

# Inferring individual cow effects, dairy system effects and feeding effects on latent variables underlying milk protein composition and cheese-making traits in dairy cattle

Christos Dadousis, Claudio Cipolat-Gotet, Stefano Schiavon, Giovanni Bittante and Alessio Cecchinato\*

Department of Agronomy, Food, Natural resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020 Legnaro, Italy

Received 30 January 2017; accepted for publication 9 October 2017; first published online 10 November 2017

We examined the latent structure of 26 cheese related phenotypes in dairy cattle. Traits related to milk yield and quality (8 traits), milk protein fractions (8 traits), coagulation and curd firmness indicators (CF, 5 traits) and cheese-making phenotypes (cheese yields (%CY) and nutrient recoveries in the curd (REC), 5 traits) were analysed through multivariate factor analysis (MFA) using a *varimax* rotation. All phenotypes were measured in 1264 Brown Swiss cows. Ten mutual orthogonal, latent variables (factors; Fs) were obtained explaining 74% of the original variability. These Fs captured basic concepts of the cheese-making process. More precisely, the first 4 Fs, sorted by variance explained, were able to capture the underlying structure of the CY percentage (F1: %CY), the CF process with time (F2: CF<sub>t</sub>), the milk and solids yield (F3: Yield) and the presence of nitrogen (N) in the cheese (F4: Cheese N). Moreover, 4 Fs (F5:  $\alpha_1$ - $\beta$ -CN, F7:  $\kappa$ - $\beta$ -CN, F8:  $\alpha_2$ -CN and F9:  $\alpha_1$ -CN-Ph) were related to the basic milk caseins and 1 factor was associated with the  $\alpha$ -LA whey protein (F10:  $\alpha$ -LA). A factor describing udder health status (F6: Udder health), mainly loaded on lactose, other nitrogen compounds in the milk and SCS, was also obtained. Further, we inferred the effects of some potential sources of variation (e.g. stage of lactation and parity) including feeding and management systems. Stage of lactation had a significant effect for 7 of the 10 Fs, followed by parity of the cow (3 Fs), dairy system and feeding (3 Fs). Our work demonstrates the usefulness of MFA in reducing a large number of variables to a few latent factors with biological meaning and representing groups of traits that describe a complex process like cheese-making. Such an approach would be a valuable tool for studying the influence of different production environments and individual animal factors on protein composition and cheese-making related traits.

**Keywords:** Multivariate factor analysis, milk protein, coagulation, curd firmness, individual cheese yield.

Cheese produced at the individual cow level is the result of a complex process with plenty of elements involved, from milk quality characteristics (e.g. percentage of protein and fat in the milk), milk coagulation properties (MCP), curd firmness (CF) and cheese-yields, such as the quantity of cheese obtained from a given amount of milk (CY) and the loss of nutrients in whey/recovery in curd (REC). Although detailed phenotyping of individual animals is required to understand the biological and genetic background of the traits of interest, inclusion of a large number of traits into a

selection index in breeding programs poses difficulties in interpretation and computations.

To overcome the problem of data dimension, and for a better understanding of complex phenomena, multivariate factor analysis (MFA) has been investigated for a variety of traits, such as milk composition, MCP and CY traits in dairy cattle, sheep and goats (Macciotta et al. 2012; Manca et al. 2016; Vacca et al. 2016), for milk fatty acids in dairy cows (Conte et al. 2016; Mele et al. 2016) and for cheese volatile organic compounds (Bergamaschi et al. 2015).

Factor analysis belongs to the general framework of multivariate analysis. The main idea of MFA is that  $n$  observed variables,  $x$ , can be expressed as linear functions of  $p$  ( $p < n$ ) latent variables (factors; Fs): this statistical approach focuses

\*For correspondence; e-mail: [alessio.cecchinato@unipd.it](mailto:alessio.cecchinato@unipd.it)

on understanding relationships (the underlying latent concept that the measured variables represent) among a set of observed variables. Thus,  $F_s$  can be considered as variables that are not measurable, but they can be extracted from a set of measured-indicator variables by making use of their covariance structure (Bollen, 1989). In such a way, only the underlying concept of interest is kept for further analysis while at the same time data reduction is achieved.

To our knowledge, there are no studies reported in the literature that deal with MFA and milk protein profile. Moreover, milk protein fractions have never been considered in combination with milk technological traits to represent the complexity of cheese-making process. In addition, besides the statistical methodology, no studies have been reported in the literature dealing with the effects of feeding regime and management system on milk proteins and cheese-making related variables. Therefore our objectives were (i) to create a new set of latent phenotypes related to milk quality, detailed protein composition measured by the RP-HPLC method, coagulation properties, and cheese-making traits, by using MFA, and (ii) to assess some individual sources of variation (i.e., stage of lactation and parity) and herd factors as well as the effects of feeding and management systems.

## Materials and methods

### *Animals, herds and dairy farming systems*

Milk samples from 1264 Italian Brown Swiss cows reared in 85 herds located in Trento Province (Italy) were collected. A full description of the sampling procedure can be found in (Cecchinato et al. 2013). In brief, ~15 cows/herd were individually sampled once (evening milking). Samples were processed within 20 h after collection. Information on cows and herds was supplied by the Breeders Association of Trento Province.

The farming systems have been previously analysed and reported in Sturaro et al. (2009, 2013). In brief, two main categories of farming systems were distinguished: (i) traditional farms and (ii) modern dairy systems. The traditional systems represent small and old barns where feeding is heavily based on meadow hay and cows are tied. Modern farms, with loose cows and milking parlour, were further distinguished in 3 sub-categories depending on the feeding system: (a) modern dairy system but without use of total mixed ratio (TMR), (b) modern dairy farms with silage-based TMR and (c) modern dairy farms without silage using water to moisturise TMR.

### *Cheese-making phenotypes*

**Analysis of milk and milk protein fractions.** Individual milk subsamples were analysed for protein, fat, lactose and casein contents using MilkoScan FT6000 (Foss, Hillerød, Denmark). The pH of the subsamples was measured using a Crison Basic 25 electrode (Crison, Barcelona, Spain).

Somatic cell count measures were obtained by Fossomatic FC counter (Foss, Hillerød, Denmark) and were logarithmic transformed [ $SCS = \log_2(SCC/100\ 000) + 3$ ] (Ali & Shook, 1980).

Protein fractions [ $\alpha_{s1}$ -,  $\alpha_{s2}$ -,  $\beta$ - and  $\kappa$ - casein (CN),  $\beta$ -lactoglobulin ( $\beta$ -LG) and  $\alpha$ -lactalbumin ( $\alpha$ -LA)] were measured by the RP-HPLC method (Bonfatti et al. 2008) and were expressed as proportions to the total milk nitrogen (N) content. Further, the phosphorylated form of the  $\alpha_{s1}$ -CN was obtained as proposed by (Bonfatti et al. 2011). The remained N milk compounds were also included in the analysis and were calculated by subtracting the sum of the protein fractions from the total N milk content.

### *Milk coagulation properties and modelling the CF.*

Measures of milk coagulation properties were obtained using the Formagraph instrument by Foss Electric A/S according to the procedure described in Cecchinato et al. (2013). Files containing 360 CF values for each milk sample, recorded every 15 sec for 90 min, were retrieved and used to estimate a set of parameters of CF at time  $t$  ( $CF_t$ ) according to equations and methodology developed by Bittante et al. (2013b). Estimated parameters included: rennet coagulation time ( $RCT_{eq}$ , min), potential asymptotical curd firmness ( $CF_p$ , mm), representing the maximum potential curd firmness of a given sample after infinite time in the absence of syneresis, curd-firming rate constant ( $k_{CF}$ ,  $\% \times \text{min}^{-1}$ ) which measures the relative increment of  $CF_p$ , that is predominant until reaching the maximum curd firmness ( $CF_{max}$ , mm; at time  $t_{max}$ , min) and before syneresis [measured by  $k_{SR}$  ( $\% \times \text{min}^{-1}$ )] became prevalent. To avoid convergence and estimation problems,  $CF_p$  was calculated multiplying  $CF_{max}$  by 1.34, that is the coefficient resulting from the linear regression between  $CF_p$  and  $CF_{max}$  values obtained in a preliminary analysis (Stocco et al. 2016). The other three  $CF_t$  model parameters ( $RCT_{eq}$ ,  $k_{CF}$ , and  $k_{SR}$ ) were estimated by curvilinear regression using the nonlinear procedure (PROC NLIN) in the SAS software (SAS Institute Inc., Cary, NC).

### *Individual cheese yield and curd nutrient recoveries.*

A detailed description of the individual model-cheese processing can be found in Cipolat-Gotet et al. (2013). The phenotypes were obtained through a model cheese-making procedure on 1500 ml of milk for each cow. In brief, the traits analysed were: (i) three %CY traits, expressing the weight (wt) of fresh curd ( $\%CY_{CURD}$ ), of curd dry matter ( $\%CY_{SOLIDS}$ ), and of water retained in the curd ( $\%CY_{WATER}$ ) as percentage of wt of milk processed, and (ii) four REC traits representing the proportion of nutrients and energy of the milk retained in the curd ( $REC_{SOLIDS}$ ,  $REC_{FAT}$ ,  $REC_{PROTEIN}$  and  $REC_{ENERGY}$  calculated as the % ratio between the nutrient in curd and the corresponding nutrient in processed milk). The recovery energy in the curd was calculated as the difference between energy in the milk and in the cheese (NRC, 2001).

### Statistical analysis

**Multivariate factor analysis.** The approach has been described in Dadousis et al. (2017a). In brief, to avoid severe multicollinearity problems, 3 out of the 29 phenotypes ( $CF_{max}$ ,  $\%CY_{CURD}$  and  $REC_{SOLIDS}$ ) were excluded. As mentioned above,  $CF_{max}$  is proportional to  $CF_P$ . Moreover, the  $\%CY_{CURD}$  is the sum of  $\%CY_{WATER}$  and  $\%CY_{SOLIDS}$ , while  $REC_{SOLIDS}$  has phenotypic correlation with  $REC_{ENERGY}$  and  $\%CY_{SOLIDS}$  greater than 0.9. The remaining twenty-six phenotypes were simultaneously analysed in the following factor model:

$$\mathbf{x} = \mathbf{\Lambda}\boldsymbol{\xi} + \boldsymbol{\delta}, \quad (1)$$

where  $\mathbf{x}$  is a vector containing the measured phenotypes,  $\boldsymbol{\xi}$  is a vector of the factors,  $\mathbf{\Lambda}$  contains the factor loadings ( $\lambda$ ) relating the factors to the original variables and  $\boldsymbol{\delta}$  is a vector of the residuals.

At a first step, the difference between Pearson and partial correlations of the measured variables was used as a hint to assess the adequacy of the data for MFA. This difference can be viewed as a way to control if the correlation between 2 variables is mediated by other variables, with a high value indicating the existence of a latent structure. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was applied to quantify this difference (Dziuban & Shirkey, 1974; Kaiser & Rice, 1974). Further, MFA was applied. The factors were extracted based on prior knowledge, biological interpretation and the proportion of original variance explained. Moreover, factor rotation (*varimax*) was used to identify simple structure. Factor loadings  $>|0.4|$  were considered as ‘significant’ to explain the factors. The MFA was performed with the *psych* package (Revelle, 2014) in R (R Core Team, 2013).

**Mixed model analysis.** To estimate sources of variation related to the factors, the 10 factor scores were analysed with the following model:

$$Y_{ijklm} = \mu + \text{dairy system}_i + \text{herd}_j(\text{dairy system})_j + \text{parity}_k + \text{DIM}_l + e_{ijklm} \quad (2)$$

where  $Y_{ijklm}$  is the observed phenotype (i.e. the factor scores);  $\mu$  is the overall mean;  $\text{dairy system}_i$  is the fixed effect of the  $i$ th dairy system ( $i = 1$  to 4);  $\text{herd}_j(\text{dairy system})_j$  is the random effect of the  $j$ th herd ( $j = 1$  to 85)  $\sim N(0, \mathbf{I}\sigma_h^2)$  nested within the  $i$ th dairy system;  $\text{parity}_k$  is the fixed effect of the  $k$ th parity ( $k = 1$  to 4 or more lactations);  $\text{DIM}_l$  is the  $l$ th 30-d class of days in milk (11 classes); and  $e_{ijklm}$  is the residual random error term  $\sim N(0, \mathbf{I}\sigma_e^2)$ ; where  $\mathbf{I}$  is an identity matrix,  $\sigma_h^2$  and  $\sigma_e^2$  are herd/date and residual variances, respectively. The significance of the dairy system was tested on the error line of the herd within dairy system, while for parity and DIM class the error line of the residual variance was used. Orthogonal post hoc contrasts ( $P < 0.05$ ) were built for dairy system: (i) the ‘Traditional’ dairy system was compared with the ‘Modern’ systems; (ii) within the

**Table 1.** Summary statistics of milk (yield and quality), protein fractions, coagulation (curd firming) and cheese-making (%CY and REC) traits

Trait†	Mean	SD‡	CV§, %
<b>Milk traits</b>			
Milk <sub>yield</sub> , kg/d	24.95	7.84	31
Fat <sub>yield</sub> , kg/d	1.09	0.41	38
Protein <sub>yield</sub> , kg/d	0.92	0.27	30
Fat, %	4.37	0.89	20
Protein, %	3.71	0.41	11
Lactose, %	4.86	0.19	4
pH	6.64	0.08	1
SCS, units	2.87	1.86	65
<b>Milk protein fractions, %</b>			
$\alpha$ S <sub>1</sub> -CN	25.69	1.79	7
$\alpha$ S <sub>1</sub> -CN-Ph	1.45	0.61	42
$\alpha$ S <sub>2</sub> -CN	9.20	1.14	12
$\beta$ -CN	32.26	2.45	8
$\kappa$ -CN	9.44	1.49	16
$\beta$ -LG	8.68	1.55	18
a-LA	2.39	0.50	21
Other N compounds	10.89	2.24	21
<b>Curd firming</b>			
RCT <sub>eq</sub> , min	20.96	6.16	29
CF <sub>P</sub> , mm	49.20	9.71	20
k <sub>CF</sub> , % × min <sup>-1</sup>	12.90	4.07	32
k <sub>SR</sub> , % × min <sup>-1</sup>	1.23	0.46	37
CF <sub>max</sub> , mm	36.91	7.25	20
t <sub>max</sub> , min	41.83	12.40	30
<b>Cheese yield (%CY)</b>			
%CY <sub>CURD</sub>	14.95	1.84	12
%CY <sub>SOLIDS</sub>	7.17	0.91	13
%CY <sub>WATER</sub>	7.77	1.26	16
<b>Nutrient recovery (REC, %)</b>			
REC <sub>SOLIDS</sub>	51.80	3.50	7
REC <sub>FAT</sub>	89.75	3.60	4
REC <sub>PROTEIN</sub>	78.16	2.44	3
REC <sub>ENERGY</sub>	67.15	3.28	5

†SCS =  $\log_2(\text{SCC}/100\,000) + 3$ . Milk protein fractions: CN, casein; LA, lactalbumin and LG, lactoglobulin. Curd firming: RCT<sub>eq</sub>, estimated RCT; CF<sub>P</sub>, asymptotical potential value of CF; k<sub>CF</sub>, curd-firming instant rate constant; k<sub>SR</sub>, syneresis instant rate constant; CF<sub>max</sub>, maximum curd firmness achieved within 90 min; and t<sub>max</sub>, time at achievement of CF<sub>max</sub>. %CY, ratios of the weight (g) of the fresh curd (%CY<sub>CURD</sub>), curd dry matter (%CY<sub>SOLIDS</sub>) and curd water (%CY<sub>WATER</sub>) vs. the weight of the processed milk (g); REC, ratio of the weight (g) of the curd constituent (dry matter, fat, protein or energy, respectively) vs. that of the same constituent in the processed milk (g).

‡Standard deviation.

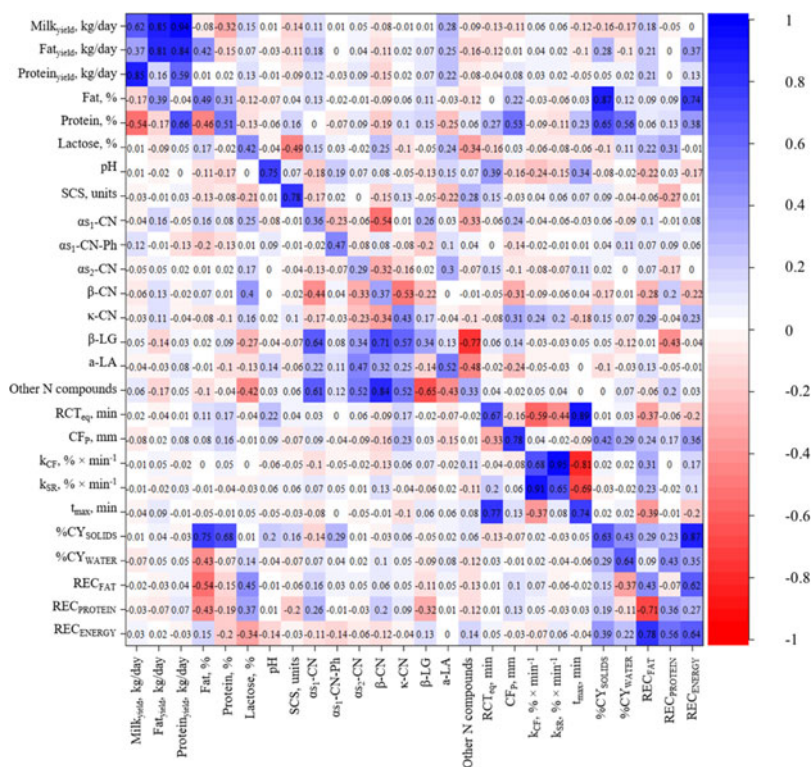
§Coefficient of variation.

modern systems, the ‘No TMR’ herds were compared with the TMR herds; (iii) within the TMR herds, those that use silage were compared with those that use water.

## Results

### Factors

Descriptive statistics of the full dataset including all 29 phenotypes are shown in Table 1. The phenotypic Pearson and



**Fig. 1.** Pearson (above the diagonal) and partial (under the diagonal) phenotypic correlations and among the 25 traits used in the factor analysis. On the diagonal the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy per trait.

partial correlations of the 26 traits analysed in MFA are presented in Fig. 1. The average KMO value in our dataset was 0.55. Ten FS, explaining 74% of the original variance, were kept for further analysis. The varimax rotated factor loadings (sorted by maximum variance explained) are shown in Tables 2 and 3.

The first factor was heavily and positively loaded on %CY<sub>SOLIDS</sub>, fat and protein (%) and REC<sub>ENERGY</sub>, thus it was considered as a factor underlying the latent concept of percentage of cheese yield (F1: %CY). The second factor was linked to all CF<sub>t</sub> traits, but CF<sub>p</sub>, and the REC<sub>FAT</sub>, reflecting the curd firmness process (F2: CF<sub>t</sub>). Due to positive loadings of the instant rate constants (k<sub>CF</sub> and k<sub>SR</sub>) and negative to the time traits (RCT<sub>eq</sub> and t<sub>max</sub>), the factor was ascribed to a favourable CF<sub>t</sub> meaning. Moreover, this factor was favourably related to REC<sub>FAT</sub>. The third factor was associated with the milk, fat and protein daily yields of individual cows, and hence was considered as a descriptor of the milk production (F3: Yield). These 3 factors were almost equally important and together explained 38% of total variability of the 26 milk traits considered. The following 7 factors were less important, each one explaining between 4 to 7% of total variability. The fourth factor was heavily, but negatively, associated with β-LG, the quantitatively most important and variable whey protein fraction, and positively to other N compounds, representing a proxy of casein number, and being positively linked also to REC<sub>PROTEIN</sub>. Hence this factor was representative of the N

found in the cheese (F4: Cheese N). The fifth factor was primary linked to α<sub>1</sub>-CN (positively) and then to the β-CN (negatively) and was considered as representative of α<sub>1</sub>- and β-CN importance in milk (F5: α<sub>1</sub>-β-CN). The sixth factor was positively associated with lactose and negatively with the SCS and the other N compounds. This factor was considered to be an indicator of the udder health status of the cow (F6: Udder health). The seventh factor was strongly and positively associated to the κ-CN and negatively, with a weaker loading, with the β-CN (F7: κ-β-CN). Thus, an increase of this factor was associated to an increased importance of κ-CN. The last 3 Fs were 1 trait-1 factor associations and were named according to the phenotype they were linked to as F8: α<sub>2</sub>-CN, F9: α<sub>1</sub>-CN-Ph and F10: α-LA, respectively.

*Sources of variation*

Table 4 summarises the results of the analysis of variance. The dairy system strongly affected F3: Yield and F1: %CY and had also a negative effect on udder health. The parity of the cow affected four Fs: F3: Yield, F2: CF<sub>t</sub>, F6: Udder health and also F9: α<sub>1</sub>-CN-Ph. The days in milk influenced all factors but F4: Cheese N, F9: α<sub>1</sub>-CN-Ph and F7: κ-β-CN. Factors F3: Yield and F6: Udder health were the only Fs influenced by all effects in the model. On the contrary, none of the effects that the model accounted for affected F4: Cheese N.

**Table 2.** Rotated factor (F) pattern, factor name and variance explained by the Fs for F1 through F5†

Trait‡	F1 %CY	F2 CF <sub>t</sub>	F3 Yield	F4 Cheese N	F5 as <sub>1</sub> -β-CN
<b>Milk traits</b>					
Milk <sub>yield</sub> , kg/d	-0.19	0.08	<b>0.96</b>	0.00	0.03
Fat <sub>yield</sub> , kg/d	0.28	0.08	<b>0.89</b>	-0.02	0.06
Prot <sub>yield</sub> , kg/d	0.00	0.00	<b>0.97</b>	-0.04	0.03
Fat, %	<b>0.90</b>	0.01	0.07	-0.04	0.06
Prot, %	<b>0.59</b>	-0.22	-0.11	-0.14	0.02
Lactose, %	-0.07	0.01	0.08	0.05	-0.01
pH	-0.08	-0.31	0.00	0.11	-0.13
SCS, units	0.06	-0.02	-0.08	0.04	-0.05
<b>Milk protein fractions, %</b>					
as <sub>1</sub> -CN	0.04	0.01	0.07	-0.14	<b>0.94</b>
as <sub>1</sub> -CN-Ph	0.04	-0.02	0.00	0.09	-0.10
as <sub>2</sub> -CN	0.02	-0.08	0.04	-0.06	0.01
β-CN	-0.10	-0.05	-0.11	0.12	<b>-0.70</b>
κ-CN	0.12	0.17	0.00	-0.09	0.05
β-LG	0.05	-0.03	0.02	<b>-0.98</b>	0.12
α-LA	-0.06	-0.02	0.18	-0.14	-0.02
Other N comp.	-0.05	0.01	-0.02	<b>0.76</b>	-0.09
<b>Curd firming</b>					
RCT <sub>eq</sub> , min	0.01	<b>-0.74</b>	-0.05	-0.06	-0.01
CF <sub>p</sub> , mm	0.38	0.03	0.01	-0.08	0.23
k <sub>CF</sub> , % × min <sup>-1</sup>	-0.01	<b>0.94</b>	0.00	0.00	-0.03
k <sub>SR</sub> , % × min <sup>-1</sup>	-0.06	<b>0.88</b>	0.00	-0.01	-0.05
t <sub>max</sub> , min	0.04	<b>-0.90</b>	-0.04	-0.03	-0.01
<b>Cheese yield (%CY)</b>					
%CY <sub>SOLIDS</sub>	<b>0.99</b>	0.02	0.04	0.00	0.04
%CY <sub>WATER</sub>	0.39	-0.05	-0.07	0.11	-0.08
<b>Nutrient recovery (REC, %)</b>					
REC <sub>FAT</sub>	0.28	<b>0.44</b>	0.19	0.01	0.11
REC <sub>PROTEIN</sub>	0.23	-0.02	-0.02	<b>0.46</b>	-0.03
REC <sub>ENERGY</sub>	<b>0.87</b>	0.25	0.14	0.10	0.08
Cumulative variance, %	0.14	0.27	0.38	0.45	0.51

†Factors have been sorted based on proportion of variance explained. Values above |0.4| in bold. F1: %CY, Factor related to the percentage of individual cheese yield; F2: CF<sub>t</sub>, Factor related to the curd firmness; F3: Yield, Factor related to the milk yield; F4: Cheese N, Factor related to the milk nitrogen that is present into the cheese curd; F5: as<sub>1</sub>-β-CN, Factor related to the as<sub>1</sub>- and β-CN contents in milk, expressed as relative contents to the total milk nitrogen.

‡SCS = log<sub>2</sub> (SCC × 100 000) + 3. Milk protein fractions: CN, casein; LA, lactalbumin and LG, lactoglobulin. Curd firming: RCT<sub>eq</sub>, estimated RCT; CF<sub>p</sub>, asymptotical potential value of CF; k<sub>CF</sub>, curd-firming instant rate constant; k<sub>SR</sub>, syneresis instant rate constant; and t<sub>max</sub>, time at achievement of CF<sub>max</sub>. %CY, ratios of the weight (g) of the curd dry matter (%CY<sub>SOLIDS</sub>) and curd water (%CY<sub>WATER</sub>) vs. the weight of the processed milk (g); REC, ratio of the weight (g) of the curd constituent (fat, protein or energy, respectively) vs. that of the same constituent in the processed milk (g).

Investigating the effect of the dairy system, Table 5 outlines the least squares means (LSM) of the 10 Fs within the 4 different dairy systems and their orthogonal contrasts. Our analysis noted major differences between traditional and modern farms, in favour of the second, for F3: Yield

and F1: %CY, and a smaller effect, in favour of the traditional farms, for F6: Udder health. Moreover, for F3: Yield, F1: %CY and F8 as<sub>2</sub>-CN, the modern farms using TMR showed greater results than those not using TMR. No difference was found in relation to the source of moisture in the TMR.

The LSM of the Fs on which the DIM had an effect are presented in Fig. 2. A ‘lactation like’ pattern (with the peak value in the second month of lactation) was observed for F6: Udder health, and also for F2: α-LA (Fig. 2e, g), revealing, after lactation peak, a decrease in lactose, and an increase of SCS. Lactose was the trait on which the factor F6: Udder health was primary loaded. On the contrary, the latent variable F3: Yield was almost linearly decreased during lactation (Fig. 2c). This pattern is primarily due to daily yield of fat and protein, and not to fresh milk (that peaked in the second month). The strong decrease at the beginning of lactation of milk fat and protein content (and also of their daily yield), paralleled by REC<sub>SOLIDS</sub>, is reflected by the ‘inverse lactation like’ pattern of the F1: %CY.

A decrease from 1st to 3rd class of DIM was observed for F2: CF<sub>t</sub> followed by a stabilisation and a slight increase at the end of the lactation (Fig. 2b). An opposite pattern was observed between F5: as<sub>1</sub>-β-CN and F8: as<sub>2</sub>-CN (Fig. 2d, f). F8: as<sub>2</sub>-CN was increasing rapidly at the beginning of lactation when F5: as<sub>1</sub>-β-CN decreased. A more stable situation was observed thereafter (Fig. 2f).

Figure 3 presents the LSM across the parity levels for the four Fs affected by parity. Factors F3: Yield and F9: as<sub>1</sub>-CN-Ph had a similar pattern with an increase from the 1st to the 3rd parity, and a slight decline from the 4th parity (Fig. 3b, d). The F2: CF<sub>t</sub> showed almost no difference between 1st and 3rd parities but an increase after (Fig. 3a). A linear decrease was observed for the factor F6: Udder health related to the mammary gland health status (Fig. 3c).

## Discussion

### Interpretation of the factors

Ten orthogonal latent variables were extracted, out of 26 milk (quantity and quality), milk coagulation and curd firmness indicators, and individual cheese yield traits. The Fs were explaining 74% of the original variability while at the same time drastically reduced the data space. The average KMO was quite close to the value (0.57) reported by Manca et al. (2016) using a similar, but smaller, dataset of 12 milk composition, MCP and udder health phenotypes in dairy sheep. Furthermore, the factor model was able to capture basic concepts of the cheese-making process. The same factor scores have been previously used for estimating (co)variance components using standard quantitative genetic model (Dadousis et al. 2017a) and for conducting genome-wide associations and gene-set enrichment analyses (Dadousis et al. 2017b). Results were coherent with the given name of the factors. More precisely, the first four Fs, sorted by variance explained, were able to capture the

**Table 3.** Rotated factor (F) pattern, factor name, communality (com)<sup>†</sup> of variables and variance explained by the Fs for F6 through F10<sup>‡</sup>

Trait <sup>§</sup>	F6 Udder health	F7 $\kappa$ - $\beta$ -CN	F8 as <sub>2</sub> -CN	F9 as <sub>1</sub> -CN-Ph	F10 $\alpha$ -LA	com
<b>Milk traits</b>						
Milk <sub>yield</sub> , kg/d	0.11	0.01	0.01	0.02	0.08	0.99
Fat <sub>yield</sub> , kg/d	0.17	-0.01	-0.01	-0.01	0.07	0.92
Prot <sub>yield</sub> , kg/d	0.02	0.02	0.05	-0.02	0.03	0.96
Fat, %	0.08	-0.06	-0.03	-0.05	0.00	0.84
Prot, %	-0.30	0.02	0.08	-0.07	-0.17	0.56
Lactose, %	<b>0.62</b>	-0.01	0.00	0.03	0.04	0.40
pH	-0.02	0.03	0.04	0.17	0.15	0.18
SCS, units	<b>-0.41</b>	0.09	0.01	0.03	-0.09	0.20
<b>Milk protein fractions, %</b>						
as <sub>1</sub> -CN	0.25	-0.04	-0.08	-0.14	-0.04	0.99
as <sub>1</sub> -CN-Ph	0.01	-0.04	-0.04	<b>0.98</b>	0.06	1.00
as <sub>2</sub> -CN	-0.03	-0.07	<b>0.98</b>	-0.04	0.14	1.00
$\beta$ -CN	0.38	<b>-0.47</b>	-0.33	-0.03	-0.05	1.00
$\kappa$ -CN	-0.11	<b>0.96</b>	-0.09	-0.04	0.01	1.00
$\beta$ -LG	0.01	0.07	-0.05	-0.10	0.00	0.99
$\alpha$ -LA	0.30	0.01	0.17	0.07	<b>0.90</b>	0.99
Other N comp.	<b>-0.60</b>	-0.10	-0.01	-0.02	-0.21	1.00
<b>Curd firming</b>						
RCT <sub>eq</sub> , min	-0.15	0.02	0.08	0.00	-0.01	0.59
CF <sub>p</sub> , mm	-0.11	0.23	-0.07	-0.11	-0.20	0.32
k <sub>CF</sub> , % × min <sup>-1</sup>	-0.09	0.08	0.00	0.00	-0.01	0.90
k <sub>SR</sub> , % × min <sup>-1</sup>	-0.08	0.06	0.00	0.01	0.00	0.79
t <sub>max</sub> , min	-0.06	-0.05	0.03	0.00	-0.01	0.82
<b>Cheese yield (%CY)</b>						
%CY <sub>SOLIDS</sub>	-0.08	0.02	0.01	0.01	-0.03	0.99
%CY <sub>WATER</sub>	-0.03	0.05	0.01	0.08	0.02	0.19
<b>Nutrient recovery (REC, %)</b>						
REC <sub>FAT</sub>	0.00	0.21	0.10	0.08	0.12	0.39
REC <sub>PROTEIN</sub>	0.23	0.00	-0.14	0.03	-0.02	0.34
REC <sub>ENERGY</sub>	-0.03	0.10	0.01	0.03	0.06	0.88
Cumulative variance, %	0.56	0.61	0.66	0.70	0.74	

<sup>†</sup>Communality, the sum of the squared factor loadings per trait.

<sup>‡</sup>Factors have been sorted based on proportion of variance explained. Values above |0.4| in bold. F6: Udder health, Factor related to the udder health of a cow; F7:  $\kappa$ - $\beta$ -CN, Factor related to the  $\kappa$ - and  $\beta$ -CN contents in milk, expressed as relative contents to the total milk nitrogen; F8: as<sub>2</sub>-CN, Factor related to the milk as<sub>2</sub>-CN, expressed as relative content to the total milk nitrogen; F9: as<sub>1</sub>-CN-Ph, Factor related to the milk as<sub>1</sub>-CN-Ph expressed as content to the total milk nitrogen; F10:  $\alpha$ -LA, Factor related to the milk  $\alpha$ -LA.

<sup>§</sup>SCS = log<sub>2</sub> (SCC / 100 000) + 3. Milk protein fractions: CN, casein; LA, lactalbumin and LG, lactoglobulin. Curd firming: RCT<sub>eq</sub>, estimated RCT; CF<sub>p</sub>, asymptotical potential value of CF; k<sub>CF</sub>, curd-firming instant rate constant; k<sub>SR</sub>, syneresis instant rate constant; and t<sub>max</sub>, time at achievement of CF<sub>max</sub>. %CY, ratios of the weight (g) of the curd dry matter (%CY<sub>SOLIDS</sub>) and curd water (%CY<sub>WATER</sub>) vs. the weight of the processed milk (g); REC, ratio of the weight (g) of the curd constituent (fat, protein or energy, respectively) vs. that of the same constituent in the processed milk (g).

underlying structure of the cheese yield (%), the curd firming process, the milk yield and the presence of N into the cheese. Moreover, 4 Fs were associated to the basic milk caseins (as<sub>1</sub>- $\beta$ -CN,  $\kappa$ - $\beta$ -CN, as<sub>2</sub>-CN and as<sub>1</sub>-CN-Ph) and 1 factor was related with the, quantitatively, most important whey protein ( $\alpha$ -LA). A factor describing the udder health status of a cow, mainly loaded on lactose, other N compounds and SCS, was also obtained.

The meaning of some of the Fs is comparable to previous studies that used MFA, albeit with smaller datasets. For example, in an experimental design study assessing the effect of DIM, body condition and milk protein genotypes on the MCP and the protein composition in Danish Holstein ( $n = 39$ ) 8 Fs were extracted out of 28 measured variables (Ostensen et al. 1997). Among the measured

phenotypes were milk yield traits (milk, fat, protein and lactose), whey and non-casein contents, proportion of caseins, proportion of alleles in genotyped milk proteins, MCP and energy balance and the body condition. The 8 Fs obtained represented MCP properties of milk, milk protein genotypes, the energy and the body condition status of the cows, and effects of stage of lactation on milk yield and protein composition. In another study, eleven milk composition, MCP and udder health phenotypes measured in Brown Swiss cows ( $n = 1200$ ) were substituted by 4 Fs, describing milk composition, coagulation, acidity and the udder health status (Macciotta et al. 2012). In small ruminants, Todaro et al. (2005) replaced 11 traits related to milk composition, MCP and udder health measured on 117 Girgentana goats with 3 Fs named as 'slow milks',

**Table 4.** Analysis of variance (*F*-values and significance) of the 10 factors (F)<sup>†</sup>

Item	Dairy system <i>F</i> -value	Parity <i>F</i> -value	DIM <i>F</i> -value	Residual RMSE
DF	3	3	10	–
F1: %CY	6.25***	2.41	13.34***	0.83
F2: CF <sub>t</sub>	1.78	7.38***	5.96***	0.84
F3: Yield	19.54***	46.37***	30.42***	0.59
F4: Cheese N	1.18	0.75	1.74	0.76
F5: as <sub>1</sub> -β-CN	1.43	0.74	7.47***	0.84
F6: Udder health	3.39*	32.88***	14.40***	0.71
F7: κ-β-CN	0.81	0.23	1.28	0.93
F8: as <sub>2</sub> -CN	1.68	1.62	4.16***	0.78
F9: as <sub>1</sub> -CN-Ph	0.48	3.60*	1.14	0.73
F10: α-LA	1.71	1.60	11.09***	0.72

\**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.001.

†F1: %CY, Factor related to the percentage of individual cheese yield; F2: CF<sub>t</sub>, Factor related to the curd firmness; F3: Yield, Factor related to the milk yield; F4: Cheese N, Factor related to the milk nitrogen that is present into the cheese curd; F5: as<sub>1</sub>-β-CN, Factor related to the as<sub>1</sub>- and β-CN contents in milk, expressed as relative contents to the total milk nitrogen; F6: Udder health, Factor related to the udder health of a cow; F7: κ-β-CN, Factor related to the κ- and β-CN contents in milk, expressed as relative contents to the total milk nitrogen; F8: as<sub>2</sub>-CN, Factor related to the milk as<sub>2</sub>-CN, expressed as relative content to the total milk nitrogen; F9: as<sub>1</sub>-CN-Ph, Factor related to the milk as<sub>1</sub>-CN-Ph expressed as content to the total milk nitrogen; F10: α-LA, Factor related to the milk α-LA.

‘milk yield’ and ‘curd firmness’. Recently, Vacca et al. (2016), by analysing 9 milk yield, composition and hygiene traits in 1050 Sardinian goats obtained 4 Fs (quality, hygiene, production, and acidity), while Manca et al. (2016) in a study on 991 Sarda ewes, extracted 4 Fs (composition and cheese yield, udder health status, coagulation and curd characteristics) from a total of 12 measured phenotypes. All the 5 above mentioned studies used an estimatory MFA with a *varimax* rotation. The proportion of variance explained by the factor model in those studies ranged between 51.2% (Todaro et al. 2005) to 97% in Ostersen et al. (1997). Part of this difference, however, could be attributed to the different ways and methods of measuring the phenotypes used in MFA.

Although a direct, factor by factor, comparison between our results and the above mentioned studies is problematic due to differences in the measured phenotypes analysed (number and type), a consistent pattern can be observed: generally, in all the studies MFA clearly distinguishes between milk quality, production, coagulation and health related concepts. These results are encouraging and confirm the ability of MFA to capture basic underlying structure of correlated variables in different species and rearing conditions.

Factor F1: %CY was positively related to fat and protein percentage as well as the %CY<sub>SOLIDS</sub> and the REC<sub>ENERGY</sub>. This is in agreement with the phenotypic and genetic correlations that have been reported in the literature for these traits (Bittante et al. 2013a) and confirms that cheese produced is

not only due to fat and protein content of milk but also to the ability to recover them in curd or lose with whey.

The inverse relationship of RCT<sub>eq</sub> and *t*<sub>max</sub> with k<sub>CF</sub> and k<sub>SR</sub> found on F2: CF<sub>t</sub> has been previously pointed out in Cecchinato & Bittante (2016) and the inclusion of REC<sub>FAT</sub> on the factor confirm that coagulation properties of milk affect its ability to retain fat in the curd or to lose it with whey. On factor F5: as<sub>1</sub>-β-CN, the as<sub>1</sub>- and β-CN were oppositely related. The antagonistic nature of those two caseins has been previously suggested (Bobé et al. 1999; Bonfatti et al. 2010). Moreover, the factor F6: Udder health was inversely related to other N compounds that are present in the milk. It is worth noting that within this N fraction, urea was also included. Urea is inversely related to lactose (Miglior et al. 2007), which was the trait most strongly related to factor F6: Udder health. Herein we considered, for the first time, the milk protein profile along with different milk technological traits. More specifically, a set of 26 relative new phenotypes were considered: 3 production traits (i.e., milk, fat and protein yield), 8 milk protein fractions, 5 traits related to curd firming modelling, 2 cheese yields and 3 nutrients recovery traits. Out of those, only 5 standard milk quality traits were in common with the study of Macciotta et al. (2012) (i.e. fat%, protein %, lactose%, pH and SCC).

To further assess the practical application in dairy farming and in breeding programs of the new set of the latent variables, an investigation on the sources affecting variation on those new traits was followed.

#### Sources of variation of the factors

**Effects of dairy system on the extracted factor scores.** Major differences between the different dairy systems were identified for 2 Fs, namely F3: Yield and F1: %CY. Modern farms were associated with a higher milk and solids yield and percentage of cheese obtained for every 100 kg of milk processed. Moreover, within the modern farms, the use of TMR in the feeding system was found beneficial for both latent variables. Only Vacca et al. (2016) on lactating goats studied the relationships between Fs and dairy system and farms characteristics: they did not find any associations between Fs and dairy system, but observed an association between altitude of the farm and the production factor and between flock size and the acidity factor.

Nonetheless, the results of our study are in agreement with previous findings on dairy cattle carried out on the measured daily milk yield production traits instead of Fs (Sturaro et al. 2009, 2013; Bittante et al. 2015). Moreover, although individual cheese traits were not included in the analysis of Bittante et al. (2015), major differences between the dairy systems were reported for milk fat and protein percentages. These two traits were both loaded on F1: %CY in our study. No significant differences between the dairy systems were observed for the coagulation properties of milk (F2: CF<sub>t</sub>). The F2: CF<sub>t</sub> was primary loaded to the k<sub>CF</sub>, but also to RCT<sub>eq</sub>, k<sub>SR</sub>, *t*<sub>max</sub> and REC<sub>FAT</sub>. These findings

**Table 5.** Effects of the dairy system (traditional with tied cows, modern with loose cows), the use of total mixed ration (TMR) within modern farms, and of the moisture source of TMR on the 10 factors (F)<sup>†</sup>

Dairy system	Dairy system LSM				Orthogonal contrasts <i>F</i> -values		
	Traditional	Modern		Orthogonal contrasts <i>F</i> -values			
		No TMR	TMR		Modern vs. traditional <sup>‡</sup>	TMR vs. no TMR <sup>§</sup>	Silage vs. water <sup>¶</sup>
Moisture source	–	Silage	Water				
Herds	29	30	9	17	–	–	–
F1: %CY	–0.25	–0.11	0.40	0.18	13.21***	10.22**	1.37
F2: CF <sub>t</sub>	0.15	–0.13	0.02	0.04	2.48	1.50	0.01
F3: Yield	–0.53	0.16	0.59	0.60	53.31***	8.38**	0.12
F4: Cheese N	0.19	–0.02	–0.26	–0.04	3.29	0.52	0.63
F5: as <sub>1</sub> -β-CN	0.21	–0.002	0.02	–0.12	3.34	0.10	0.42
F6: Udder health	0.28	–0.25	–0.18	0.05	5.98*	1.28	0.90
F7: κ-β-CN	–0.002	–0.08	0.19	–0.01	0.10	1.75	1.17
F8: as <sub>2</sub> -CN	–0.12	–0.21	0.22	0.14	1.22	4.66*	0.09
F9: as <sub>1</sub> -CN-Ph	0.08	0.02	–0.24	0.08	0.52	0.23	1.14
F10: α-LA	–0.12	–0.11	0.35	0.20	2.43	3.95	0.26

\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

<sup>†</sup>F1: %CY, Factor related to the percentage of individual cheese yield; F2: CF<sub>t</sub>, Factor related to the curd firmness; F3: Yield, Factor related to the milk yield; F4: Cheese N, Factor related to the milk nitrogen that is present into the cheese curd; F5: as<sub>1</sub>-β-CN, Factor related to the as<sub>1</sub>- and β-CN contents in milk, expressed as relative contents to the total milk nitrogen; F6: Udder health, Factor related to the udder health of a cow; F7: κ-β-CN, Factor related to the κ- and β-CN contents in milk, expressed as relative contents to the total milk nitrogen; F8: as<sub>2</sub>-CN, Factor related to the milk as<sub>2</sub>-CN, expressed as relative content to the total milk nitrogen; F9: as<sub>1</sub>-CN-Ph, Factor related to the milk as<sub>1</sub>-CN-Ph expressed as content to the total milk nitrogen; F10: α-LA, Factor related to the milk α-LA.

<sup>‡</sup>Contrast between the 'Traditional' dairy system vs. the three 'Modern' ones.

<sup>§</sup>Contrast between the 'Modern No TMR' dairy system vs. the two 'Modern TMR' ones.

<sup>¶</sup>Contrast between the 'Modern TMR Silage' dairy system vs. the 'Modern TMR Water' one.

are consistent with Bittante et al. (2015), where no significant differences were noted among the dairy systems for  $k_{CF}$  and  $RCT_{eq}$ . A significant effect of the dairy systems existed in that study for  $k_{SR}$  ( $P < 0.05$ ) and  $t_{max}$  ( $P < 0.01$ ). However, since a factor is a mixture of phenotypes, and in our case F2: CF<sub>t</sub> was dominated by the  $k_{CF}$ , this effect was diluted.

**Effect of stage of lactation and parity on the extracted factor scores.** The stage of lactation significantly influenced most of the Fs. As expected, the F3: Yield was decreased during lactation, in agreement to Bittante et al. (2015). The absence of the peak of lactation is probably due to the length of the DIM classes, but it should be noted that, analysing mixed-breed farms in the same area, Stocco et al. (2016) observed the presence of peak of lactation for cows reared in farms characterised by high average milk yield, but not in those with a low milk productivity.

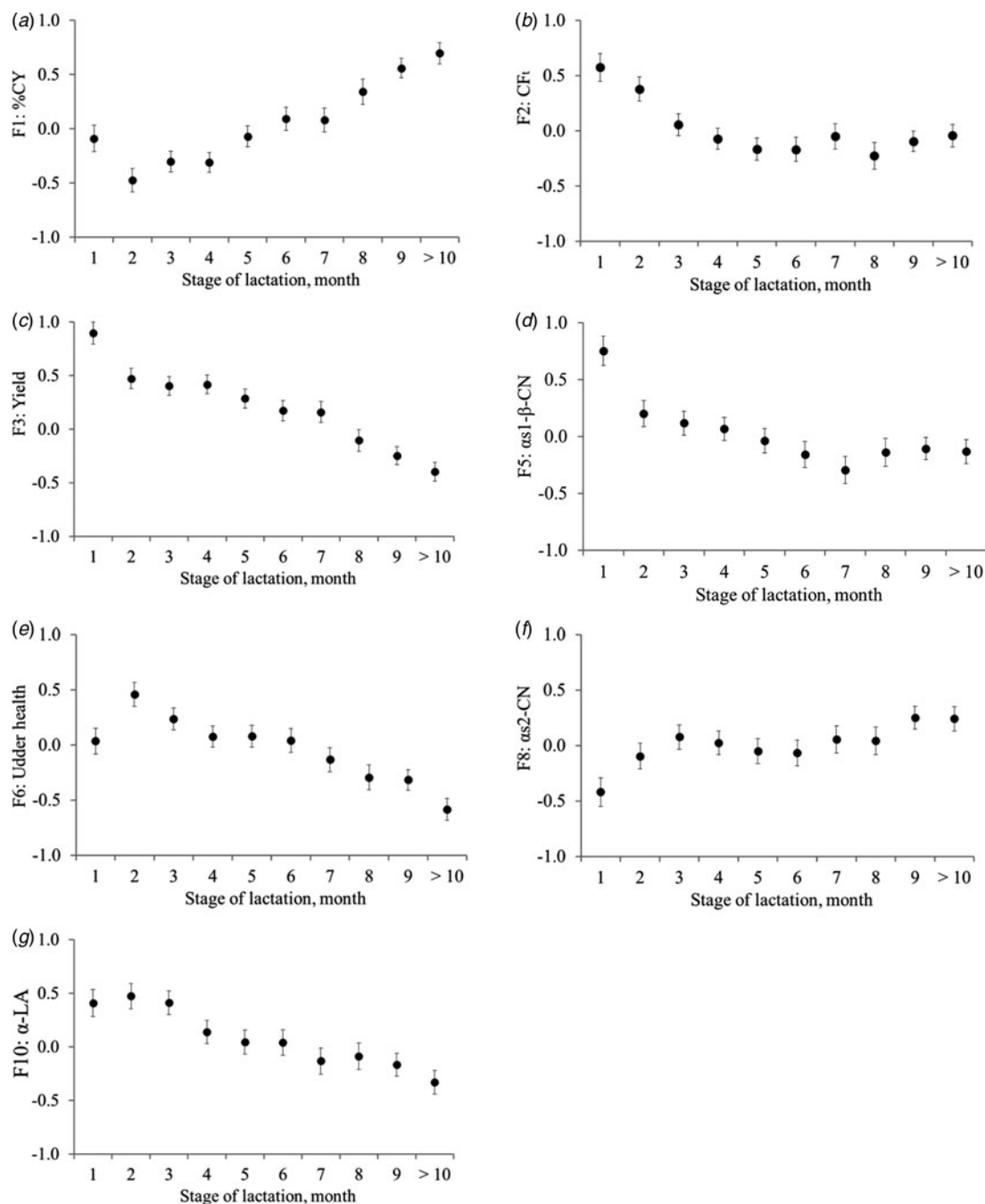
A tendency towards worsening of milk coagulation during lactation is known. This pattern has been reported in Macciotta et al. (2012) using a factor as indicator of milk coagulation (albeit based on the traditional MCP values). Moreover, the same trend has been observed in previous studies using the traditional curd-firmness value at 30 min after the beginning of coagulation, generally known as  $a_{30}$  (Ikonen et al. 2004). Being consistent, the coagulation

ability (F2: CF<sub>t</sub>) was smoothly decreased (worsening) during early lactation, stabilised between 4th and 8th class of DIM, with an evidence of improvement at the end of the lactation. Moreover, the factor related to the percentage of cheese yield (F1: %CY) showed a decrease from the first to the second month of lactation and then a linear increase. Not surprisingly, this pattern during lactation is consistent with the trend of %CY<sub>SOLIDS</sub> on which this factor was positively and strongly related (Cipolat-Gotet et al. 2013).

In line with our expectations, the F10: α-LA was decreased during lactation having an opposite trend compared to the F1: %CY. The same pattern has been reported for the measured α-LA in Ostensen et al. (1997). A smooth decrease of F5: as<sub>1</sub>-β CN during lactation and the inverse pattern of F8: as<sub>2</sub>-CN was observed, as expected.

The pattern of the health status of the mammary gland mainly related to the lactose (F6: Udder health) was worsened across the course of lactation. As it is known, lactose in milk is inversely related to mastitis or somatic cell count (Shuster et al. 1991) and SCS has an opposite trend within lactation compared to milk yield (Walsh et al. 2007). Similar results have been reported using MFA by Macciotta et al. (2012) in dairy cattle and in Manca et al. (2016) in dairy sheep. Also in those studies, lactose was the most related trait to the factor underlying the health status of the mammary gland.

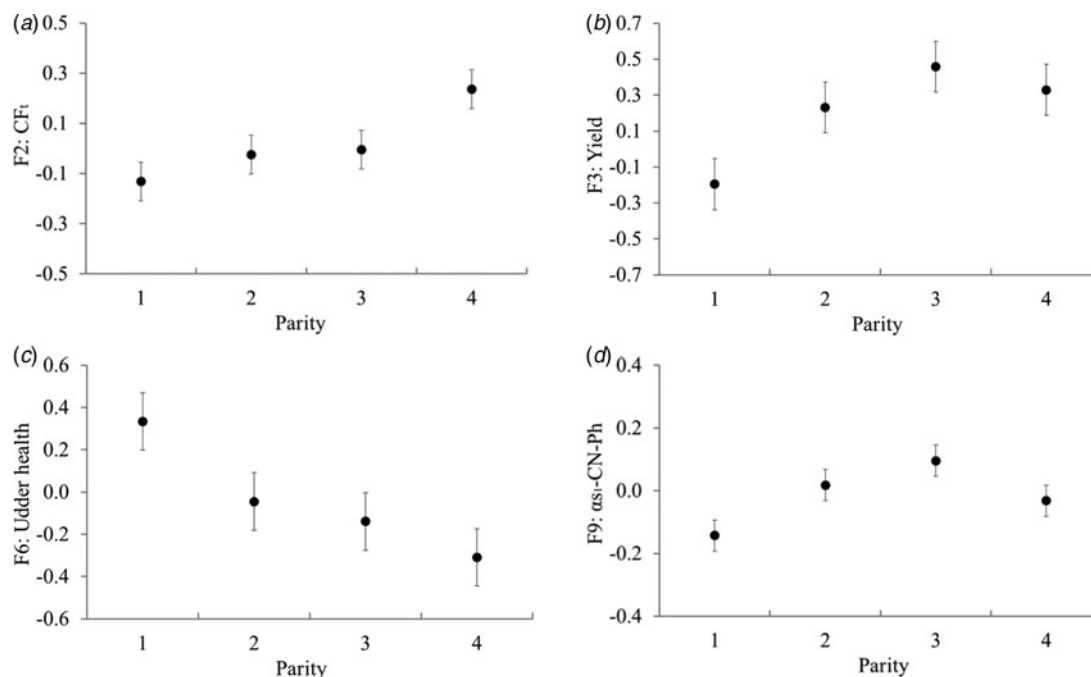




**Fig. 2.** Least squares means of the seven factors across stage of lactation. Description: F1: %CY, Factor related to the percentage of individual cheese yield; F2:  $CF_t$ , Factor related to the curd firmness; F3: Yield, Factor related to the milk yield; F5:  $as_1\text{-}\beta\text{-CN}$ , Factor related to the  $as_1$ - and  $\beta\text{-CN}$  contents in milk, expressed as relative contents to the total milk nitrogen; F6: Udder health, Factor related to the udder health of a cow; F8:  $as_2\text{-CN}$ , Factor related to the milk  $as_2\text{-CN}$ , expressed as relative content to the total milk nitrogen F10:  $\alpha\text{-LA}$ , Factor related to the milk  $\alpha\text{-LA}$ .

The trend of parity effect on F3: Yield was similar to previous results on milk yield traits (Bittante et al. 2015). Moreover, in the same study the  $k_{CF}$  values were lower in first lactation cows compared to following lactations, similar to the effect of parity on the F2:  $CF_t$  found in our

analysis. Furthermore, a decline of F6: Udder health with the increase of parity is in agreement with the literature where a decrease in milk lactose percentage and an increase of SCS has been observed in multiparous cows (Yang et al. 2013). Parity also affected F9:  $as_1\text{-CN-Ph}$



**Fig. 3.** Least squares means of the four factors across the levels of parity. Description: F2: CF<sub>t</sub>, Factor related to the curd firmness; F3: Yield, Factor related to the milk yield; F6: Udder health, Factor related to the udder health of a cow; F9: as<sub>1</sub>-CN-Ph, Factor related to the milk as<sub>1</sub>-CN-Ph expressed as content to the total milk nitrogen.

latent variable in our analysis with a pattern among different levels of parities mimicking the trend of F3: Yield.

## Conclusions

By applying multivariate factor analysis, 26 original phenotypes were substituted by 10 latent variables. These variables captured important concepts of the cheese-making process as well as basic bovine health indicators related to the mammary gland. Results of ANOVA were in agreement to the given name of the factor and reflected the underlying structure that each factor was representing. This approach, therefore, represents a valuable tool for studying the effects of different production systems, feeding regimes, and health status on the milk protein composition and cheese-making related traits, and for identifying strategies for altering such technological traits in accordance with dairy industry.

The authors wish to thank the Trento Province (Italy), the Italian Brown Swiss Cattle Breeders Association (ANARB, Verona, Italy) and the Superbrown Consortium of Bolzano and Trento for financial and technical support. C. Dadousis benefitted from financial support of the CARIPARO (Cassa di Risparmio di Padova e Rovigo) Foundation (Padua, Italy). The authors also wish to acknowledge Valentina Bonfatti for her cooperation in assessing the milk protein fractions.

## References

Ali A & Shook G 1980 An optimum transformation for somatic cell concentration in milk. *Journal of Dairy Science* **63** 487–490

- Bergamaschi M, Biasioli F, Cappellin L, Cecchinato A, Cipolat-Gotet C, Cornu A, Gasperi F, Martin B & Bittante G 2015 Proton transfer reaction time-of-flight mass spectrometry: a high-throughput and innovative method to study the influence of dairy system and cow characteristics on the volatile compound fingerprint of cheeses. *Journal of Dairy Science* **98** 8414–8427
- Bittante G, Cipolat-Gotet C & Cecchinato A 2013a Genetic parameters of different measures of cheese yield and milk nutrient recovery from an individual model cheese-manufacturing process. *Journal of Dairy Science* **96** 7966–7979
- Bittante G, Contiero B & Cecchinato A 2013b Prolonged observation and modelling of milk coagulation, curd firming, and syneresis. *International Dairy Journal* **29** 115–123
- Bittante G, Cipolat-Gotet C, Malchioldi F, Sturaro E, Tagliapietra F, Schiavon S & Cecchinato A 2015 Effect of dairy farming system, herd, season, parity, and days in milk on modeling of the coagulation, curd firming, and syneresis of bovine milk. *Journal of Dairy Science* **98** 2759–2774
- Bobé G, Beitz D, Freeman A & Lindberg G 1999 Effect of milk protein genotypes on milk protein composition and its genetic parameter estimates. *Journal of Dairy Science* **82** 2797–2804
- Bollen K 1989 *Structural Equations with Latent Variables*. John Wiley & Sons, New York
- Bonfatti V, Grigoletto L, Cecchinato A, Gallo L & Carnier P 2008 Validation of a new reversed-phase high-performance liquid chromatography method for separation and quantification of bovine milk protein genetic variants. *Journal of Chromatography A* **1195** 101–106
- Bonfatti V, Di Martino G, Cecchinato A, Vicario D & Carnier P 2010 Effects of  $\beta$ - $\kappa$ -casein (CSN2-CSN3) haplotypes and  $\beta$ -lactoglobulin (BLG) genotypes on milk production traits and detailed protein composition of individual milk of simmental cows. *Journal of Dairy Science* **93** 3797–3808
- Bonfatti V, Cecchinato A, Gallo L, Blasco A & Carnier P 2011 Genetic analysis of detailed milk protein composition and coagulation properties in simmental cattle. *Journal of Dairy Science* **94** 5183–5193
- Cecchinato A & Bittante G 2016 Genetic and environmental relationships of different measures of individual cheese yield and curd nutrients

- recovery with coagulation properties of bovine milk. *Journal of Dairy Science* **99** 1975–1989
- Cecchinato A, Cipolat-Gotet C, Casellas J, Penasa M, Rossoni A & Bittante G** 2013 Genetic analysis of rennet coagulation time, curd-firming rate, and curd firmness assessed over an extended period using mechanical and near-infrared instruments. *Journal of Dairy Science* **96** 50–62
- Cipolat-Gotet C, Cecchinato A, De Marchi M & Bittante G** 2013 Factors affecting variation of different measures of cheese yield and milk nutrient recovery from an individual model cheese-manufacturing process. *Journal of Dairy Science* **96** 7952–7965
- Conte G, Serra A, Cremonesi P, Chessa S, Castiglioni B, Cappucci A, Bulleri E & Mele M** 2016 Investigating mutual relationship among milk fatty acids by multivariate factor analysis in dairy cows. *Livestock Science* **188** 124–132
- Dadousis C, Cipolat-Gotet C, Bittante G & Cecchinato A** 2017a Inferring genetic parameters on latent variables underlying milk yield and quality, protein composition, curd firmness and cheese-making traits in dairy cattle. *Animal*. doi: 10.1017/S1751731117001616
- Dadousis C, Pegolo S, Rosa GJM, Bittante G & Cecchinato A** 2017b Genome-wide association and pathway-based analysis using latent variables related to milk protein composition and cheesemaking traits in dairy cattle. *Journal of Dairy Science* **100** 9085–9102
- Dziuban CD & Shirkey EC** 1974 When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin* **81** 358
- Ikonen T, Morri S, Tyrisevä A, Ruottinen O & Ojala M** 2004 Genetic and phenotypic correlations between milk coagulation properties, milk production traits, somatic cell count, casein content, and pH of milk. *Journal of Dairy Science* **87** 458–467
- Kaiser HF & Rice J** 1974 Little jiffy, Mark IV. *Educational and Psychological Measurement* **34** 111–117
- Macciotta NPP, Cecchinato A, Mele M & Bittante G** 2012 Use of multivariate factor analysis to define new indicator variables for milk composition and coagulation properties in brown Swiss cows. *Journal of Dairy Science* **95** 7346–7354
- Manca M, Serdino J, Gaspa G, Urgeghe P, Ibba I, Contu M, Fresi P & Macciotta N** 2016 Derivation of multivariate indices of milk composition, coagulation properties, and individual cheese yield in dairy sheep. *Journal of Dairy Science* **99** 4547–4557
- Mele M, Macciotta N, Cecchinato A, Conte G, Schiavon S & Bittante G** 2016 Multivariate factor analysis of detailed milk fatty acid profile: Effects of dairy system, feeding, herd, parity, and stage of lactation. *Journal of Dairy Science* **99** 9820–9833
- Miglior F, Sewalem A, Jamrozik J, Bohmanova J, Lefebvre D & Moore R** 2007 Genetic analysis of milk urea nitrogen and lactose and their relationships with other production traits in Canadian Holstein cattle. *Journal of Dairy Science* **90** 2468–2479
- NRC** 2001 *Nutrient Requirements of Dairy Cattle*, 7th rev. Washington, DC: Natl. Acad. Press
- Ostensen S, Foldager J & Hermansen JE** 1997 Effects of stage of lactation, milk protein genotype and body condition at calving on protein composition and renneting properties of bovine milk. *Journal of Dairy Research* **64** 207–219
- R Core Team** 2013 *R: A Language and Environment for Statistical Computing*. <http://www.R-project.org/>
- Revelle W** 2014 *Psych: Procedures for Personality and Psychological Research*. Evanston: Northwestern University, R Package Version. 1
- Shuster D, Harmon R, Jackson J & Hemken R** 1991 Suppression of milk production during endotoxin-induced mastitis. *Journal of Dairy Science* **74** 3763–3774
- Stocco G, Cipolat-Gotet C, Bobbo T, Cecchinato A & Bittante G** 2016 Herd productivity and breed of cows affect milk composition, coagulation properties and curd firming and syneresis modeling. *Journal of Dairy Science* **100** 129–145
- Sturaro E, Cocca G, Gallo L, Mrad M & Ramanzin M** 2009 Livestock systems and farming styles in eastern Italian Alps: An on-farm survey. *Italian Journal of Animal Science* **8** 541–554
- Sturaro E, Marchiori E, Cocca G, Penasa M, Ramanzin M & Bittante G** 2013 Dairy systems in mountainous areas: Farm animal biodiversity, milk production and destination, and land use. *Livestock Science* **158** 157–168
- Todaro M, Scatassa ML & Giaccone P** 2005 Multivariate factor analysis of girgentana goat milk composition. *Italian Journal of Animal Science* **4** 403–410
- Vacca G, Paschino P, Dettori M, Bergamaschi M, Cipolat-Gotet C, Bittante G & Pazzola M** 2016 Environmental, morphological, and productive characterization of sardinian goats and use of latent explanatory factors for population analysis. *Journal of Animal Science* **94** 3947–3957
- Walsh S, Buckley F, Berry D, Rath M, Pierce K, Byrne N & Dillon P** 2007 Effects of breed, feeding system, and parity on udder health and milking characteristics. *Journal of Dairy Science* **90** 5767–5779
- Yang L, Yang Q, Yi M, Pang Z & Xiong B** 2013 Effects of seasonal change and parity on raw milk composition and related indices in Chinese Holstein cows in northern China. *Journal of Dairy Science* **96** 6863–6869