Using multivariate analysis to explore the relationships between color, composition, hygienic quality and coagulation of milk from Manchega sheep

Journal:	*Journal of Dairy Science*
Manuscript ID	JDS.2019-17201.R1
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Figueroa, Antonio; Universidad de Córdoba, Producción Animal Caballero-Villalobos, Javier; Universidad de Córdoba, Producción Animal Angón, Elena; University of Córdoba, Animal Production Arias, Ramón; CERSYRA, Consejería de Agricultura y Desarrollo Rural Garzón, Ana; Universidad de Córdoba, Producción Animal Perea, José; Universidad de Córdoba, Producción Animal
Key Words:	sheep, multivariate analysis, colorimetry, milk coagulation

Using multivariate analysis to explore the relationships between color, composition, hygienic quality and coagulation of milk from Manchega sheep.

Figueroa, A.[*], Caballero-Villalobos, J.[*1], Angón, E.[*], Arias, R.[†], Garzón, A[*] and Perea, J.M.[*]

* Departamento de Producción Animal, Universidad de Córdoba, Córdoba 14071, Spain.

† Centro Regional de Selección y Reproducción Animal de Castilla-La Mancha, Valdepeñas, Ciudad Real 13300, Spain.

ABSTRACT

This study explores the relationships between composition, hygienic quality and color values of milk, in order to evaluate if colorimetry could be used as a valid predictor of the quality of raw milk and its coagulation process. For this purpose, analyses were performed in 1,200 individual Manchega sheep milk samples from 4 flocks from the region of Castilla-La Mancha (Spain).

Based on the measured variables, differences and similarities between coagulating and non-coagulating samples were determined by performing discriminant analysis techniques. The variables with a higher discriminant ability were lactose content, somatic cell score (SCS), pH and the color values L* (lightness) and a* (red/green value). The model based on color values shows a predictive ability similar to that found in the model based on milk composition and hygienic quality variables.

Canonical correlation analysis allowed to explore the relationships among both sets of variables. Canonical correlations for the first and second pair of canonical values were

[1] Corresponding author: v22cavij@uco.es

0.794 and 0.438, respectively. Both values were significant and represented 92.82% of the observed variability. The correlation structure evidenced that color values have a strong correlation with fat and protein content and total solids, and a low correlation with lactose content and SCS.

The two first combinations of standardized canonical variability could be considered as a predictable measure of composition and, to a lesser extent, of hygiene of milk. Thus, measurement of color values could be a rapid and effective tool to supplement standard analyses, in order to determine coagulation ability of Manchega sheep milk.

Key words: *sheep, multivariate analysis, colorimetry, milk coagulation.*

INTRODUCTION

Perception of color by the human eye is not exempt of subjectivity. For this reason, the International Commission on Illumination **(CIE)** has developed Color Space Systems that can provide reliable information through the measurement of coordinates that can objectively represent human visual colors. For instance, the CIELAB color space (ISO/CIE 11664-4, 2019) expresses color as 3 values: **L*** (lightness), **a*** (red/green value), **b*** (blue/yellow value). In addition, CIELCh (the cylindrical representation of the CIELAB system) combines lightness with 2 other additional values to represent color: **C*** (chroma or saturation) and **h*** (hue). Traditionally, instrumental measurement of color using the CIELAB color space has been used to characterize composition and quality of food products, and can be also used as an indicator of consumers' preferences (Ramírez-Navas, 2010).

Regarding milk, most of the studies focused on colorimetry have aimed on both the characterization of different types of cheeses during ripening (Álvarez et al., 2007; Ávila et al., 2017; Vargas-Uscategui et al., 2017) and the relationship between color and

fat content (Winkelman et al., 1999; McDermott et al., 2016) or hygienic quality (Espada et al., 2002) of bovine milk. Nevertheless, colorimetry studies in relation to ewe milk are scarce and very few literature references can be found. To this end we can highlight the work of Jiménez Sobrino et al. (2018), who analyzed the raw material and the influence of different animal management practices on composition and color of bulk tank milk. However, this study does not explore in depth the existing relationship between milk quality and color.

The importance of sheep milk production in Spain is unquestionable. In 2017, 510 millions of liters were produced, and a 31% of this production corresponded to the breed Manchega (Oviespaña, 2017). Milk from this breed is intended almost entirely for elaboration of **PDO** (Protected Designation of Origin) Manchego Cheese (Arias et al., 2016). This production is sheltered by the current regulations of the European Union, where marketing and exportation of a wide range of dairy sheep products is acknowledged with quality marks such as PDOs and **PGIs** (Protected Geographical Indications) (Cipolat-Gotet et al., 2016). In order to guarantee these high standards for ewe milk, several factors such as milk sanitary status, physicochemical quality and technological performance, must be considered. However, the current methods to analyze these parameters are not only slow and expensive, but also unable to determine *in situ* milk coagulation capacity.

For this reason, we consider a priority objective to develop a rapid, affordable and effective tool for predicting milk coagulation. In this sense, we believe the measurement of color values in milk could become a useful tool to quickly evaluate the technological capacity of milk from individual ewes, providing also valuable information about its composition and sanitary status (Ramírez-Navas, 2010).

Other previous references in literature have been found to support this hypothesis. For instance, (García-Pérez et al., 2005) established that lightness (L*) of milk mainly depends on the presence of colloidal particles (fat globules and casein micelles). On the other hand, Espada et al. (2002) reported that mastitis caused by *Streptococcus esculin* seemed to cause a yellowish-red coloring of milk, and that other changes in color were also evident in cases of *Streptococcus dysgalactiae* udder infections. In addition, milk color values seem to be affected by other factors such as herd nutrition. In particular, Faulkner et al. (2018) found that, in grazing herds with a diet based on grass or grass/clover, the higher contents of ß-carotene induced changes in perception of milk color, evidencing higher values of b*.

Considering all these facts, we expect color values could be used as quality benchmarks for milk and other dairy products. Thus, this study aims to analyze the relationship between composition, hygienic status and colorimetry of milk, in order to evaluate the possibility of developing predictive models based on color space systems to effectively estimate the quality of milk and its coagulation process in Manchega sheep.

MATERIAL AND METHODS

Data Set and Collection of Milk Samples

This study includes the analysis of 1200 individual milk samples from Manchega ewes from 4 farms located in the region of Castilla-La Mancha, Spain (50 animals/flock). Milk was collected monthly during two different seasons (3 time points in Spring and other 3 in Autumn). Samples were collected during the morning milking and were stored at 4ºC in hermetically sealed containers until analysis.

Laboratory Analysis

All analyses were performed in the Dairy Small Ruminant Laboratory (Departamento de Producción Animal, Universidad de Córdoba, Spain), within 5 hours from sample collection. Native pH of milk was measured with a Crisson Basic20 pH meter (Crisson Instruments S.A., Barcelona, Spain) and milk composition (fat, protein, lactose and total solids) was determined by mid-infrared spectroscopy using a MilkoScan FT120 (Foss Electric, Hillerød, Denmark). Technological traits were monitored at 32ºC using a Formagraph viscometer (Foss Electric, Hillerød, Denmark), following the method first developed by McMahon and Brown (1982), with the modifications described by Caballero-Villalobos et al. (2018). Briefly, this method is based on the oscillatory movement of pendula immersed in milk during coagulation. Coagulation was initiated by adding 50 µL of a 4% single-strength liquid animal rennet solution to 10 mL of milk. The test was set up at 60 min and information was transferred to a computer and represented in a coagulation diagram including the following parameters: rennet clotting time **(RCT)**, curd firming time **(k_{20})** and curd firmness at 30 and 60 minutes **(A_{30}** and **A_{60})**.

Somatic cell count **(SCC)** was measured using a Fossomatic FC (Foss Electric, Hillerød, Denmark). Subsequently, a logarithmic transformation was applied to normalize the distribution (Ali and Shook, 1980), and the variable was expressed as somatic cell score **(SCS)**.

Color values of milk were measured using the CIELAB color space and its cylindrical representation (CIELCh). Variables included lightness (L*), red/green value (a*) and blue/yellow value (b*), which were measured using a PCE-CSM2 Color Meter (PCE Instruments Ltd., Southampton, UK) by placing the lens directly over a capsule containing the milk sample. Chroma (C*) and hue (h*) were obtained from the

following mathematical formulae: $C^* = (a^{*2} + b^{*2})^{1/2}$ and $h^* = \tan^{-1}(b^*/a^*)$, according to Daszkiewicz et al. (2012), and were also included as color variables.

Statistical analysis

Preliminary testing of data was carried out to determine outliers to be discarded before further analysis. Pearson correlations were also analyzed in order to avoid variables presenting a correlation coefficient with an absolute value higher than 0.9. Because data had different measurement units, they were standardized to zero mean and a unit standard deviation. The common descriptive characteristics of the studied variables are shown in Table 1.

Subsequently, multivariate analysis techniques were used to: (1) analyze the differences and similarities in color values and milk composition and hygienic quality due to the coagulation process; (2) evaluate the specific relationships between composition, hygienic quality and color values of milk. In order to discriminate between the two groups (coagulating / non-coagulating) three complementary techniques were applied: canonical discriminant analysis (CDA), stepwise discriminant analysis (SDA) and discriminant analysis (DA).

CDA is a dimension-reduction technique related to principal component analysis and canonical correlation. Given a classification character and several variables, CDA derives a set of new variables, called canonical functions (CAN), which are linear combinations of the original variables which summarizes between-groups variation within the data, highlighting their differences.

The minimum number of variables able to discriminate between the 2 groups was obtained performing SDA, which was performed on 3 sets of variables: those related to milk composition and hygienic quality, those related to color, and the whole set of

variables. The efficiency of the discriminant power of a given model was determined by the test of significance of Wilks' Lambda. The effective separation of groups was assessed using the Mahalanobis distance and the corresponding Hotelling's T-squared test (De Maesschalck et al., 2000; Mardia, 1975).

The most discriminant variables obtained in SDA were selected and used for CDA and DA. The predictive ability of each model was tested using the absolute assignment of samples to the preassigned group (Mardia et al., 2000).

The second step was to study the existing relationships between colorimetry and quality traits of milk. Canonical correlation analysis (CCA) was deemed appropriate because it provides not only the magnitude of the relationships that may exist between groups of variables but also a quantification of the relative contribution of each variable to the relationships (Tabachnick and Fidell, 1996). Moreover, CCA complements the discriminant analysis because the latter explores only associations between data without explaining why they exist (Caballero-Villalobos et al., 2018a).

CCA is a multivariate analysis method based on the linear relationship between two multidimensional variables, X and Y. The aim is to find linear combinations $U = a^T X$ and $V = a^T Y$ so that correlation between U and V is maximized. Such linear combinations reflect the relationship between both sets of variables (Yin, 2004). The basic principle of CCA is the construction of subsequent pairs of canonical variables (U_i, V_i), that are linear combinations of the originals, so that each pair is orthogonal to the previous and represents the best explanation of the Y set (formed by q dependent variables) respect to the X set (formed by p independent variables), that has not been obtained by the previous pairs (Liu et al., 2009).

All statistical analyses were performed using SAS version 5 for Windows (SAS Institute Inc., Cary, NC).

RESULTS AND DISCUSSION

Table 2 shows results obtained by CDA with all the measured variables for milk composition and hygienic quality, colorimetry and the whole set of variables. The variables selected by SDA that better discriminate between coagulation/non-coagulation are highlighted in bold.

For milk composition and hygienic quality, the most discriminant variables were lactose, SCS and pH. Fat and protein contents were also significant, although they showed lower discriminant power. For milk color values, the most discriminant variables were L*, a* and b*. Finally, considering together both sets of variables, those with a greater discriminant ability were lactose, SCS, pH and L*.

Non-coagulating milk samples were characterized by a higher pH, lower lactose content, higher fat and protein yields and higher SCS. These results agree with those obtained in milk from Sarda sheep by (Pazzola et al., 2018), which reported lactose, SCS and pH to affect simultaneously and also separately both milk composition and milk coagulation properties.

CDA was applied to the selected variables for each of the three sets (Table 3). In all cases, the extracted CAN significantly discriminated the two groups (coagulating/non-coagulating) (P-value from Hotelling's t-test < 0.0001). The F-statistics revealed a higher discriminating ability for those variables related to milk composition and hygienic quality. This can as well be noted in Figure 1 which shows the values for the Mahalanobis distance between both coagulation groups (1.116 for colorimetric

188 variables, 2.657 for composition variables and 2.998 for both sets). All pairwise
189 distances were significant.

190 DA for the whole set of variables was able to correctly classify 83.7% of the samples
191 into their original groups (Table 4). The model constructed with composition and
192 hygienic quality variables classified correctly 82.4% of the samples, while the model
193 based on color variables classified correctly 76.3% of them. The classification error for
194 positive predictions is below 2.3% in all 3 models. For negative predictions, the
195 classification error varies between 81.9% obtained by the model based on composition
196 and hygienic quality, and 87.1% obtained by the model based on color values.

197 Thus, the model based on color values shows a predictive ability similar to that found in
198 the model based on milk composition and hygienic quality variables. These results
199 indicate that the variation pattern of color variables reflects many of the changes in
200 composition and hygienic quality of milk causing coagulation defects. Therefore, color
201 values could be used as a potential indicator of deterioration of the coagulation process.

202 Results obtained from CCA are presented in Table 5. The model extracted 84.15% of
203 the variance from the set of composition and hygienic quality variables, and 100.0% for
204 the set of color variables. Canonical correlations for the first and second pair of
205 canonical variables were 0.794 and 0.438, respectively. These values were significant
206 and represented 92.82% of the variability observed in the data.

207 The correlation structure in Figure 2 evidenced that color values are strongly correlated
208 with fat and protein content and total solids, and weakly correlated with SCS and
209 lactose content. No strong correlations with pH was observed. The first pair of
210 canonical variables links color variables and milk composition and shows a negative
211 relation between lactose content and fat and protein yields. The second pair of canonical

components links color values to SCS and lactose content, evidencing a negative association between SCS and lactose. Thus, the first combination of standardized canonical variates could be considered as a predictable measure of milk composition, while the second combination could be regarded as a predictable measure of milk hygienic properties.

pH is one of the most important factors to impact on milk coagulation, mainly due to its effect on the stability of casein micelles (Pirisi et al., 2007). Caballero-Villalobos et al. (2018b) reported that pH is the parameter that most conditions the efficiency of the coagulation process in Manchega ewe milk. Other studies have reported that coagulation time increases and, overall, coagulation properties worsen with an increase of pH of milk (Bittante et al., 2017). The present study shows that non-coagulating milk samples are significantly differentiated by higher values of pH, which agrees with previous investigations. Results also indicate that variations in pH are not appropriately reflected by color variables.

Faulty coagulation of sheep and goat milk has been previously linked to higher levels of SCC (Leitner et al., 2008; Caballero-Villalobos et al., 2018a) and has been attributed to the proteolytic effect of some enzymes (Poulsen et al., 2015). Several studies have evidenced that high levels of SCC delay the coagulation process and, as a result, deteriorate its properties (Rovai et al., 2015; Vacca et al., 2015; Pazzola et al., 2018). In agreement with these authors, the present results show that non-coagulating milk samples are differentiated by their higher SCS values. Furthermore, there seems to be a negative relation between SCS and lactose content. This decrease in lactose has been proposed as an indicator of the udder health status (Vivar-Quintana et al., 2006; Manca et al., 2016).

Pazzola et al. (2018) suggested that the mechanism that explains the association between lactose and coagulation is related to the role of lactose as an osmotic regulator in milk (proved by Poulsen et al., 2015) and changes in the percentage and composition of minerals and proteins. Results from the present study are in agreement with those obtained by Pazzola et al. (2018), who reported a worsening of coagulation properties in relation to the low lactose content, and the negative association between lactose and protein contents.

According to (De Marchi et al., 2013), the determination of milk chemical composition and/or coagulation properties requires the use of different devices that are not suitable for quality control systems or for dairy plant cheese production lines due to disadvantages such as the high cost of equipment, the need for highly qualified personnel and the considerable investment of time. However, color variables can be rapidly measured with portable low-cost devices that do not require the use of any reagents or skilled staff. In other words, the model based on color parameters developed in the present study can predict faulty coagulation with the same precision as the model based on milk composition and hygienic quality, with less than 3% of the investment and with no direct costs per milk sample.

For these reasons, colorimetry could be of great interest for the dairy sector. Color variables would therefore provide a rapid and economic tool to analyze Manchega milk and, according to the results obtained by Ferragina et al. (2017) using FTIR infrared spectroscopy, it would be possible to implement this model directly on those samples collected during milk recording, in order to gather a large amount of data from each animal. In addition, as reported by Winkelman et al. (1999) in a previous study performed on dairy cattle, milk color values seem to have a high heritability (≈ 0.40), so

color could be considered on selection schemes for the improvement of major aspects of the dairy sheep industry.

CONCLUSIONS

SCS, pH and lactose content are the features that better discriminate coagulation properties of Manchega sheep. However, color values have been found to have similar predictive ability, so they could be used as a potential indicator of the coagulation ability of Manchega ewe milk.

Color values in Manchega milk are strongly correlated with fat content, protein content and total solids, and also slightly correlated with SCS and lactose concentration. Consequently, colorimetry could offer a predictable measure of milk composition, although it provides lesser information on hygienic quality.

Thus, the measurement of color values in milk could be considered a rapid and effective tool to supplement standard analyses, in order to determine coagulation ability of Manchega sheep milk.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the Spanish Association of Manchega Sheep Breeders (AGRAMA), for their implication in this study.

BIBLIOGRAPHY

Ali, A.K.A., and G.E. Shook. 1980. An Optimum Transformation for Somatic Cell Concentration in Milk. J. Dairy Sci. 63:487–490. doi:10.3168/jds.S0022-0302(80)82959-6.

Álvarez, S., V. Rodríguez, M.E. Ruiz, and M. Fresno. 2007. Correlaciones de

textura y color instrumental con la composición química de quesos de cabra canarios. Arch. Zootec. 56:663–666.

Arias, R., R. Gallego, S. Altares, A. Garzón, J. Romero, L. Jiménez, B. Oliete, C. Arias, J. Caballero, A. Martínez, N. Núñez, A. García, M. Ramón, V. Montoro, and M. Pérez-Guzmán. 2016. Quality of milk in Manchego sheep flocks. A review.. Arch. Zootec. 65:469–473.

Ávila, M., N. Gómez-Torres, D. Delgado, P. Gaya, and S. Garde. 2017. Effect of high-pressure treatments on proteolysis, volatile compounds, texture, colour, and sensory characteristics of semi-hard raw ewe milk cheese. Food Res. Int. 100:595–602. doi:10.1016/J.FOODRES.2017.07.043.

Bittante, G., C. Cipolat-Gotet, M. Pazzola, M.L. Dettori, G.M. Vacca, and A. Cecchinato. 2017. Genetic analysis of coagulation properties, curd firming modeling, milk yield, composition, and acidity in Sarda dairy sheep. J. Dairy Sci. 100:385–394. doi:10.3168/JDS.2016-11212.

Caballero-Villalobos, J., A. Figueroa, K. Xibrraku, E. Angón, J.M. Perea, and A. Garzón. 2018a. Multivariate analysis of the milk coagulation process in ovine breeds from Spain. J. Dairy Sci. 101:10733–10742. doi:10.3168/jds.2018-14752.

Caballero-Villalobos, J., J.M. Perea, E. Angón, R. Arias, and A. Garzón. 2018b. Coagulation efficiency and its determinant factors: A case study for

302 Manchega ewe milk in the region of Castilla-La Mancha, Spain. J. Dairy Sci.
303 101:3878–3886. doi:10.3168/jds.2017-13816.

304 Cipolat-Gotet, C., A. Cecchinato, M. Pazzola, M.L. Dettori, G. Bittante, and G.M.
305 Vacca. 2016. Potential influence of herd and animal factors on the yield of
306 cheese and recovery of components from Sarda sheep milk, as determined
307 by a laboratory bench-top model cheese-making. Int. Dairy J. 63:8–17.
308 doi:10.1016/J.IDAIRYJ.2016.07.013.

309 Daszkiewicz, T., D. Kubiak, R. Winarski, and M. Koba-Kowalczyk. 2012. The effect
310 of gender on the quality of roe deer (Capreolus capreolus L.) meat. Small
311 Rumin. Res. 103:169–175. doi:10.1016/J.SMALLRUMRES.2011.09.044.

312 Espada, E., H. Vijverberg, and N.V. Maasland. 2002. Milk color analysis as a tool
313 for the detection of abnormal milk. Pages 28–38 in North American
314 Conference on Robotic Milking, Toronto. Canada.

315 Faulkner, H., T.F. O'Callaghan, S. McAuliffe, D. Hennessy, C. Stanton, M.G.
316 O'Sullivan, J.P. Kerry, and K.N. Kilcawley. 2018. Effect of different forage
317 types on the volatile and sensory properties of bovine milk. J. Dairy Sci.
318 101:1034–1047. doi:10.3168/JDS.2017-13141.

319 Ferragina, A., C. Cipolat-Gotet, A. Cecchinato, M. Pazzola, M.L. Dettori, G.M.
320 Vacca, and G. Bittante. 2017. Prediction and repeatability of milk
321 coagulation properties and curd-firming modeling parameters of ovine milk

using Fourier-transform infrared spectroscopy and Bayesian models. J. Dairy Sci. 100:3526–3538. doi:10.3168/JDS.2016-12226.

García-Pérez, F.J., Y. Lario, J. Fernández-López, E. Sayas, J.A. Pérez-Alvarez, and E. Sendra. 2005. Effect of orange fiber addition on yogurt color during fermentation and cold storage. Color Res. Appl. 30:457–463. doi:10.1002/col.20158.

ISO/CIE 11664-4:2019. Colorimetry - Part 4: CIE 1976 L*a*b* colour space.

Jiménez Sobrino, L., J.M. Poveda Colado, A.I. Garzón Sigler, A.L. Martínez Marín, N. Núñez Sánchez, J.R. Asensio, M.D. Pérez-Guzmán Palomares, and R. Arias Sánchez. 2018. Composition and colour indices of sheep's bulk-tank milk are influenced by production practices. Ital. J. Anim. Sci. 17:477–488. doi:10.1080/1828051X.2017.1383860.

Leitner, G., N. Silanikove, and U. Merin. 2008. Estimate of milk and curd yield loss of sheep and goats with intrammamary infection and its relation to somatic cell count. Small Rumin. Res. 74:221–225. doi:10.1016/J.SMALLRUMRES.2007.02.009.

Liu, J., W. Drane, X. Liu, and T. Wu. 2009. Examination of the relationships between environmental exposures to volatile organic compounds and biochemical liver tests: Application of canonical correlation analysis. Environ. Res. 109:193–199. doi:10.1016/J.ENVRES.2008.11.002.

De Maesschalck, R., D. Jouan-Rimbaud, and D.L. Massart. 2000. The Mahalanobis distance. Chemom. Intell. Lab. Syst. 50:1–18. doi:10.1016/S0169-7439(99)00047-7.

Manca, M.G., J. Serdino, G. Gaspa, P. Urgeghe, I. Ibba, M. Contu, P. Fresi, and N.P.P. Macciotta. 2016. Derivation of multivariate indices of milk composition, coagulation properties, and individual cheese yield in dairy sheep. J. Dairy Sci. 99:4547–4557. doi:10.3168/jds.2015-10589.

De Marchi, M., V. Toffanin, M. Cassandro, and M. Penasa. 2013. Prediction of coagulating and noncoagulating milk samples using mid-infrared spectroscopy. J. Dairy Sci. 96:4707–4715. doi:10.3168/JDS.2012-6506.

Mardia, K. V, F.L. Bookstein, and I.J. Moreton. 2000. Statistical assessment of bilateral symmetry of shapes. Biometrika 87:285–300. doi:10.1093/biomet/87.2.285.

McDermott, A., G. Visentin, S. McParland, D.P. Berry, M.A. Fenelon, and M. De Marchi. 2016. Effectiveness of mid-infrared spectroscopy to predict the color of bovine milk and the relationship between milk color and traditional milk quality traits. J. Dairy Sci. 99:3267–3273. doi:10.3168/JDS.2015-10424.

McMahon, D.J., and R.J. Brown. 1982. Evaluation of Formagraph for Comparing Rennet Solutions,. J. Dairy Sci. 65:1639–1642. doi:10.3168/JDS.S0022-0302(82)82390-4.

362 Oviespaña. 2017. No Title. Accessed May 11, 2017.
363 https://www.oviespana.com/informacion-de-ovino/servicio-diario-de-
364 noticias/noticias/castilla-la-mancha-produjo-durante-el-pasado-ano-158-
365 millones-de-litros-de-leche-de-oveja-el-31-de-los-510-millones-de-litros-
366 que-se-generaron-en-espana.

367 Pazzola, M., C. Cipolat-Gotet, G. Bittante, A. Cecchinato, M.L. Dettori, and G.M.
368 Vacca. 2018. Phenotypic and genetic relationships between indicators of
369 the mammary gland health status and milk composition, coagulation, and
370 curd firming in dairy sheep. J. Dairy Sci. 101:3164–3175.
371 doi:10.3168/JDS.2017-13975.

372 Pirisi, A., A. Lauret, and J.P. Dubeuf. 2007. Basic and incentive payments for goat
373 and sheep milk in relation to quality. Small Rumin. Res. 68:167–178.
374 doi:10.1016/J.SMALLRUMRES.2006.09.009.

375 Poulsen, N.A., A.J. Buitenhuis, and L.B. Larsen. 2015. Phenotypic and genetic
376 associations of milk traits with milk coagulation properties. J. Dairy Sci.
377 98:2079–2087. doi:10.3168/JDS.2014-7944.

378 Ramírez-Navas, J.S. 2010. Espectrocolorimetría en caracterización de leche y
379 quesos. Tecnol. Láctea Latinoam. 61:52–58.

380 Rovai, M., N. Rusek, G. Caja, J. Saldo, and G. Leitner. 2015. Effect of subclinical
381 intrammamay infection on milk quality in dairy sheep: I. Fresh-soft cheese

produced from milk of uninfected and infected glands and from their blends. Small Rumin. Res. 125:127–136. doi:10.1016/J.SMALLRUMRES.2015.02.019.

Tabachnick, B.G., and L.S. Fidell. 1996. Using Multivariate Statistics. 3rd ed. Harper Collins, New York (NY).

Vacca, G.M., M. Pazzola, M.L. Dettori, E. Pira, F. Malchiodi, C. Cipolat-Gotet, A. Cecchinato, and G. Bittante. 2015. Modeling of coagulation, curd firming, and syneresis of milk from Sarda ewes. J. Dairy Sci. 98:2245–2259. doi:10.3168/JDS.2014-8902.

Vargas-Uscategui, R., A. Arenas-Clavijo, and J.S. Ramírez-Navas. 2017. Efecto del proceso de acidificación sobre el color de queso cottage. Agron. Mesoam. 28:677–690.

Vivar-Quintana, A.M., E. Beneitez De La Mano, and I. Revilla. 2006. Relationship between somatic cell counts and the properties of yoghurt made from ewes' milk. Int. Dairy J. 16:262–267. doi:10.1016/J.IDAIRYJ.2005.03.006.

Winkelman, A.M., D.L. Johnson, and A.K.H. MacGibbon. 1999. Estimation of Heritabilities and Correlations Associated with Milk Color Traits. J. Dairy Sci. 82:215–224. doi:10.3168/JDS.S0022-0302(99)75226-4.

Yin, X. 2004. Canonical correlation analysis based on information theory. J. Multivar. Anal. 91:161–176. doi:10.1016/S0047-259X(03)00129-5.

403

404 **Table 1.** Description of composition and color variables used to evaluate the coagulation
405 process of Manchega sheep milk (n = 920)

Variable	Description	Unit / Range	Mean	SD
Composition				
Fat	Fat content	%	6.61	1.94
Protein	Protein content	%	5.74	0.84
Lactose	Lactose content	%	4.88	0.40
TS	Total solids	%	18.13	2.47
SCS	Somatic cell score	$\log_{10}$ (cells/mL)	5.33	0.68
pH	pH	-log [H^+]	6.64	0.29
Colorimetry				
L*	Lightness	[0, 100]	82.86	1.81
a*	Red/Green value	[-60, +60]	-2.46	0.70
b*	Blue/Yellow value	[-60, +60]	4.79	1.79
C*	Chroma	$(a^{*2} + b^{*2})^{1/2}$	5.56	1.32
h*	Hue	$\tan^{-1}(b^*/a^*)$	-0.53	0.27

Table 2. Results obtained by canonical discriminant analysis with all variables measured for composition and hygienic quality, colorimetry and both groups of variables (variables selected by the stepwise discriminant analyses are shown in bold)

Variable	Non-coagulating samples	Coagulating samples	Wilks' Lambda	F-value	P-value	CAN[1]
Both groups						
Fat	7.25 ± 1.83	6.57 ± 1.94	0.994	5.334	0.021	-0.188
Protein	6.06 ± 0.83	5.73 ± 0.84	0.992	7.061	0.008	-0.217
Lactose	**4.51 ± 0.52**	**4.90 ± 0.39**	**0.956**	**42.355**	**< 0.0001**	**0.521**
TS	18.70 ± 2.38	18.11 ± 2.48	0.997	2.519	0.113	-0.130
SCS	**6.06 ± 0.76**	**5.28 ± 0.63**	**0.940**	**58.451**	**< 0.0001**	**-0.607**
pH	**6.92 ± 0.17**	**6.63 ± 0.29**	**0.955**	**43.327**	**< 0.0001**	**-0.526**
L*	**82.08 ± 1.96**	**82.90 ± 1.80**	**0.990**	**9.038**	**0.003**	**0.245**
a*	-2.33 ± 0.63	-2.47 ± 0.70	0.998	1.625	0.023	-0.104
b*	4.93 ± 1.78	4.78 ± 1.79	1.000	0.312	0.577	-0.046
C*	5.60 ± 1.38	5.56 ± 1.32	1.000	0.043	0.836	-0.017
h*	-0.49 ± 0.25	-0.53 ± 0.28	0.999	0.682	0.409	-0.068
Composition						
Fat	7.25 ± 1.83	6.57 ± 1.94	0.994	5.334	0.021	0.203
Protein	6.06 ± 0.83	5.73 ± 0.84	0.992	7.061	0.008	0.233
Lactose	**4.51 ± 0.52**	**4.90 ± 0.39**	**0.956**	**42.355**	**< 0.0001**	**-0.560**
TS	18.70 ± 2.38	18.11 ± 2.48	0.997	2.519	0.113	0.140
SCS	**6.06 ± 0.76**	**5.28 ± 0.63**	**0.940**	**58.451**	**< 0.0001**	**0.653**
pH	**6.92 ± 0.17**	**6.63 ± 0.29**	**0.955**	**43.327**	**< 0.0001**	**0.567**
Colorimetry						
L*	**82.08 ± 1.96**	**82.90 ± 1.80**	**0.990**	**9.038**	**0.003**	**0.431**
a*	**-2.33 ± 0.63**	**-2.47 ± 0.70**	**0.998**	**1.625**	**0.023**	**-0.184**
b*	4.93 ± 1.78	4.78 ± 1.79	1.000	0.312	0.577	-0.080
C*	5.60 ± 1.38	5.56 ± 1.32	1.000	0.043	0.836	-0.030
h*	-0.49 ± 0.25	-0.53 ± 0.28	0.999	0.682	0.409	-0.119

[1] Correlation of each variable with the canonical function.

413
414 **Table 3.** Discriminant canonical models for composition variables, colorimetry and both sets of
415 variables

Model	Variables in the model	Number of groups	Wilks' Lambda	F-value	P-value
Composition	Lactose, SCS, pH	2	0.888	$F(38.624) = 2.615$	< 0.0001
Colorimetry	L*, a*, b*	2	0.950	$F(16.216) = 2.615$	< 0.0001
Both sets	Lactose, SCS, pH, L*	2	0,875	$F(32.647) = 2.382$	< 0.0001

416

417 **Table 4.** Assignation percentages in the predefined groups and classification errors

	Non-coagulating samples	Coagulating samples
Both groups		
Non-coagulating samples	63.04	36.96
Coagulating samples	15.22	84.78
Level of error	0.82	0.02
Prior probability	0.50	0.50
Composition		
Non-coagulating samples	71.74	28.26
Coagulating samples	17.05	82.95
Level of error	0.82	0.02
Prior probability	0.50	0.50
Colorimetry		
Non-coagulating samples	65.22	34.78
Coagulating samples	23.11	76.89
Level of error	0.87	0.02
Prior probability	0.50	0.50

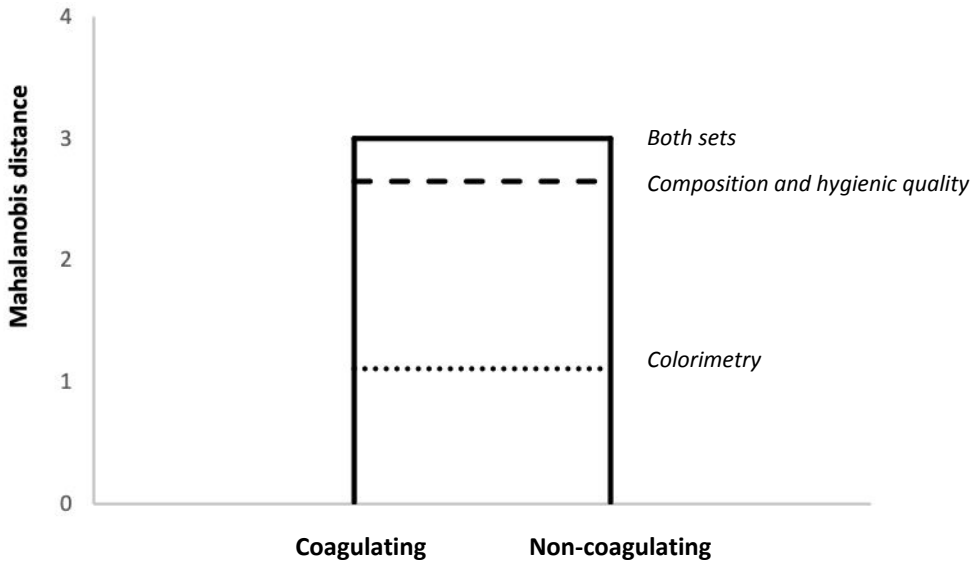
418

419 **Table 5**. Results for composition and color values obtained by canonical correlation analysis
420 with all the measured variables

Factors	Canonical correlations	Eigen value	Variability (%)	Wilks' Lambda	F-value	P-value
F_1	0.794	0.630	71.17	0.280	33.428	< 0.000
F_2	0.438	0.192	21.65	0.758	8.559	< 0.000
F_3	0.212	0.045	5.09	0.937	4.503	< 0.000

421
422
423
424
425
426

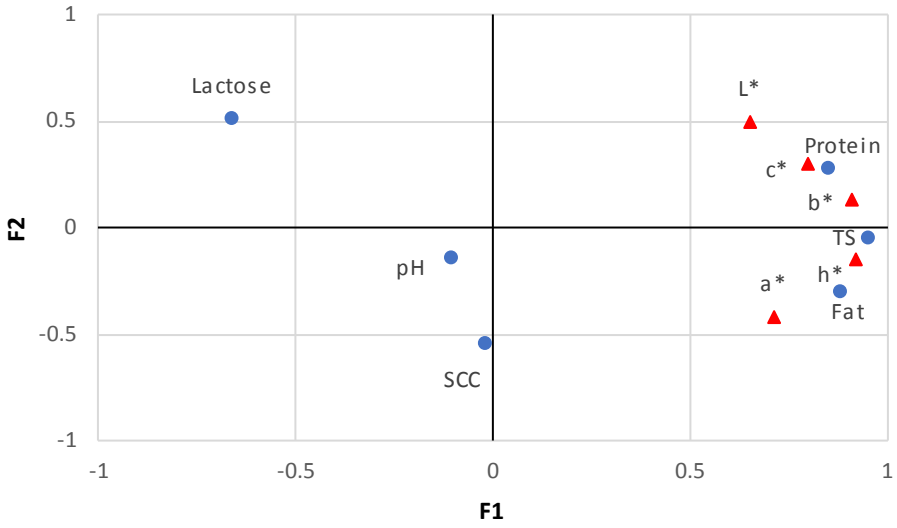
428
429
430



431 **Figure 1.** Dendrogram showing composition and colorimetry relationship between samples

432

433



434

435 **Figure 2.** Correlation structure between the first two pairs of canonical variables (F_1, F_2) and the

436 composition and hygienic quality (●) and colorimetry (▲) variables of milk