscopy can be used to study the structure of proteins and peptides, including primary and secondary structures, extensions, and side bonds. In particular, of aliphatic *CH* groups (Chernukha et al. 2023).

Signal 970 cm⁻¹, responsible for unsaturated bonds of fatty acids, enters PC1 and PC2 with weight coefficients of different (opposite) signs – -0.3258 and 0.2419, respectively. *C-C* interaction is also described by signal 1127 cm⁻¹. In our case, this signal enters PC1 and PC2 with the same signs of weight coefficients (0,3396 и 0,2899) (Saleem et al. 2021; Broadhurst et al. 2018).

The smallest values of weight coefficients are highlighted in red in the Table 3.

Figure 4 shows the zones of the FA profiles of samples formed by the PCA method.

In figure 4, we observe that points of Samples 1 and Samples 3 are grouped together in all three graphs (a,b,c). All graphs show a tightly grouped set of points of Sample 3. The points of Sample 2 are distributed chaotically in all graphs, and only in figure 4a they form a group, which is highlighted by an ellipse of the corresponding color. The grouping of points of Samples 2 is noted only in figure 4a. At that, points corresponding to Sample 1 are in II and III quadrants (positive and negative according to PC1 and negative

according to PC2), those corresponding to Sample 3 are in III and IV quadrants (negative according to PC1 and positive and negative according to PC2), and those corresponding to Sample 2 are in all four quadrants. There is an overlap of the area of points of Sample 1 with the area of Samples 2. Points of Sample 1 form a group on each graph, but the points are not located close to each other. This grouping is probably due to the different breeds of pigs from which samples were taken for the study and, consequently, to different quality characteristics of lard, which was previously confirmed in the work(Motoyama et al. 2010).
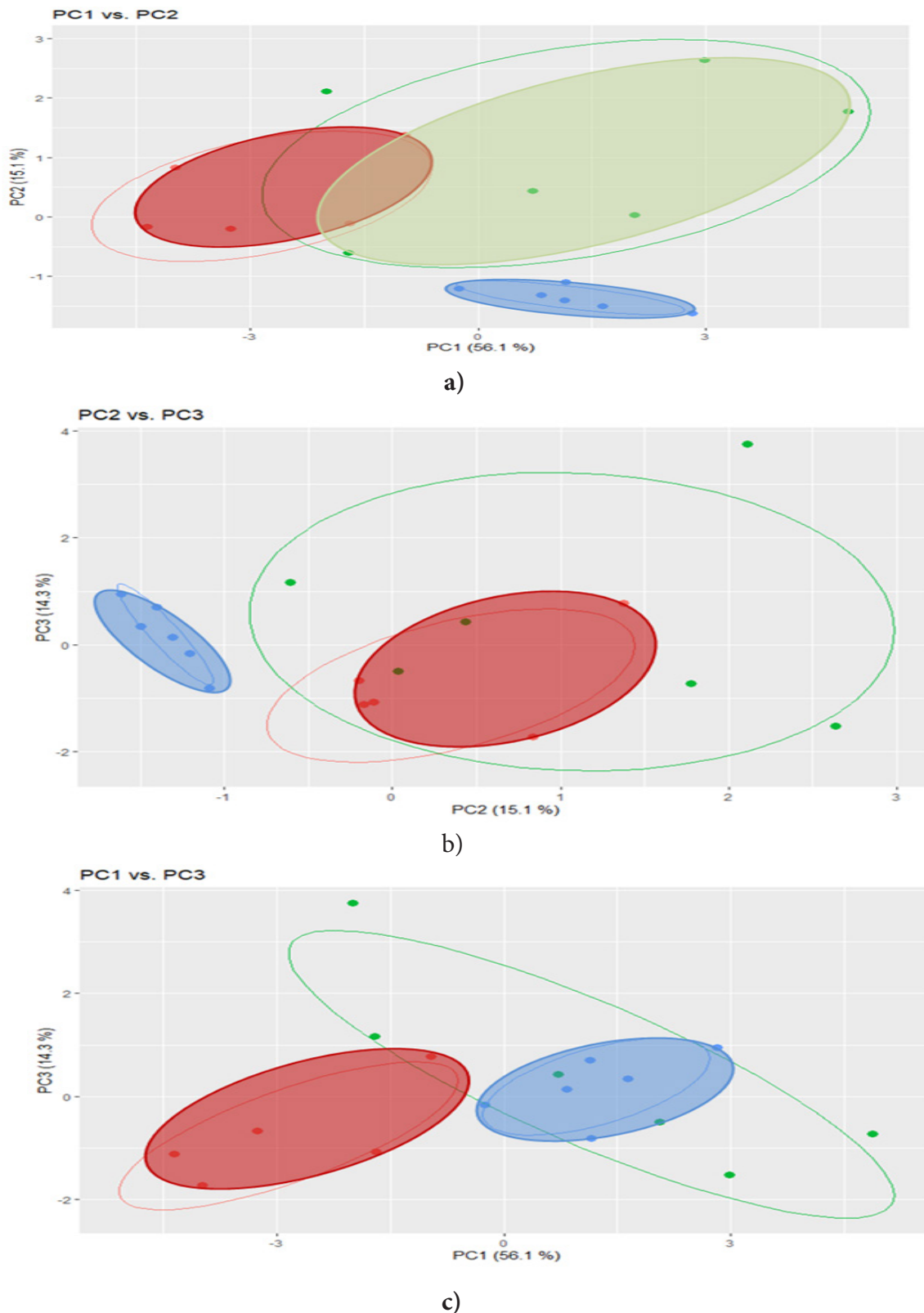

Fig. 4: Dot plot of the dependencies of the "significant" principal components and confidence regions (unhatched ellipse): a) PC1 vs. PC2; b) PC2 vs. PC3; c) PC1 vs. PC3 (red – Sample 1, green – Sample 2, blue – Sample 3)

Despite the slight overlap, the PCA score plot (see figure 4a) shows that the two Samples are clearly separated from each other. Sample 1 spectra were clustered in negative PC2 space (II and III quadrants), while Sample 3 spectra were clustered in negative PC1 space (III and IV quadrants). Sample 2 spectra did not tend to cluster and were located in all PC1 and PC2 spaces (I, II, III and IV quadrants), which may indicate that there are no significant differences in the obtained values.

Spectra of Sample 1 in the graph PC2 vs. PC3 (see figure 4b) are grouped in the positive and negative space of PC2 (I and II quadrants), positive and negative space of PC3 (I, III and IV quadrants), and spectra of Sample 3 are grouped in the positive and negative space of PC2 (II and III quadrants) and negative space of PC3 (II and III quadrants).

In the figure 4c, spectra of Sample 1 are grouped in the negative space of PC1 and the positive and negative space of PC3, which corresponds to II and III quadrants on the coordinate plane. Spectra of Sample 3 – in the positive and negative space of PC1 (I, III and IV quadrants) and the

positive and negative space of PC3 (III and IV quadrants). Besides, a grouping of spectral signal blocks was revealed (see figure 5 a,b,c). Group 1 (figure 5a) united signals 1080, 1266, 1438, 1650 and 1740 cm^{-1} on all graphs (see figure 5 a,b,c). Group 2 (highlighted with a green ellipse) – 1061, 1127, 1300 cm^{-1}. Signals of spectra 970, 868, 1418 and 1368 cm^{-1} are not included in any groups, or in neither of them. At that, peak 868 cm^{-1} is located in II quadrant, which corresponds to the positive space of PC1 and the negative space of PC2. Peak 1368 cm^{-1} is located in IV quadrant, which corresponds to the negative space of PC1 and the positive space of PC2.

In the graph PC1 vs. PC3 (figure 5b), signals of spectra 868, 1368, 1418 and 1438 cm^{-1} do not belong to any of the formed groups.

In the graph PC2 vs. PC3 (figure 5c), peak 868 cm^{-1} is located in IV quadrant, and peak 1368 cm^{-1} is in III quadrant, which corresponds to the negative space in both cases of PC2. However, in PC3 space, peak 868 cm^{-1} is positive, and peak 1368 cm^{-1} is negative.
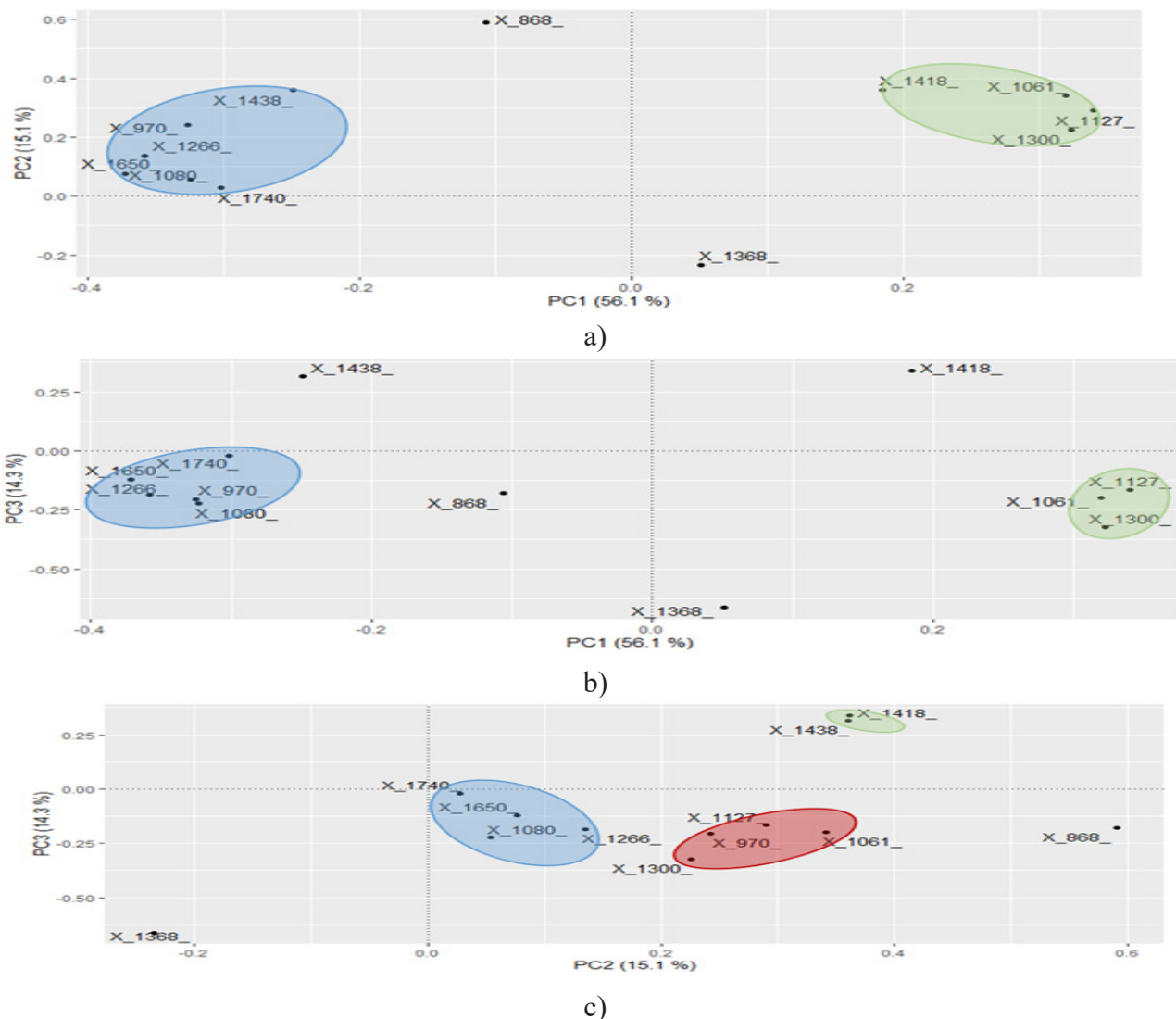

Fig. 5: Graph of variables (spectra) in coordinates of significant principal components: a) PC1 vs. PC2; b) PC2 vs. PC3; c) PC1 vs. PC3

In all graphs, spectra 1061, 1127 and 1300 cm^{-1} form a separate cluster of FA groups. These data are similar to the results of Szykuła (2023), who suggested that these spectra are responsible for the interspecies identification of pork. Of further note is the presence of a peak in region 1410-1418 cm^{-1} in all samples (see figure 6), which, according to Motoyama (2010), is a unique marker of pork and is detected at a pork concentration in a mixture of more than 50%. At the same time, Indastri (2010) showed that the presence of trans c18:3 ω3t, **c**20:3 ω3t and **c**20:2 ω6 FA is a marker of pork, and docosenoic (erucic) FA is a marker of beef.

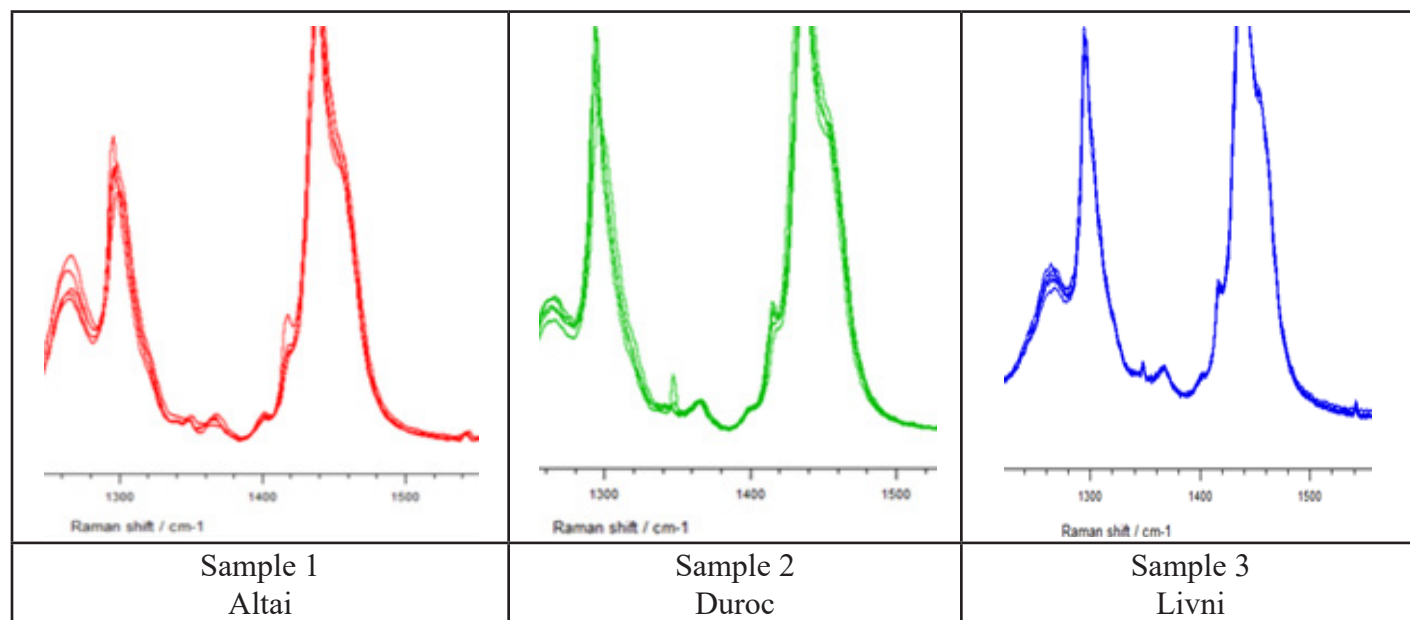| Sample 1	Sample 2	Sample 3
Altai	Duroc	Livni

Fig. 6: Fragments of Raman spectra of fat profiles. The sector with wavelength 1410-1420 cm^{-1} is highlighted in an oval

Thus, spectra of meat (fat) raw material samples can be divided into qualitative groups, breeds, and classified.

FUNDING

CONCLUSION

Raman spectra of pork lard of three different breeds were compared using PCA. The contribution of "significant" principal components to the total variance is 85%. The use of PCA to Raman spectral signals allowed us to: 1) remove unnecessary "disturbances" and identify the main patterns; 2) reduce the number of variables that carry little information. As a result of transformation, we reduced the matrix of the original data (17×12) to 3 "significant" variables for all signals of the studied spectra (12×3); 3) identify the most important spectral signals that explain the largest part of the data variability; 4) show the differences between animals of the same species but different breeds; 5) confirm the presence of pig-specific signals - wavelength 1410-1420 cm^{-1}; 6) determine the differences between the data, which is necessary for classification and clustering of the studied samples.

The advantage of PCA is that it is a mathematical analytical tool that allows us to identify natural changes in a large data set without prior knowledge of this data structure. Thus, using PCA analysis, we have demonstrated the possibility of comparative evaluation of Raman profiles of pork lard, identifying the most characteristic signal values for fat and assessing the quality of fat from the standpoint of the ratio of saturated and unsaturated FAs.

Visualization of the data showed clearly samples divided into clusters corresponding to breeds. This confirms the significant differences in fat composition between samples. Samples that fell out of clusters indicated anomalies that may be related to individual variations.

This will be studied further.

REFERENCES

Abbas O, Fernández Pierna JA, Codony R, von Holst C, Baeten V (2009) Assessment of the discrimination of animal fat by FT-Raman spectroscopy. Journal of Molecular Structure. 924-926: 294–300

Bartlett MS (1937) Properties of sufficiency and statistical tests. Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences. 160(901): 268–282.

Berhe DT, Eskildsen CE, Lametsch R, Hviid MS, van den Berg F, Engelsen SB (2016) Prediction of total fatty acid parameters and individual fatty acids in pork backfat using Raman spectroscopy and chemometrics: Understanding the cage of covariance between highly correlated fat parameters. Meat Science. 111: 18-26.

Boyaci İH, Uysal RS, Temiz T, Shendi EG, Yadegari RJ, Rishkan MM et al (2014) A rapid method for determination of the origin of meat and meat products based on the extracted fat spectra by using of Raman spectroscopy and chemometric method. European Food Research and Technology. 238(5): 845–852.

Broadhurst C, Schmidt WF, Nguyen JK, Qin J, Chao K, Kim MS (2018) Gradient temperature raman spectroscopy of fatty acids with one to six double bonds identifies specific carbons and provides systematic three dimensional structures. Journal of Biophysical Chemistry. 9(1): 1-14.

Cattell RB (1966) The Scree Test for the number of factors. Multivariate Behavioral Research. 1(2): 245–276.

Chernukha I, Kotenkova E, Pchelkina V, Ilyin N, Utyanov D, Kasimova T et al (2023) Pork fat and meat: A balance between consumer expectations and nutrient composition of four pig breeds. Foods. 12(4): 690.

Czamara K, Majzner K, Pacia M, Kochan K, Kaczor A, Baranska M (2014) Raman spectroscopy of lipids: A review. Journal of Raman Spectroscopy. 46: 4-20.

Deegan KC, Holopainen U, McSweeney PLH, Alatossava T, Tuorila H (2014) Characterisation of the sensory properties and market positioning of novel reduced-fat cheese. Innovative Food Science and Emerging Technologies. 21: 169-178.

Dhanapal L, Erkinbaev C (2024) Portable hyperspectral imaging coupled with multivariate analysis for real-time prediction of plant-based meat analogues quality. Journal of Food Composition and Analysis. 126: 105840.

Gao F, Ben-Amotz D, Zhou S, Yang Z, Han L, Liu X (2020) Comparison and chemical structure-related basis of species discrimination of animal fats by Raman spectroscopy using near-infrared and visible excitation lasers. LWT. 134: 110105.

Gómez-Mascaraque LG, Kilcawley K, Hennessy D, Tobin JT, O'Callaghan TF (2020) Raman spectroscopy: A rapid method to assess the effects of pasture feeding on the nutritional quality of butter. Journal of Dairy Science. 103(10): 8721-8731.

Indrasti D, Che Man YB, Mustafa S, Hashim DM (2010) Lard detection based on fatty acids profile using comprehensive gas chromatography hyphenated with time-of-flight mass spectrometry. Food Chemistry. 122(4): 1273–1277.

Kabacoff RI (2022) R in action. Publishing: Simon and Schuster.

Kaiser HF (1960) The application of electronic computers to factor analysis. Educational and Psychological Measurement. 20(1): 141–151.

Li F, Lu Y, He Z, Yu D, Zhou J, Cao H et al (2024) Analysis of carcass traits, meat quality, amino acid and fatty acid profiles between different duck lines. Poultry Science. 103(7): 103791.

Li-Chan E, Nakai S, Hirotsuka M (1994) Raman spectroscopy as a probe of protein structure in food systems. In: Yada RY, Jackman RL, Smith JL (eds). Protein Structure-Function Relationships in Foods. Springer, Boston.

Liu Y, Xiang Y, Sun W, Degen A, Xu H, Huang Y et al (2024) Identifying meat from grazing or feedlot yaks using visible and near-infrared spectroscopy with chemometrics. Journal of Food Protection. 100295.

Logan BG, Hopkins DL, Schmidtke L, Morris S, Fowler SM (2020) Preliminary investigation into the use of Raman spectroscopy for the verification of Australian grass and grain fed beef. Meat Science. 160: 107970.

Lyndgaard LB, Sørensen KM, Berg F, Engelsen SB (2011) Depth profiling of porcine adipose tissue by Raman spectroscopy. Journal of Raman Spectroscopy. 43(4): 482–489.

Mastickij SE, Shitikov VK (2015). Statistical Analysis and Data Visualization with R. Moscow: DMK-Press. (In Russian)

Motoyama M, Ando M, Sasaki K, Hamaguchi H-O (2010) Differentiation of animal fats from different origins: Use of polymorphic features detected by Raman spectroscopy. Applied Spectroscopy. 64: 1244-1250.

Olsen EF, Rukke E-O, Flatten A, Isaksson T (2007) Quantitative determination of saturated, monounsaturated, and polyunsaturated fatty acids in pork adipose tissue with non-destructive Raman spectroscopy. Meat Science. 76: 628-634.

Ostovar pour S, Fowler SM, Hopkins DL, Torley PJ, Gill H, Blanch EW (2019) Investigation of chemical composition of meat using spatially off-set Raman spectroscopy. The Analyst. 144(8): 2618–2627.

Paula AM, Conti-Silva AC (2014) Texture profile and correlation between sensory and instrumental analyses on extruded snacks. Journal of Food Engineering. 121: 9–14.

Pchelkina VA, Chernukha IM, Korotkiy IA, Ilyin NA (2022) Study of adipose tissue of Kemerovo piglets: Detection of beige adipocytes. Theory and Practice of Meat Processing. 7(4): 265-272.

Pearson K (1901) On lines and planes of closest fit to systems of points in space. Philosophical Magazine. 2: 559-572.

Pomeranzev AL (2008) Principal Component Analysis. Retrieved from https://web.archive.org/web/20090528101113/http://www.chemometrics.ru/materials/textbooks/ pca.htm. Accessed May 19, 2024 (In Russian)

Robert C, Fraser-Miller SJ, Jessep WT, Bain WE, Hicks TM, Ward JF et al (2020) Rapid discrimination of intact beef, venison and lamb meat using Raman spectroscopy. Food Chemistry. 343: 128441.

Rodionova OY, Fernández Pierna JA, Baeten V, Pomerantsev AL (2021) Chemometric non-targeted analysis for detection of soybean meal adulteration by near infrared spectroscopy.

Food Control. 119: 107459.

Saleem M, Amin A, Irfan M (2021) Raman spectroscopy based characterization of cow, goat and buffalo fats. Journal of Food Science and Technology. 58: 234–243.

Sylvester JJ (1889) On the reduction of a bilinear quantic of the nth order to the form of a sum of n products by a double orthogonal substitution. Messenger of Mathematics. 19: 42-46.

Szykuła KM, Offermans T, Lischtschenko O, Meurs J, Guenther D, Mattley Y et al (2023) Predicting animal welfare labels from pork fat using raman spectroscopy and chemometrics. AppliedChem. 3(2): 279-289.

Vilanova M, Genisheva Z, Masa A, Oliveira JM (2010) Correlation between volatile composition and sensory properties in Spanish Albariño wines. Microchemical Journal. 95(2): 240-246.

Xiao L, Qi L, Fu R, Nie Q, Zhang X, Luo W (2024) A large-scale comparison of the meat quality characteristics of different chicken breeds in South China. Poultry Science. 103(6): 103740.

Zhilinskaya NV, Sarkisyan VA, Vorobieva VM, Vorobieva IS, Kochetkova AA, Smirnova EA et al (2018) Development of a marmalade for patients with type 2 diabetes: Sensory characteristics and acceptability. Food Science and Technology International. 24(7): 617-626.