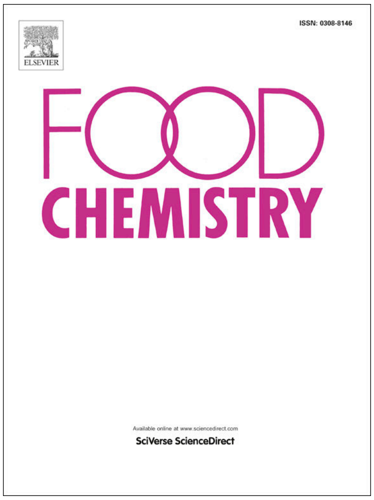

Analytical Methods

Effect of season, feeding, and anatomical region on the triacylglycerol profile of Iberian pig fat

Antonio Garrido-Fernández, Manuel León-Camacho

PII: S0308-8146(21)01076-1
DOI: https://doi.org/10.1016/j.foodchem.2021.130070
Reference: FOCH 130070

To appear in: *Food Chemistry*

Received Date: 23 November 2020
Revised Date: 11 April 2021
Accepted Date: 8 May 2021

Please cite this article as: Garrido-Fernández, A., León-Camacho, M., Effect of season, feeding, and anatomical region on the triacylglycerol profile of Iberian pig fat, *Food Chemistry* (2021), doi: https://doi.org/10.1016/j.foodchem.2021.130070

Effect of season, feeding, and anatomical region on the triacylglycerol profile of Iberian pig fat

Antonio Garrido-Fernández and Manuel León-Camacho*

[1] Food Biotechnology Department; [2] Lipid Characterization and Quality Department

Instituto de la Grasa, Campus Universitario Pablo de Olavide, Building, 46.

Ctra. Sevilla-Utrera, km 1. 41013, Sevilla

Running title: Compositional data analysis of Iberian pig fat triacylglycerols.

Corresponding author: Manuel León Camacho, Ph.D. e-mail: melon@ig.csic.es

Abstract

The work studies the effects of season, feeding type, and anatomical region on the Iberian pig fat triacylglycerol (TAG) profiles, considered as compositional data (CoDa). The analysis consisted of applying exploratory tools in the simplex and standard multivariate techniques to data transformed into the Euclidean space (*ilr coordinates*). Compositional biplot showed differences in TAG containing palmitic (P) and oleic (O) acids between the 2005 and 2003/2004 seasons but not within these. PPP (*clr* variance, 0.139), OLL (0.095), PPS (0.075), POPo+PLP (0.074), and PSS (0.629) showed high CoDa variability among treatments. The ANOVA analysis found significant ($p \leq 0.05$) effects of season, feeding type, and anatomical region on pig fat TAG profile, but only that for 2005 season was well predicted (97.5%) by discriminant analysis (DA). Overall, season was more influential on

the Iberian pig fat TAG profile than *montanera* length and sampling region, which effects were not significant for some statistical techniques.

Key words: Compositional data analysis, *montanera*, ham, rump, season, triacylglycerol.

1.Introduction

Triacylglycerol (TAG) is the primary lipid class (about 86%) in the adipose tissues of Iberian pigs (Perona & Ruiz-Gutierrez, 2005). Because of the low volatility of the TAG, the high-performance liquid chromatography (HPLC) was the habitual method used for TAG analysis (Díaz, García-Regueiro, Casillas, & De Pedro, 1996; Tejeda, Gandemer, Antequera, Viau, & García, 2002). The main identified components in pig fat by Díaz et al. (1996), using reverse-phase HPLC and light scattering detector, were (see Table 1 for full names of fatty acids) OOL, POL, OOO, POO (the most abundant), POP, SOO, and POS, followed by PPS, SSO, and PSS POO, in lower proportions. The three types of nutrition (*montanera*, *recebo,* and *cebo*) provided to pigs significantly affected the OOL/SOO and OOO/SOO TAG ratios.

However, new stationary phases' development permitted the gas chromatography (GC) use for the TAG's analysis. The new method allowed identifying in the Iberian pig fat seven more components (PPP, MOP, POPo + PLP, PLPo + MLO, PLL + PoLO, SOL and OOL) than the HPLC techniques. (Viera-Alcaide, Vicario, Graciani-Constante, & León-Camacho,_2007). The column resolves by atom carbon number (CN) and unsaturation degree, although several TAGs still elute together. POO was again recognised as the major component (Viera-Alcaide et al., 2007). Since then, the technique has been widely used. Among others, for studying the changes in the TAG profiles in the subcutaneous fat of Iberian ham during the dry-curing process (Narváez-Rivas, Vicario, Graciani-Constante, & León Camacho, 2008), the TAG composition of the subcutaneous adipose tissue and its relationship with the fattening diet and genotype (Viera-Alcalce, Vicario, Escudero-Gilete, Graciani-Constante, & León-Camacho, 2008), or the differentiation between intensive and extensive fattening systems by TAG profiles (Gallardo, Narváez-Rivas, Pablos, Jurado, & León Camacho, 2012). Then, the TAG analysis by GC has been validated by numerous publications in relevant analytical chemistry and food science Journals and is extensively used nowadays.

Several authors have studied the influence of fattening, *montanera* length (46 and 154 days), and the anatomical region on the TAG profile of pig fat. Most of them reported significant effects of these factors (Viera-Alcaide et al. 2007; Narváez-Rivas, León-Camacho, & Vicario, 2009; Viera-Alcaide, Vicario, Escudero-Gilete, Graciani-Constante, & León-Camacho, 2008). To notice the significantly higher amount of oleic acid in purebred than in crossbred pig fat, even considering animals fed on the same diet (Ventanas, Ventanas, Jurado, & Estevez, 2006). However, other authors only found a limited influence of the growing conditions (Tejeda et al., 2002). Overall, the literature on the TAGs in pig fat is still scarce since most of the research in this field was preferentially devoted to fatty acids. Furthermore, the previous studies have always use standard multivariate analysis.

The triacylglycerols in Iberian pig fat is usually expressed as proportions, and their sum is 100. Aitchison (1986) classified this type of information as compositional data, characterized by non-negative elements (e.i., have a sample space restriction to R_{+}^{D}) and a constant sum, usually 1 or 100. Geometrically, compositional data with D components belong to a sample space of the regular unit D-simplex, S^D (Bacon-Shone, 2011). However, the standard multivariate analysis assumes that the sample's space is R^D. Therefore, its application to these data is inappropriate, at least formally.

The concept of compositional data is nowadays focused on the original idea of Aitchison (1986): the relative information they carry. For Filzmoser et al. (2018), relative information refers to representing the quantitatively described contribution on the whole. In most cases, the units in which data are expressed already indicate their relative nature, and, regardless of the chosen definition, they should be treated as compositional data and be subjected to the specific methodologies developed for such a type of information.

The work aims to study the effect of season, *montanera* length, and anatomical region on the Iberian pig fat's triacylglycerol composition. The statistical analysis assumed that this data set represents a clear example of compositional data (sum constraints to 100). Hence, the strategy followed consisted of i) first study by exploratory tools specially developed for working-in-the-simplex, and ii) the application of the standard multivariate methods to the original values transformed into *coordinates* (in the Euclidean space). Results also were interpreted according to the CoDa principles.

2.-Material and methods

2.1 Experimental design

The experiment was performed along three successive seasons (2003, 2004, and 2005). The number of pigs was always chosen, in agreement with the farmer, among those having similar growing conditions and weights. Then, the animals were randomly assigned to two *montanera* lengths (short and long) for final fattening. After slaughtering, the animals were sampled at two anatomical regions, the subcutaneous fat from the rump (R) and the adipose tissue covering the *B. femoris* muscle (H). Hence, the experiment consisted of a complete factorial design for season (three levels), *montanera* length (two), and anatomical sampling region (two), which makes a total of 12 treatments. The distribution of animals between seasons was 208, 136, and 56 for 2003, 2004, and 2005, respectively. The assignation to *montanera* length within season was as follows: 2003, M1 (short montanera)=15, M2(long montanera)=89; 2004, M1=30, M2=38; 2005, M1=1 (because of farmer compromises), M2=27 for 2005..

2.2 Sampling

From each animal slaughtered, two samples of adipose tissue were taken. They consisted of a chunk (3 x 3 - 5x5 cm) of at least 6 mm thickness from the rump (R) and the subcutaneous adipose tissue covering the *B. femoris* muscle of both hams (H). In the first case, the sample was withdrawn at approximately 10 cm of the tail, following the back line, and consisted of skin, adipose tissue, and some loin. In the laboratory, the skin and the loin were removed, and the remaining chunks punctured and homogenised before extraction. In the second, the homogenisation was achieved after mixing the two portions. Then, the fats from the two anatomical regions were obtained by melting their corresponding samples for 3 min in a microwave oven at 360w power, filtered, and stored at -20ºC . The sample's analysis was achieved just after each season experiment whereas publication required a confidential period.

2.3. Triacylglycerol analysis

TAGs were determined by GC (Viera-Alcaide et al, 2007) in a Varian 3800 (Varian Co, Palo Alto, CA, USA) using a fused silica capillary DB-17HT column (30m x 0,32 mm I.D., 0,15 µm film thickness). The oven temperature was kept at 330 ºC, then raised up to 360 ^{o}C at a rate of 2.0 ^{o}C/min, and hold isothermally for 10 min. The injector temperature was kept at 360 ^{o}C, while the detector temperature was set 370 ^{o}C. Hydrogen (30.0 psi column head pressure) was used as the carrier gas and nitrogen as make-up gas.

The assignment of the chromatographic peaks was done by means of standards, which allowed deducing the carbon number of the components associated with each peak group and the difference between the retention times of the triacylglycerols (Table 1). Figure 1S shows a representative chromatogram reporting the TAGs from the subcutaneous fat and the identification of peaks. The quantification of the TAGs was achieved by using trimiristin as an internal standard and assuming an equal response factor for all TAG species.

2.4 Triacylglycerol database

The results from the analyses were gathered in a data set, where rows corresponded to samples (experimental combinations per pig and columns to their TAG concentrations. Before statistical analysis, the data set was checked for abnormal and atypical values (assuming Additive Normal Distribution), using a threshold of p=0.05. The depurated data set consisted of a table with

400 rows (samples) and 20 columns (season, *montanera* lenght, and anatomical sampling location plus 17 TAGs, some of which were the combination of two components). Each data cell (x_{ij}) represents the proportion in sample *i* of the *j* TAG.

2.4. Statistical analysis

2.4.1. Central tendency and dispersion

The classical mean and standard deviation (developed for data in the Euclidean space) have no meaning in CoDa. Instead, the geometric mean and logratio variances are used.

a) Central tendency. For a dataset (X) of size n and D parts, the best representation of the center is the closed geometric mean (barycenter), defined as:

$$g_m = C[g_1, g_2, \dots g_D] \text{ where } g_j = \left(\prod_{i=1}^{n} x_{ij}\right)^{1/n} \qquad \text{(Eq. 1)}$$

where the *C* operator stands for closure; that is, for rescaling g_j to constant *k* (usually 1 or 100).

*b) Dispersion/variability.*The range of values is evaluated by estimating the corresponding 0, 25, 50 (median), 75, and 100 percentiles in the usual way, although the parameter has a limited application because of its dependence on the scale. Instead, because of these data's relative information, the matrix of logratio variances is used. For a matrix X, the relative variability is defined as:

$$T = \begin{pmatrix} t_{11}\, t_{12} \dots t_{1D} \\ t_{21}\, t_{22} \dots t_{2D} \\ \dots\ \dots\ \dots\ \dots \\ t_{D1} t_{D2} \dots t_{DD} \end{pmatrix} \text{ where } t_{ij} = var\left(ln\frac{x_i}{x_j}\right) \qquad \text{(Eq. 2)}$$

The matrix is symmetric, and its diagonal terms, by definition, are zeros. In the so-called variation array, the upper half of this matrix consists of the logratios variances while the bottom half (image of the upper one) is substituted with the information on the logratio $\left(ln\frac{x_i}{xj}\right)$ means.

The total variance, which summarises the variability of the whole dataset in a single value, is:

$$totvar(X) = \frac{1}{2D}\sum_{i=1}^{D}\sum_{j=1}^{D} var\left(ln\frac{x_i}{x_j}\right) \qquad \text{(Eq. 3)}$$

2.4.2. Transformation of compositional data into *coordinates*

The most common transformations are the centered logratio, used for building the CoDa biplot, and the isometric logratio, which leads to orthonormal *coordinates* in the Euclidean space.

a) *Centered logratio transformation (clr).* It is defined as:

$$Z = clr(X) = \left[ln\frac{x_1}{g_m(x)}, \ldots, ln\frac{x_D}{g_m(x)}\right] \qquad \text{(Eq. 4)}$$

which is estimated row-wise. The variability of each *clr* part (*clr* variance) is estimated, column-wise, using the standard procedure. The sum of all clr variances also represents the total variance.

b) *Isometric logratio transformation (ilr)*. The most common process for its calculous is the so-called sequential binary partition (SBP) (Egozcue, Pawlowsky-Glahn, Mateo-Figueras, & Barceló-Vidal, 2003; Pawlowsky-Glahn, Egozcue, & Tolosana-Delgado, 2015). It consists of a successive segregation of the composition into two mutually exclusive groups. The calculous ends when the group has only two parts (hence, the number of balances is always D-1). The *coordinates* (balances) are defined as the normalised log ratio of each group of parts, that, row-wise, is estimated by:

$$z_i = \sqrt{\frac{rs}{r+s}}\, ln\frac{(x_{i1}x_{i2\ldots}x_{ir})^{1/r}}{(x_{i1}x_{i2\ldots}x_{is})^{1/s}} \qquad \text{(Eq. 5)}$$

For each balance, the numerator, denominator, or absent parts are coded by 1, -1, and 0, respectively. They are usually gathered in a matrix, which is helpful for back transformation. Regardless of the basis used for the transformation, the new data can be subjected to the standard multivariate analysis. The standard multivariate techniques were always applied to the corresponding *ilr* transformed data (*ilr coordinates*).

Other CoDa techniques used in this work are briefly introduced in the text. For more extensive details on CoDa analysis, readers may consult specialised books (Pawlowsky-Glahn et al., 2011; Pawlowsky-Glahn et al. 2015; Van den Boogaart & Tolosana Delgado, 2013; Filzmoser, Hron, & Templ 2018; Greenacre, 2018). Ros-Freixedes & Estany (2014) also includes an extensive overview of the essential CoDa tools.

For the CoDa analysis, CoDaPack (Comas-Cufí & Thió-Henestrosa, 2011) and the R packages *compositions* (Van den Boogaart & Tolosana-Delgado, 2013) and *robCompositions* (Templ, Hron, & Filzmoser, 2011) were used. R packages were run in R v 4.03 (R Core Team, 2020), under RStudio v.1.4.1103 (RStudio Team ,2020). Besides, IBM SPSS Statistics for Windows v.22 (IBM Corp., Armonk, N.Y., U.S.A) performed bootstrapping DA.

3. Results and discussion

3.1 Overall triacylglycerol composition of pig fat (central tendency)

The geometric means according to season levels (Table 2) revealed that the more abundant triacylglycerols in Iberian pig fat (1-100 scale) were POO, PSO, OOO, and PLO, all of them abundant in oleic acid. Differences between the standard overall arithmetic means (last column) and the geometric means are scarce. However, this should not be an argument to discredit CoDa analysis since the arithmetic means does not fit the Atchison geometry (Aitchison, 1986).

A rough measure of the effect of variables can be achieved by comparing the geometric means at each level and the overall value. Regarding season, POO and OOO have substantially higher geometric means in the 2005 season, whereas, simultaneously, many other components (PPP, POP, PSO, among others) show lower values. The differences point out to season as an influential variable. However, the changes may also be caused by the estimation methodology (sum constrained to 100), which automatically lead to spurious negative correlations (Pawlowsky-Glahn et al., 2015; Filzmoser et al., 2018; Greenacre, 2018). Regarding *montanera* or anatomical location, one may still observe slight modifications in the contents of some TAGs with respect to the overall geometric mean or within the levels of each factor, but always at a lower scale than for season.

In summary, the comparison of the overall central tendency (geometric means) and those according to variable levels revealed some clues on the influential variables on pig fat TAG profiles. However, one should be cautious when interpreting the changes since they can be a consequence of the data structure (usually not considered when applying standard multivariate techniques). Besides, their correlations could be spurious.

3.2. Evaluation of the triacylglycerol variability among treatments (compositional data dispersion)

Percentiles are appropriate for the appreciation of the data dispersion, but they do not play any role in CoDa analysis because of their dependence on the scale. On the contrary, the variation array (defined in Section 2.4.1) is strongly associated with these data's relative information and is used by several stay-in-the-simplex statistical tools. The logratios of PPP and OLL over the remaining variables showed the highest variance (Table 1S, **in bold**). Besides, an estimation of the variance associated with each triacylglycerol is obtained by the standard procedure from its *clr* coefficient (non-orthonormal), and it is usually included at the right of the variation array matrix. The highest values corresponded to PPP (Table S1). In some circumstances, few logratios may condense most variability and be useful for segregating treatments (Greenacre, 2018).To notice, however, that these variances might not always reflect the effect of the design variables (Egozcue & Pawlosky-Glahn, 2005; Pawlowsky-Glahn & Egozcue, 2011) but instead determination errors.

Low values in the variation array also indicate a similar evolution (correlation) in the two TAGs involved in the logratio. However, there is no complete agreement on this interpretation (Filzmoser et al., 2018). Data below the diagonal are the corresponding components' logratio means over treatments and show their relative abundance.

3.3. Graphical exploration of factors' effects on the pig fat triacylglycerol composition

The tetrahedral, based on four components, is the most complex allowed plot. Those TAGs with the highest *clr* variance presented the highest segregation power (Fig. 2S A), placing 2005 samples separately from the overalping 2003/2004 samples. In the graph for *montanera* length (Fig. 2S B), the segregated samples (bottom left) corresponded to the 2005 season; that is, the season's effect predominated over that of *montanera*. Plotting the data according to the sampling region, there was a complete overlapping (not shown) because the composition in both anatomical parts was the same regardless of the other variables.

Bivariate plots are sometimes a useful alternative. As the greatest dispersions corresponded to PPP and OOL, a plot as a function of their logs could be suitable (Fig. 3S). The segregations obtained were worse than in the tetrahedral display because seasons were poorly segregated and *montanera* and anatomical levels strongly overlapped. In general, tetrahedral plots achieved better segregation by groups than bivariate plots (frequently used in the literature using standard statistical), possibly due to the extra variance contribution of the two additional axes.

3.4. Main transformations of the triacylglycerol into the Euclidean space and their uses

The search for transformations to move the original data (in the simplex) into the Euclidean space, where the standard multivariate tools can analyse them, was continuous since the compositional concept was formulated (Aitchison, 1986).

The *clr* transformation has the advantage of being easily related to the original variables but has some mathematical drawbacks. Filzmoser et al. (2018) suggest using the term *coefficient*, instead of *coordinate*, for the value obtained (not in orthonormal axes). The *ilr* transformation is the most widely recommended, but the interpretation of the standard multivariate analysis strongly depends on the basis and, frequently, is not straightforward. Apart from other applications, *coefficients* and *coordinates* support interesting exploratory techniques.

3.4.1. Application of centered logratio (*clr*) and compositional biplot to evaluate the effects

The compositional biplot was developed by Aitchison & Greenacre (2002). In it, row and column points are both centered at the origin of the display. It comprises two options: covariance (appropriate to study the components) and form biplot (which preserves row distances) (Fig. 1 A and B, respectively). The two components account for 81.74% of the total variance and may adequately reflect the data structure. In the covariance biplot, the distance between the ends of two rays is the standard deviation of the corresponding logratios. The largest distances (variances) correspond to the PPP/OLL or PPP/OOL logratios, which agree with their variation array values. Their angles, in turn, are close to zero, indicating a strong correlation between them. When the ends of two rays coincide (e.g. OLL, OOL, SOL, OOO, or POO), their logratios are almost constant and, subsequently, their variances are zero or near zero. Therefore, these variables provide redundant information, and only one could be enough for statistical analysis.

The loadings of the first two PCs of the biplot (Table 2S) are not free to vary, but linked to each other through their null sum. In general, low values should not be interpreted as an absence or dependence, but in terms of a ratio. For example, PPP /SOS increases along the PC1 with a slope of 0.49432 (0.50137-0.00705) or PPS/OLL remains almost constant along PC1, with a slope of 0.00436 (0.23768-0.23332). Similar comments apply regarding PC2 and the rest of the components (presented in Fig. 4S). It is also possible to plot in a graph the evolution, relative to one component, of the complete set of variables, presenting the observations as dots and the changes as lines (Fig. 5S). There was a general increase of all the components (relative to PPP) along PC1, which become linear when expressed as logratios.

In the form biplot (Fig. 1 B), distances between points represent those between cases. The highest segregation power relies on PC1, which is positive and negatively related to TAGs containing P, and O, respectively. The 2005 samples are well segregated (except for some overlapping with samples from other seasons). To notice the evident, opposed, correlation between SOS and PLPo+MLO, which logratio is, in turn, highly correlated to PC2.

CoDa biplot results only showed TAG changes because of season but not for the fattening period or sampling location. However, Viera-Alcaide et al. (2008) segregated *montanera* and *cebo* fattening, possibly due to the highest major differences in the feeding stuff's composition. The most differentiating variables were OOO, OOL, and PSO, reaching a perfect separation between fattening types. The fist two TAGs also were selected in the logratios (OOO/PPP, OOL/PPP, and OLL/PPP, among others) responsible for the separation of season 2005 samples. On the contrary, our results agree with those reported by Segura, Cambero, Cámara, Loriente, Mateos, & Lòpez Bote (2015), who observed that the dietary intervention during the fattening phase hardly affects TAG composition. However, the effect of season is hardly mentioned despite de strong impact that climate can have on the natural feeding stuffs, as found in this work.

3.4.2. Application of the isometric logratio transformation (*ilr*) and CoDa dendrogram to study the effect of factors on the triacylglycerol profile of Iberian pig fat

The *ilr* transformation moves the original data into *ilr coordinates* in the Euclidean space. The corresponding *ilr coordinates* may then be subjected to the standard multivariate methods. Interpretation of these analyses is preferably made after back transformation into the original units. The *ilr* values (*coordinates*) are usually obtained by SBP (defined in section 4.2.2), and the results from the further multivariate analysis are invariant regarding the basis. Nevertheless, the *coordinates* and their explained variances depend on the TAGs selected in each step.

a) Sequential binary partition following the decreasing order of variances. The SBP (defined in Material and Methods) was applied by selecting, one by one, the TAGs in the numerator according to their decreasing order of *clr* variances. In this way, the successive *ilr coordinates* (or balances) approximately explains less and less variance. As an illustration of the calculous, the estimation of the first coordinate (z_1) (b_1) was as follows:

$$z_1 = \sqrt{\frac{1 \cdot 16}{1+16}}\, ln \frac{(PPP)^{1/1}}{(MOP \cdot PPS \cdot POP \cdot (POPo + PLP) \cdots\cdots SOO \cdot OOO \cdot SOL \cdot OOL \cdot OLL)^{1/16}}$$

For the estimation of the second balance, the PPP component was discarded, and the formula was then applied likewise to the rest of TAGs, up to the last two. The identification of the successive TAGs in the numerator, denominator, or absent is summarised in a matrix of 1, -1, or 0, respectively (Table 3S). Finally, the concentration of each compound was substituted in the original database, row-wise, with the equivalent *ilr* (*ilr1*, *ilr2*, … *ilr(D-1)*, except the last two, which were replaced by only one column. The new data set was then composed of 400 rows and 16 columns (D-1). Their averages and variances (Table 4S, under the column "Order of variances") can now be estimated following the standard methods for data in the Euclidean space. Notice the existence of negative and positive *coordinates* (or balances), as corresponds to the new data set class. Besides, the first *coordinate* or balance (0.1481) accounts for the highest proportion (19.5%) of variance (0.7591), the following three also represent an important contribution (approx. 10%), but the remaining ones usually have low variances. The SBP following the decreasing order of *clr* variances had the advantage of explaining most of the variance in the first balances and facilitating the effects' analysis.

The means of balances (intersection on the horizontal line) and variances (vertical line above the mean) are usually plot in the so-called CoDa dendrogram (Fig.6S). When the balances are estimated according to a grouping variable (season, *montanera* length, or anatomical region), their averages by levels (in boxplots) are plotted below the overall balance. The comparison within them or their average is then straightforward (Fig.6S). The box for the 2005 is displaced regarding 2003/2004 seasons, mainly in the four first balances. The plot also points out to the high dispersion (larg green vertical bar) within seasons.

b) SBP following the clustering sequences. In this option, SBP roughly followed the segregation suggested by the cluster dendrogram (Ward D.2 method) (Fig. 7S). The first partition corresponds to the two groups with the most remarkable dissimilarity. Later, each cluster's components were successively separated into two groups until the last balance (Table 5S). Therefore, most of the sequences consisted of several triacylglycerols in the numerator and denominator. According to Fig. 7S and Table 5S, the first balance was:

$$z_1 = \sqrt{\frac{6 \cdot 11}{6+11}}\, ln \frac{(PPP \cdot MOP \cdot PPS \cdot POP \cdot (POPo + PLP) \cdot PSS)^{1/6}}{((PLPo + MLO) \cdot PSO \ldots\ldots SOO \cdot OOO \cdot SOL \cdot OOL \cdot OLL)^{1/11}}$$

The following *coordiantes* (z_i) or balances were built similarly from Fig. 7S and Table 5S to end with:

$$z_{16} = \sqrt{\frac{1 \cdot 1}{1+1}}\, ln \frac{(SOO)^{1/1}}{(SOL)^{1/1}}$$

As in the previous case, the number of balances was 16 (with both positive and negative values), and the first accounted for the highest variance (0.3756) and contribution (49.5%). Notice that now, the first four balances had more segregation power (Fig 2, 2005 is displaced regarding 2004/2003 seasons) than that obtained following the decreasing order of *clr* variances y explains a greater proportion of variances (Fig. 6S). Then, the SBP based on clustering would be preferable to that obtained by decreasing order of variances. No differences between *montanera* lengths and sampling locations were observed in this analysis.

The results of the CoDa dendrogram were similar tho those observed in the CoDa biplot and are in contrast to those reported by Viera-Alcaide, Vicario, Escudero-Gilete, Graciani-Constante, & León-Camacho, 2008), who found significant effects of the fattening diet (*montanera*, *cebo* and *recebo*) and genotypes (Iberian purebred and IberianxDuroc crossbred) on the TAG composition of the subcutaneous adipose tissue. However, other authors have reported a limited influence of the genotype on the fat TAG composition (Tejeda et al., 2002). Several studies are mainly devoted to changes in fatty acids. Daza, Olivares, Latorre, Rey, Callejo, & López Bote (2017) reported high differences in the subcutaneous backfat saturation, inner layer, and hepatic fat. Ventanas, Ventanas

(J), Jurado, & Estevez (2006) also reported significantly higher amount of oleic acid in purebred than in crossbred pigs, even for animals fed on the same diet. These results suggest that the changes could be due to the diverse Δ^9-desaturase enzyme activity in each genotype.

3.5. Evaluating the effect of treatments on the pig fat triacylglycerol profile by compositional analysis of variance (ANOVA)

The compositional analysis of variance follows the same model as a regression (van der Boogaart and Tolosana-Delgado, 2013): $Y_i = a \oplus \mathrm{b} x_i \oplus \varepsilon i$, with a and b_g, $g \epsilon$ $\{g_1, \dots, g_m\}$ being compositional constants in S^D, and ε_i a compositional random variable with the neutral expected value, modelled as a normal distribution on the simplex ($\mathcal{N}_S^D(1,\Sigma)$). Since in ANOVA one must fix the contrasts for the factors, there are several different compositional contrasts: a, b_{g2}, b_{g3}, …, b_{gm} (a different one for each group g_i, i=1 ,…., m), which are distinguished by their subscripts.

As generally suggested for performing the analysis, *ilr* transformation is applied to map the data to a multivariate linear model:

$$\mathrm{ilr}Y_i=\mathrm{ilr}(a \oplus X_i \oplus \varepsilon_i)=\mathrm{ilr}(a)+\mathrm{ilr}(b_{Xi})+\mathrm{ilr}(\varepsilon_i)$$

The analysis can be applied to multiple variables by adding the appropriate terms into the modelling equation:

$$\mathbf{Y}_i=a \oplus (X_1) \odot b_1+X_{2i} \odot b_2 \oplus b_3(X_3) \oplus b_4(X_4) \oplus \varepsilon_i$$

where the covariables can be continuous or discrete. The linear model in R is accessed through the command *lm*. When several covariables are involved, each line of the ANOVA corresponds to a test of the form: H_0: given the preceding (i-1) variables, the i^{th} covariable has no influence (the composition has the same expectation for levels); H_1: the i variable, in addition to the already variables in the model, will improve the model. Then, to check each variable's additional contribution, it is necessary to situate each of them successively in the last position. The variable will be removed when its contribution is not significant. In this work, the analysis only included categorical covariables.

The function was: model=lm(ilr(Y)~X_1+X_2+X_3+$X_1 \cdot X_2$+$X_1 \cdot X_3$+$X_2 \cdot X_3$). The joint significance of the covariables was computed by a generalised ANOVA table, using the function anova.mlm. After performing it with each term in the last position, only the interaction $X_2 \cdot X_3$ was not significant and, subsequently, removed.

In the final model, all the terms were highly significant (Table 6S) and explained 40.65% of the variance (R-square=0.4065485). However, the models with only X_1, X_1+X_2, X_1+X_2+X_3, and X_1+X_2+X_3+$X_1 \cdot X_2$, explained 37.51, 38.89, 40.27, 40.28%, respectively; that is, the most important contribution to variability was season while that of the progressive incorporation of the rest of the terms was limited; however, their low p-value is an argument for their retention. In fact, this is the only analysis in the study that detected differences among *montanera* and anatomical regions. In summary, season was the most influential variable, although the three factors and the interaction of season and *montanera* with anatomical sampling location also were significant.

For the interpretation of results, the models' parameters in *ilr coordiantes* were back-transformed into the original units (Table 7S). The intercept should be interpreted as the expected composition at the lowest levels of all factors, while the values of the remaining columns represent the contribution of each factor or interaction beyond the first. The coefficients can also be presented as bar charts (Fig. 8S).

The model goodness of fit, plotted in 2D according to season and *montanera* (the most relevant variables), is appropriate (Fig. 9S). The graph also shows the increase of the most abundant component (POO) and the decrease of those present at intermediate levels in the 2005 season (in blue), with the discrepancies being slightly higher for the moderately abundant TAGs. The plot of residuals presented as boxplots displays reasonably symmetry (Fig. 10S). Besides, it confirms that PPP and PSS have some far observations (not considered as atypical values) but, in general, agrees with the assumption of normality and shows a limited presence of outliers. The quantile-quantile plot also presents a good fit for most of the data, with only a few samples separated from the linear trend.

The model can also be used for predictions. The concentrations obtained (in the original units, percentages) for the diverse combinations of factors agree with those observed and reproduce the high content of POO as well as the decrease (always in lower proportions) in other TAGs (Table 3).

Results from this work agree with those of Díaz et al. (1996) in identifying the most abundant TAGs and the detection of significant differences between *montanera* lengths. They assigned the greatest segregating power to the ratios OOL/SOO and OOO/SOO but did not check the season's effect. Interestingly, by using the TAG ratios, in some way, these authors assume that data carry relative information. Viera-Alcaide et al. (2008) found a significant correlation between most TAGs and the type of feeding (*montanera*, *cebo*, and *recebo*), but the standard multivariate analysis did not show clear differentiation among feeding groups. Narváez-Rivas et al. (2009) reported differences in the composition of fat from ham and rump for most TAG, concluding that those from the rump showed low saturation; apparently, the differences were related to the intensity of the oxidative metabolism. They also studied season's effect, reporting an increase in the ratio oleic/stearic when the natural resources were low (high feeding stuff). Notice that, unconsciously, the ratio of components (as proposed by CoDa) again appears as a segregation parameter in this work.

3.6. Evaluation of the effect of factors on pig fat triacylglycerol profile by discriminant analysis

Classification assumes prior information of the class membership of cases. The task of DA is to predict the class for a test set of observations. Bayes and Fisher are the most-used rules, while ~~k~~ linear (LDA) or quadratic (QDA) functions are the most common methods. In general, LDA tends to underfit, whereas QDA tends to overfit. LDA is the most widely used method since it requires the lowest number of parameters. As recommended in CoDa analysis, the DA was applied after transforming the original data into *ilr coordinates*.

3.6.1 Bayes discriminant rule

LDA (classical or robust procedures) and QDA, based on Bayes' rule, were applied for the classification of the samples in *robCompositions* (Filzmoser et al., 2018), using *pivot coordinates* (a particular case of *ilr* transformation). Regardless of the procedure, the best classification of samples was always obtained for the season (Table 8S) with apparent errors of 2.5 (LDA, ten misclassified samples) and 1.75% (QDA, seven). Therefore, DA supports the previous observations (Fig. 2S, 3S, 1, 6S, and 2), which separated the 2005 but showed only slightly different trends between 2003 and 2004 season. However, the apparent misclassifications regarding *montanera* and sampling location were above 16 and 23%, respectively. The differences in TAGs within *montanera* lengths and anatomical regions were detected by the ANOVA (Table 6S) but were not enough to lead to good segregation by DA.

For LDA, the package *compositions*, using standard *ilr* transformation (van den Boogaart & Tolosano-Delgado, 2013), led to similar discrimination of samples according to season (Fig. 3) than *robCompositions*. Samples from 2005 were differentiated from the rest, while 2003 and 204 seasons overlapped to some extension. The back transformation (into original units) of the mean *ilr* values deduced for the classification informs about the origin of the differences or similarities (Table 10S). The cause of the segregation of season 2005, as previously observed, was based on the high proportions of POO (mainly), PLO, OOO, OOL, and OLL, as well as the lower levels of the remaining TAGs. These data contrast with the almost similar values of TAGs assigned to *montanera* lengths and anatomical sampling regions by the DA (Table 10S).

Furthermore, using IBM SPSS 22 for a standard LDA with bootstrapping, there was also a 97.5% correct classification (97.3% for validation), 83.8% (81.5%), and 74.0% (71.8%) for season, *montanera* length, and anatomical region, respectively, regardless of the criterion used for the *ilr* estimation (decreasing order of *clr* variance or clustering). These outputs agree with the CoDa property that results from multivariate analysis using *ilr coordinates* do not depend on the basis chosen for the transformation and confirms the results obtained using the R packages *compositions* and *robCompositions*.

3.6.2. Fisher discriminant rule

This technique uses the idea of searching for projection/directions, which allows for maximum separation. It then allows investigating the grouping structure in the projections. Classification is obtained through the computation of the *Fisher discriminant scores,* including robust estimations. Its application to the TAG data, according to the season, also led to the correct classification of 97.5% of the samples, with a plot (Fig. 11S) quite similar to that obtained previously using Bayes' rule and the package *compositions* (Fig. 3). It presents a complete separation from 2005 samples but shows confusion between several samples from 2004 and 2003 (the ten misclassified samples are distinguished by filled green dots). For *montanera* length and anatomical region, the proportion of success was 83.75% and 94.75%, respectively. The robust option (Fig. 12S) led to a similar correct classification 97.5%, 68.75%, and 82.25%, respectively. Notice that the Fisher procedure increased the number of successful assignations but its robust alternative reduced them.

These classification results, mainly those following the Bayes' rule, show that *montanera* length and anatomical region have somewhat different TAG profiles, but discrimination may have error percentages close to 20%. In part, these results agree with those from Narvaez-Rivas et al. (2009), who only found significant differences between ram and rump for some TAG (PPS, PLPo+MLO, PLO, PLL+PoLo, SOS, SOL, and OLL). Narvaez-Rivas et al. (2009) also reported changes due to the agronomical season characteristics. According to Díaz et al. (1996), OLL/SOO, and OOO/SOO ratios showed apparent differences between four feeding systems, but their discrimination power was not checked. Viera-Alcaide et al. (2007) also reported the authentication of the fattening diet of Iberian pig according to their TAG profile from subcutaneous fat, based on

OOO and PSO as the components with the best segregation power, recommending the PSO vs OOO plot as a simple option for distinguishing *montanera* and *cebo* feedings. According to Tejeda et al. (2002), a fattening diet largely affected muscle lipid composition with total lipids higher in *montanera* than in those using commercial feeds, the first containing more oleic acid and less stearic and palmitic acids and, accordingly, less PSO and more POL, POO, and OOO.

Viera-Alcaide et al. (2008) also found TAG differences between genotypes, *montanera* lengths, types of fattening, or anatomical regions, with comparable successful assignations (63%, for montanera; 99.5%, cebo; and 54%, recebo) to those observed in this work for fattening diets. Recently, it was observed a significant effect of local and commercial breed on the pork fatty acid profiles but not of sex, although the effect was exclusively related to fatty acids. Similarly, Pena, Noguera, García-Santana, González, Tejeda, Ros-Freixedes, & Ibáñez-Escriche (2019) reported five geneomic regions in the pork with major impact of fat composition in Iberian pigs. Besides, the fatty acid profile allowed authentication of the three categories of the Iberian ham according to the rearing systems (*Jamón de Bellota*, *Jamón de Cebo de Campo*, and *Jamón de Cebo)*, with oleic acid being the most important variable driving the differentiation (González-Domínguez, Sayago, & Fernández-Recamales, 2020).

Conclusions

The work studies the effect of season, *montanera* length, and anatomical region on the Iberian pig fat's TAGs. The profiles were considered compositional data and subjected to specif ~~statistical~~ strategies developed for this statistic. The results from exploratory compositional tools, together with CoDa ANOVA, showed that season was the most influential factor on the pig fat TAG composition since it accounted for most of the variance. On the contrary, *montanera* length, or anatomical region, had a scarce influence because of their low contribution to variability. A reliable classification according to the three factors was achieved only for seasons (97.5% correct assignations) but reached lower proportions of successful assignations for *montanera* and anatomical region (80% and 65%, respectively). Hence, tracing back pig products to feeding type or anatomical region, based on TAG profiles, could be unreliable. The commercial transcendence for clarifying this aspect suggests the convenience of further studies in this field, preferably using appropriate statistical tools.

References

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*, Chapman & Hall, reprinted in 2003 with additional material. New Jersey (USA): The Blackburn Press,.

Aitchison, J., & Greenacre, M. (2002). Biplots for compositional data. Journal of the Royal Statistical Society, Series C (Applied Statistics), 51, 375-392. https://doi.org/10.1111/1467-9876.00275

Bacon-Shone, J. (2011). A short history of compositional data analysis. In, V. Pawlowsky-Glahn, & A. Buccianti (Eds). Compositional Data Analysis. Theory and application (pp 3.12). Chichester (United Kingdom): John Whiley & Sons Ltd.

Comas-Cufí, M., & Thió-Henestrosa S., 2011. CoDaPack 2.0: a stand-alone, multi-platform compositional software. In Egozcue JJ, Tolosana-Delgado R, Ortego MI (Eds), Proceedings of the 4th International Workshop on Compositional Data Analysis (CoDaWork'11) (pp 1-10). Oral presentation, Sant Feliu de Guíxols, Lleida (Spain).

Daza, A., Olivares, A., Latorre, M.A., Rey, A.I., Callejo, A., & López-Bote, C.J. (2017) Fatty acid composition of different adipose tissues in heavy pigs. Italian Journal of Food Science, 29,657-666.

Díaz, I., García Regueiro, J. A., Casillas, M., & De Pedro, E. (1996). Triglyceride composition of fresh ham fat from Iberian pigs produced with different systems of animal nutrition. Food Chemistry, 55, 383-387. https://doi.org/10.1016/0308-8146(95)00140-9

Egozcue, J.J., & Pawlowky-Glahn, V. (2011). Exploring compositional data with the CoDa-dendrogram. Australian Journal of Statistics, 40, 103-113.

Egozcue, J.J., Pawlowsky-Glahn, V., Mateo-Figueras, G., & Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. Mathematical Geology, 35, 279-300. https://doi.org/10.1023/A:1023818214614

Gallardo, E., Narváez-Rivas, M., Pablos, F., Jurado, J. M., & León-Camacho, M. (2012). Subcutaneous fat triacylglycerols profile from Iberian pigs as a tool to differentiate between intensive and extensive fattening systems. Journal of Agricultural and Food Chemistry, 60, 1645-1651. https://doi.org/10.1021/jf2045312

González-Domínguez, R., Sayago, A., Fernández Recamales, A. (2020). Fatty acid profiling for the authentication of Iberian hams according to feeding regime. Foods,9,149. https://doi.org/10.3390/foods9020149.

Kasprzyk, A., Tyra, M., and Babicz, M. (2015) Fatty acid profile of pork from a local and a commercial breed. Archives Animal Breeding, 58, 379-385. https://doi.org/10.5194/aab-58-379-2015.

Narváez-Rivas, M., León-Camacho, M., & Vicario, I. M. (2009). Fatty acid and triacylglycerol composition of the subcutaneous fat from Iberian pigs fattened on the traditional feed. Grasas y Aceites, 60, 238-247. https://doi.org/10.3989/gya.130308

Narváez-Rivas, M., Vicario, I. M., Graciani Constante, E., & León-Camacho, M. (2008). Changes in the fatty acid and triacylglycerol profiles in the subcutaneous fat of Iberian ham during the dry-curing process.Jornal of Agriculture and Food Chemistry, 56, 7131–7137. https://doi.org/10.1021/jf800990u

Pawlowski-Glahn, V., Egozcue, J.J., & Tolosana-Delgado, R. (2011). Lecture notes on Compositional data analysis. Girona (Spain): Department of Computer Science and Applied Mathematics, University of Girona.

Pawlowsky-Glahn, V., & Egozcue, JJ. (2006). Compositional data and their analysis: An introduction, Geographical Society,London (UK), Special publication, 264, 1-10. https://doi.org/10.1144/GSL.SP.2006.264.01.01

Pena, R. N., Noguera, J. L., García-Santana, M. J., González, E., Tejeda, J. F., Ros-Freixedes, R., & Ibáñez-Escriche, N. (2019) Five genomic regions have a major impact on fat composition in Iberian pigs. Scientific Reports, 9:2031. https://doi.org/10.1038/s41598-019-38622-7

Perona, J. S. & Ruíz-Gutierrez, V. (2005). Quantitative lipid composition of Iberian pig muscle and adipose tissue by HPLC. Journal Liquid Chromatography and Related Technologies, 28, 2445–2457. ~~https://doi.org/10.1080/10826070500187707~~https://doi.org/10.1080/10826070500187707

R Core Team (2020). R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. https://www.R-project.org/.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio. Boston (MA): PBC. http://www.rstudio.com/.

Ros-Freixedes, R., & Estany, J. 2014. On the compositional analysis of fatty acids in pork. Journal of Agricultural, Biological, and Environmental Statistics, 19, 136-155. https://doi.org/10.1007/s13253-0.13-0162-x

Segura, J., Cambero, M. I., Cámara, L., Loriente, C., Mateos, G. G., & López-Bote, C. J. (2015) Effect of sex, dietary glycerol or dietary fat during late fattening, on fatty acid composition and positional distribution of fatty acids within the triglyceride in pigs. Animals, 9:11, 1904-1911. https://doi.org/10.1017/S1751731115001639

Tejeda, J. F., Gandemer, G., Antequera, T., Viau, M., & García, C. (2002). Lipid traits of muscles as related to genotype and fattening diet in Iberian pigs: total intramuscular lipids and triacylglycerols. Meat Science, 60, 357-363. https://doi.org/10.1016/S0309-1740(01)00143-7

Templ, M., Hron, K., & Filzmoser, P. (2011). robCompositions: An R-package for Robust Statistical Analysis of Compositional Data. In V. Pawlowsky-Glahn & A. Buccianti (Edts), Compositional Data Analysis: Theory and Application (pp 341-354). London (U.K.): Wiley & Sons.

Van den Boogaart, K.G., & Tolosana-Delgado, R. (2013). Analysing compositional data with R. Berlin-Heidelberg (Germany): Springer-Verlag.

Ventanas S, Ventanas J, Jurado A, & Estevez M. (2006). Quality traits in muscle biceps femoris and back-fat from purebred Iberian and reciprocal Iberian x Duroc crossbred pigs. Meat Science, 73, 651-659. https://doi.org/10.1016/j.meatsci.2006.03.009

Viera-Alcaide, I., Vicario, I. M., Escudero-Gilete, M. L., Graciani Constante, E., & León-Camacho, M. (2008). A multivariate study of the triacylglycerols composition of the subcutaneous adipose tissue of Iberian pig in relation to the fattening diet and genotype. Grasas y Aceites, 59, 327-336. https://doi.org/10.3989/gya.2008.v59.i4.526

Viera-Alcaide, I., Vicario, I. M., Graciani Constante, E., & León-Camacho, M. (2007). Authentication of fattening diet of Iberian pig according to their triacylglycerols profile from subcutaneous fat. Analytica Chimica Acta, 596, 319–324. https://doi.org/ 10.1016/j.aca.2007.06.026

Figure captions

Figure 1. Compositional biplot according to season, based on the triacylglycerol profile from Iberian pig fat samples. A) covariance biplot; B) form biplot. Both plots segregated samples from the 2005 season whereas showed overlapping between 2003 and 2004

Figure 2. CoDa dendrogram obtained by SBP, choosing the triacylglycerols in the numerator, one by one, according to clustering sequence derived from Fig. 7S. It shows that the first three balances

account for most of the variance (vertical lines). Notice the displacement of the 2005 season box in the first balance, but without reaching significant difference (overlapping limits) with respect to those from 2003 and 2004.

Figure 3. LDA segregation of Iberian pig fat samples according to seasons. 3, 4, and 5 stand for season 2003, 2004 and 2005, respectively. The technique segregated samples from the 2005 season but only discriminated partially between 2003 and 2004.

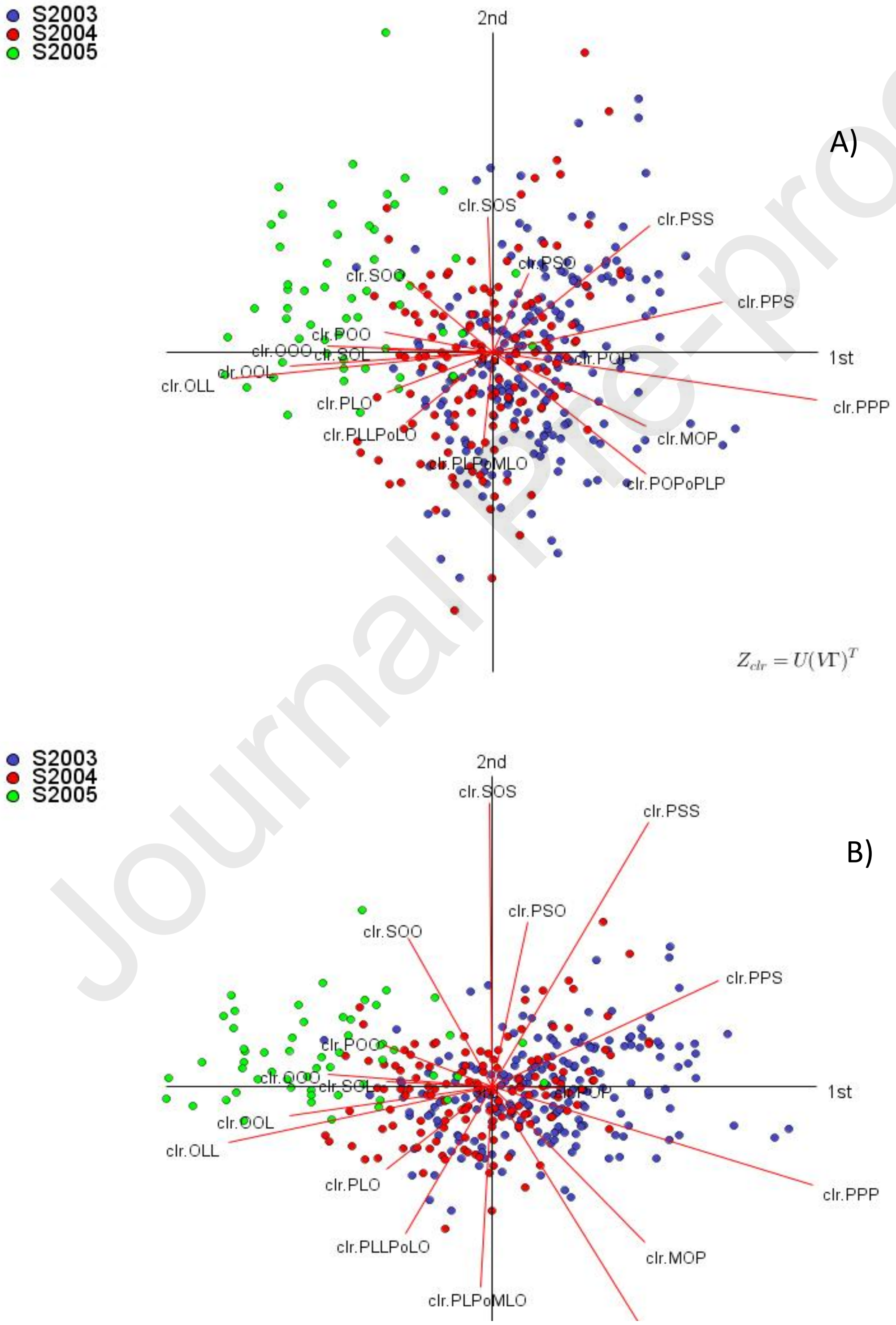

Figure 1

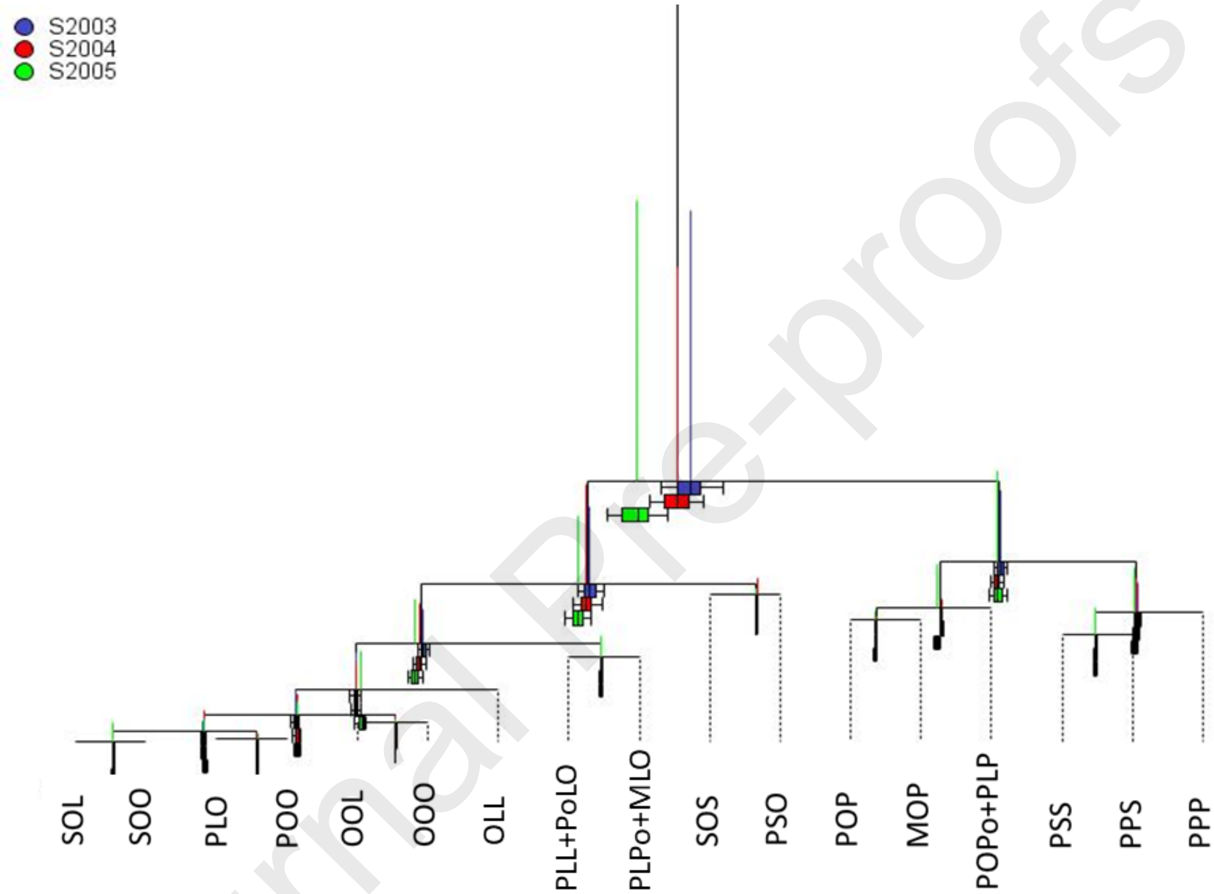

Figure 2

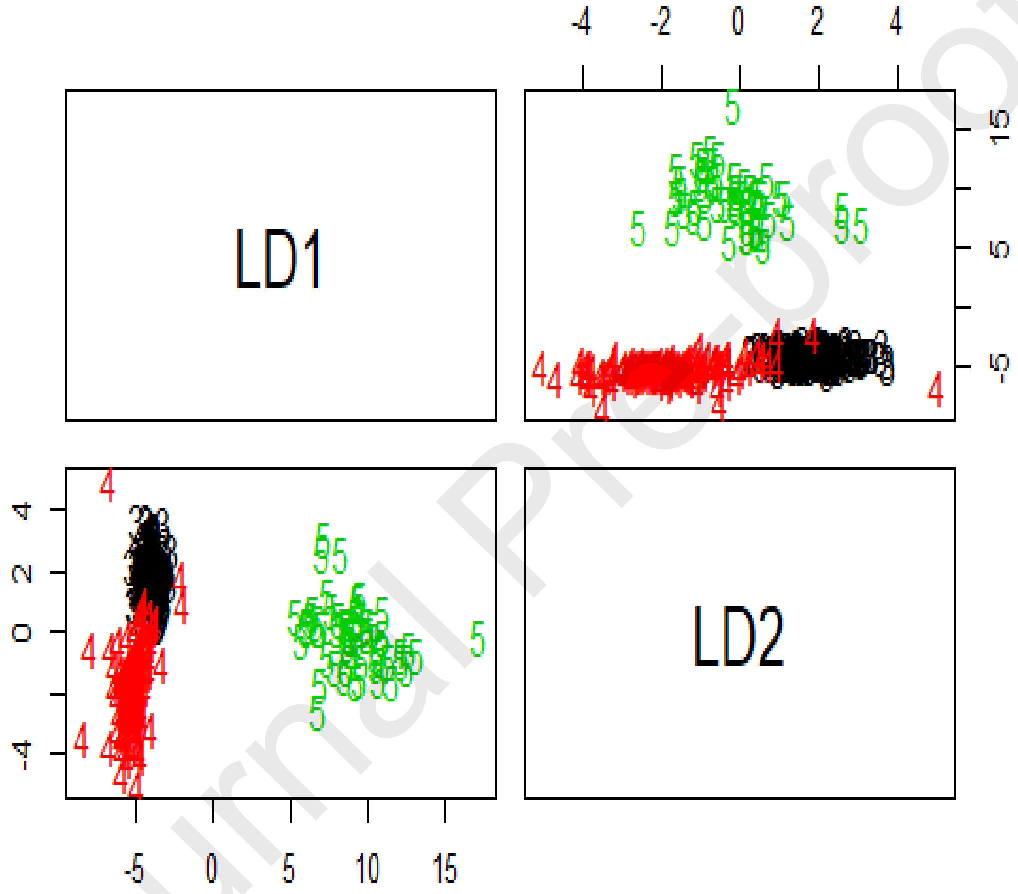

Figure 3

Table 1. Triacylglycerol species found in the Iberian pig fat samples analysed by the GC. Identification was based on the relative retention time with respect to trimiristin. See figure 1S.

Peak	Carbon number	Triacylglycerol	RRT
1	**48**	**PPP**	2.04
2		**MOP**	2.10
3	**50**	**PPS**	2.49
4		**POP**	5.54
5		**POPo+PLP**	2.62
6		**PLPo+MLO**	2.75
7	**52**	**PSS**	2.99
8		**PSO**	3.02
9		**POO**	3.08
10		**PLO**	3.24
11		**PLL+PoLO**	3.44
12	**54**	**SOS**	3.66
13		**SOO**	3.71
14		**OOO**	3.78
15		**SOL**	3.85
16		**OOL**	3.95
17		**OLL**	4.14

RRT = stands for relative retention time; the following symbols stand for their respective fatty acids: M= Myristic, P = Palmitic, Po = Palmitoleic, S = stearic, O = Oleic, L= Linoleic.

Table 2. Average geometric mean contents of the triacylglycerol compounds found in Iberian pig fat, according to season, *montanera* length, and anatomical region.

	Overall	Season			*Montanera* length		Anatomical region		Overall
Triacylglycerol	center	2003	2004	2005	M1	M2	Rump	Ham	mean
PPP	0.18	0.23	0.16	0.09	0.19	0.17	0.18	0.18	0.19
MOP	0.70	0.80	0.69	0.45	0.75	0.69	0.71	0.70	0.72
PPS	0.61	0.72	0.58	0.38	0.65	0.60	0.59	0.64	0.64
POP	5.00	5.36	4.80	4.13	5.13	4.96	4.95	5.04	5.02
POPo+PLP	4.00	4.55	4.50	1.82	4.43	3.88	4.08	3.92	4.16
PLPo+MLO	0.74	0.80	0.71	0.64	0.71	0.76	0.77	0.72	0.75
PSS	0.88	0.97	0.88	0.60	0.93	0.86	0.85	0.90	0.91
PSO	13.64	14.35	13.51	11.28	14.05	13.52	13.17	14.12	13.66
POO	32.57	31.83	31.70	36.90	32.10	32.70	32.56	32.57	32.32
PLO	10.15	9.70	10.63	10.44	9.96	10.20	10.34	9.95	10.09
PLL+PoLO	1.32	1.36	1.29	1.23	1.22	1.35	1.38	1.26	1.32
SOS	1.31	1.38	1.27	1.15	1.35	1.30	1.26	1.37	1.33
SOO	6.94	6.97	6.91	6.73	7.06	6.91	6.90	6.98	6.91
OOO	11.33	10.95	11.46	12.17	11.33	11.33	11.43	11.23	11.32
SOL	3.75	3.69	3.78	3.80	3.62	3.79	3.84	3.66	3.73
OOL	5.81	5.37	6.11	6.75	5.58	5.88	5.90	5.73	5.84
OLL	1.05	0.97	1.02	1.44	0.93	1.09	1.08	1.02	1.07

Table 3. Predicted triacylglycerol contents by the linear model (model =lm(ilr(Y)~X1+X2+X3+X1·X2+X1·X3) (values in percentages), according to the factor combination levels (treatments).

s	Triacylglycerols														
X_3	PPP	MOP	PPS	POP	POPo +PLP	PLPo +MLO	PSS	PSO	POO	PLO	PLL+PoLO	SOS	SOO	OOO	
Ham	0.264	0.851	0.839	5.730	4.506	0.733	1.105	15.372	31.810	9.159	1.131	1.434	6.956	10.843	3
ump	0.261	0.874	0.814	5.719	4.651	0.788	1.054	14.447	31.580	9.368	1.270	1.400	7.009	11.060	3
Ham	0.264	0.851	0.839	5.730	4.506	0.733	1.105	15.372	31.810	9.159	1.131	1.434	6.956	10.843	3
ump	0.222	0.800	0.697	5.294	4.610	0.830	0.927	13.807	31.707	9.878	1.472	1.356	6.992	11.050	3
Ham	0.164	0.704	0.601	4.903	4.441	0.660	0.892	14.011	31.943	10.036	1.169	1.401	7.218	11.434	3
ump	0.168	0.711	0.558	4.821	4.549	0.712	0.856	13.438	32.113	10.545	1.284	1.257	7.009	11.537	3
Ham	0.152	0.679	0.593	4.794	4.455	0.696	0.906	13.620	31.349	10.640	1.273	1.298	6.853	11.393	3
ump	0.155	0.685	0.550	4.710	4.560	0.749	0.869	13.052	31.489	11.170	1.397	1.164	6.650	11.486	3
Ham	0.101	0.497	0.461	4.803	1.522	0.604	0.641	12.047	38.275	10.180	1.083	1.152	6.494	11.689	3
ump	0.098	0.490	0.345	4.412	1.698	0.673	0.560	10.198	38.791	10.937	1.082	0.924	6.223	12.051	3
Ham	0.090	0.448	0.438	4.290	1.734	0.611	0.639	12.263	36.573	10.060	1.235	1.288	6.887	11.994	3
ump	0.086	0.441	0.328	3.937	1.933	0.679	0.557	10.371	37.032	10.799	1.232	1.032	6.594	12.355	4