Food Research International 239 (2026) 119495

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

Untargeted ^{1}H NMR metabolomics to distinguish milk from different feeding systems for haymilk authentication

Dilek Eltemur [a,b], Enmanuel Cruz Muñoz [c], Davide Ballabio [c], Ksenia Morozova [a,*], Demian Martini-Lösch [b], Giovanni Peratoner [b], Daniela Eisenstecken [b], Elena Venir [b], Peter Robatscher [b], Michael Oberhuber [b], Matteo Scampicchio [a]

[a] Faculty of Agricultural, Environmental and Food Sciences, Free University of Bolzano-Bozen, Piazza Università 5, Bolzano, Italy
[b] Laimburg Research Centre, Laimburg 6, Pfatten (Vadena), 39040 Auer, Italy
[c] Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1, 20126 Milano, Italy

ARTICLE INFO

Keywords:
Biomarker
Food authenticity
Milk
Nuclear magnetic resonance
Silage feeding

ABSTRACT

This study investigated the distinction between haymilk and milk from silage feeding using proton nuclear magnetic resonance (^{1}H NMR) spectroscopy, given that the use of silage in cow diet is not allowed in haymilk production as well as in several PDO cheese regulations. A total of 245 raw milk samples from three feeding systems – haymilk, with no fermented feed in the ration ($n = 49$), maize silage milk with maize silage as the only fermented feed ($n = 98$), and grass silage milk with grass silage as the only fermented feed (n = 98) – collected between 2019 and 2021 in South Tyrol (Northern Italy) were analyzed using untargeted ^{1}H NMR of both aqueous and lipid fractions. PCA showed clustering of milk samples by season and sampling period in the aqueous ^{1}H NMR spectra, especially for haymilk, whereas lipid spectra showed no such clustering, indicating stable fatty acid profiles across seasons and years of collection. Untargeted ^{1}H NMR metabolomics combined with multivariate supervised classification based on milk lipid profile showed a good discriminant capacity between haymilk and milk from silage feeding, achieving sensitivity and specificity for haymilk samples equal to 87% and 94%, respectively. Furthermore, the proposed modelling approach enabled discrimination between milk from maize silage and grass silage feedings. Conjugated linoleic acid and cyclopropane fatty acids were identified as potential biomarkers to discriminate among the three milk groups. Overall, these results demonstrate that ^{1}H NMR lipid metabolomics represents a robust approach for the authentication of haymilk for the presence of silage in the ration.

1. Introduction

Feeding system is one of the major factors influencing milk composition, together with breed, season, and stage of lactation, all of which affect milk nutritional and technological quality (Magan et al., 2021). This effect is particularly pronounced when comparing pasture-based Alpine milk production with silage and hay-based lowland milk production systems (Tata et al., 2022). Increased levels of beneficial unsaturated fatty acids, such as oleic acid and α-linolenic acid, as well as conjugated linoleic acid (CLA), have been reported in milk from cows grazing on Alpine summer pastures compared to milk obtained from cows fed indoor with a silage-based diet in the lowland (Roda et al., 2015), as well as when comparing hay-based diets against silage-based diets (Oever et al., 2021; Staszak, 2005). These long-chain unsaturated fatty acids, especially CLA, have been associated with potential health-promoting effects in humans related to cardiovascular disease, cancer, infant development etc. (Butler et al., 2011; Staszak, 2005; Tsiafoulis et al., 2019). In parallel, consumer demand for grassland-based labelled dairy products has increased, often willing to pay a premium price for such products (Rojas-Gómez et al., 2025).

Haymilk is a high value-added alpine dairy product, for which hay farming has been a traditional practice for centuries. In recognition of its agricultural heritage, the European Union registered haymilk as "Traditional Specialties Guaranteed" (TSG) in 2016. This designation ensures strict production specifications, particularly in feed management. According to haymilk production guidelines, the use of fermented fodder, such as silage, is not allowed as well as genetically modified feeds (Commission Implementing Regulation (EU) 2016/304, 2016).

* Corresponding author.
E-mail address: ksenia.morozova@unibz.it (K. Morozova).

https://doi.org/10.1016/j.foodres.2026.119495
Received 3 March 2026; Received in revised form 18 May 2026; Accepted 19 May 2026
Available online 23 May 2026

Moreover, the roughage proportion must be at least 75% of the annual dry feed ratio (Commission Implementing Regulation (EU) 2016/304, 2016). Consequently, the use of haymilk as silage-free milk has been established in the production of several Protected Designation of Origin (PDO) hard cheeses by the EU regulations (Commission Implementing Regulation (EU) 2016/304, 2016)).

Given the nutritional, technological, and regulatory importance of feeding composition in dairy systems, robust analytical methods are required for the quality control and authenticity of milk produced according to haymilk standards. Nuclear magnetic resonance spectroscopy (NMR) has emerged as a promising technique for food authenticity, as it is rapid, non-destructive and does not require sample separation or chemical derivatization (Eltemur, Robatscher, Oberhuber, Scampicchio, & Ceccon, 2023; Sobolev et al., 2017; Wishart, 2019). Its untargeted nature enables the identification of a wide range of milk metabolites, especially low-molecular weight compounds, which may be used as potential biomarkers for milk quality control (Sundekilde et al., 2013). Untargeted NMR metabolomics typically requires the application of multivariate statistical analyses to process and interpret the vast and complex generated dataset. NMR-based metabolomics combined with chemometrics has been widely applied in milk authenticity studies to discriminate milk according to species (Garcia et al., 2012; Wei et al., 2022), geographical origin (Renou et al., 2004; Sacco et al., 2009; Tenori et al., 2018), feeding systems (Lanza et al., 2021; O'Callaghan et al., 2018), organic milk production (Erich et al., 2015; Phuenpong et al., 2021; Tsiafoulis et al., 2019), as well as milk adulteration (Cui et al., 2019; Lachenmeier et al., 2009; Li et al., 2017).

The aim of the present study is to discriminate milk based on cow feeding systems using untargeted ^{1}H NMR metabolomics. Both the lipid and polar fractions of 245 milk samples from South Tyrol (Northern Italy) were analyzed to obtain a complete metabolic fingerprint of haymilk, given that the polar fraction has been less investigated than fatty acid composition in the studies on the effects of different feeding systems (i.e., based on silage, hay and fresh herbage) on milk composition (Lanza et al., 2021; Renou et al., 2004). Notably, dietary differences are mainly reflected in milk fatty acid profiles (Wölk et al., 2021). Chemometric analysis was performed using a hierarchical classification approach, first discriminating haymilk from milk from silage feeding and subsequently differentiating between maize silage and grass silage feedings. This approach allows assessing the potential of ^{1}H NMR-based metabolomics as a reliable tool to support the authentication of haymilk by verifying compliance with their production standards based on feeding practices.

2. Materials and methods

2.1. Chemicals and reagents

Deuterium oxide (99.9%) and sodium 3-(trimethylsilyl) propionate-2,2,3,3-d4 (TSP) was obtained from Sigma Aldrich (Saint Louis, MO, USA). Deuterated chloroform ($CDCl_3$, 0.03% v/v tetramethylsilane (TMS) as an internal standard, CAS no. 865–49-6, >98% D) was purchased from Sigma-Aldrich (Merk Life Science S.r.l., Milano, Italy). All solvents and reagents were of analytical grade.

2.2. Milk samples

Bulk bovine milk samples were collected from 29 dairy farms located in South Tyrol, Northern Italy, spanning altitudes ranging from 542 to 1470 m above sea level. Sample collection was distributed in four different periods: winter 2019 (from the end of October to the end of May), summer 2020 (from the beginning of June to the beginning of September), winter 2020 (from the middle of November to the middle of March), and summer 2021 (from the middle of May to the end of July). Sampling was repeated on a weekly basis over three consecutive weeks (only in two cases a fourth sampling week was added), with samples collected in 50 mL of aliquots directly from the bulk tank of each farm. Cow feeding data was recorded for each farm and sampling period. Sampling plan is explained in detail by Fava et al. (2025).

A total of 245 collected raw milk samples were categorized into three groups based on the different types of forage in the cow's diet: 98 milk samples from diets including maize silage with no other fermented feed (maize silage milk, MSM), 98 including grass silage with no other fermented feed (grass silage milk, GSM), and 49 haymilk (HM) samples. As a control, haymilk samples were selected from the farms that fed cows according to the EU regulation (Commission Implementing Regulation (EU) 2016/304, 2016) for hay milk production. A complete list of the investigated samples including the respective feeding regime (hay, maize silage and grass silage), the silage type fed (maize, grass and a mixture of maize and grass), the season, the sampling period and the sampling week are reported in Table S1. HM samples were collected from 4 farms, MSM from 11 farms, and GSM from 14 farms. Milk samples from cows fed a mixture of maize and grass silage (7 observations) were assigned to the maize silage milk. The samples were kept refrigerated at 4 °C during transportation and later stored at −80 °C until the time of analysis.

2.3. Freeze-drying of the milk samples

Freeze-drying of milk samples was performed using the Epsilon 2-6D LSC plus laboratory freeze-dryer (Martin Christ, Osterode, Germany) containing a double stage, rotary vane vacuum pump with a condenser capacity of 6 kg. The vacuum chamber volume of 0.5 m^3 is equipped with four shelves that can be cooled down. The vacuum cooler includes pressure and temperature sensors for monitoring. Temperature monitoring inside the freeze-dryer utilized four type-T thermocouples and three wireless sensors (WTMplus, Germany). Pressure inside the chamber was monitored using two vacuum sensors: a Pirani sensor (Thyracont VSP63, Thyracont Vacuum Instruments GmbH, Passau, Germany) and a capacitive sensor (MKS 722B, Andover, MA USA). Pressure and temperature data were recorded using a LyoLog plus software (Martin Christ, Osterode, Germany).

Milk samples were thawed overnight in the refrigerator at 4 °C prior to the freeze-drying process. The freeze-drying followed these temperature settings: the initial freeze-drying step of the samples for 15 h at a starting temperature of −40 °C in a vacuum set at 0.05 mbar. The second and third steps lasted for 10 h each, at the temperatures of −30 and −10 °C, respectively, under the same vacuum conditions. The final drying phases were conducted for 4 h at 5 °C and 20 h at 15 °C.

2.4. ^{1}H NMR analysis

Milk samples were analyzed by ^{1}H NMR on both lipid and aqueous fractions. For the NMR analysis on milk lipid fraction, 50 mg of freeze-dried milk samples, were mixed with 1.0 mL of CDCl3 (containing 0.03% v/v TMS, as an internal standard). The mixture was shaken for a few minutes and the supernatant transferred into 5 mm NMR tubes for ^{1}H NMR analysis. For the NMR analysis on milk aqueous fraction, 50 mg of freeze-dried milk samples were dissolved in 1.0 mL of D_2O (containing 0.5% w/v TSP, as an internal standard) applying the same procedure to prepare the samples for NMR analysis.

^{1}H NMR spectra were recorded using a 600 MHz spectrometer (JNM-ECZ from JEOL Ltd., Tokyo, Japan), equipped with a room-temperature "Royal" HFX/FGSQ probe and an autosampler cooled at 4 °C. All 1D ^{1}H NMR experiments were performed using a slight in-house modification of the pulse sequence "qnmr_experiment" (JEOL Ltd.). All the spectra were acquired at 25 °C, with 32,000 complex points, using a 45° pulse length and a relaxation delay of 5 s. A total of 1024 scans were acquired with a spectral width of 15.024 kHz and an acquisition time of 1.707 s. The NMR spectra were processed and analyzed with Delta NMR Data Processing software (JEOL Ltd., Tokyo, Japan). 64 K data points were collected, and the FIDs were treated using a line broadening exponential

function of 0.2 Hz. The spectra were automatically processed with a batch processing template using automatic phase and baseline corrections, alignment of the spectra to the internal standard: the TSP signal at 0 ppm for aqueous (D_2O as solvent) samples and the TMS signal at 0 ppm for lipidic ($CDCl_3$ as solvent) samples.

2.5. Data preprocessing

^{1}H NMR spectra were preprocessed by automatic bucketing (binning) using Delta NMR Data Processing software (JEOL Ltd., Tokyo, Japan). Spectral normalization was applied to total spectral integral and reduced to integrating regions to bin width of 0.02 ppm within the spectral region of 9.0 to −1.00 ppm. ^{1}H NMR data were further divided into two separate datasets corresponding to the aqueous (polar) and lipid (non-polar) fractions for subsequent multivariate analysis.

Metabolites in the milk profiles were assigned in the ^{1}H NMR spectra by using authentic standards of cyclopropane fatty acid (all-*cis*-methyleneoctadecanoic acid, microbial cyclopropane fatty acid), Cas no. 5711-28-4, neat purity 98%, Abcam, Cambridge, UK) and conjugated linoleic acid (CLA, octadecadienoic acid, conjugated, Sigma Aldrich, ST. Louis, USA). All assignments were further confirmed by using the Chenomx built-in library ver. 10.0, the Human Metabolome Database (https://hmdb.ca), and previous references (Eltemur, Robatscher, Oberhuber, & Ceccon, 2023; Erich et al., 2015; Knothe, 2006; Scano et al., 2011; Tsiafoulis et al., 2019). The CPFA contents in samples were determined in positive samples referring to values > limit of detection (230 mg kg^{-1} fat) (Iannone et al., 2024), while for CLA based on the signal-to-noise ratio ≥ 10 on Delta Software (JEOL Ltd.), corresponding to the limit of quantification ($LOQ = \frac{10C}{SNR}$), where C and SNR are the observed spin molar concentration and the signal-to-noise ratio) (Olson et al., 1995).

2.6. Multivariate analysis

Principal Component Analysis (PCA) was initially used to explore the NMR data structure. This benchmark method consists of projecting the data into a reduced hyperspace defined by the principal components (PCs), which are linear combinations of the original variables. Partial Least Squares Discriminant Analysis (PLS-DA) was employed as a multivariate classification technique to discriminate between different feedings. PLS-DA is widely regarded as one of the most effective linear classification methods: it identifies latent variables (LVs) that exhibit maximum covariance with the class labels, which are encoded in the form of a binary dummy matrix, and the predictions are quantitative, yielding values between 0 and 1 (Pérez et al., 2009). The optimal number of latent variables was determined through internal cross-validation, minimizing the classification error. NMR spectra were preprocessed by applying the first derivative and autoscaling.

Since NMR spectra of milk samples with different forage in the cows' diet showed no evident differences, overlap between classes was expected. In this framework, hierarchical models offer an effective alternative to classical classification approaches by gradually separating classes, one or a few at a time (Marchi et al., 2022). This approach can simplify the discrimination task, particularly when dealing with highly overlapping classes. To improve the discrimination of milk samples based on feeding, a PLS-DA-based hierarchical approach was therefore employed. Specifically, two independent models were developed using PLS-DA (Wold et al., 2001): model A was trained to discriminate haymilk from non-haymilk samples, that are samples belonging to the MSM and GSM classes, while model B was trained to differentiate between MSM and GSM. The prediction of a target sample was then carried out with the following hierarchical approach: initially, the target sample was predicted with model A as haymilk or non-haymilk. In the latter case, the sample was subsequently predicted with model B and therefore labelled as belonging to the MSM or GSM classes. Hierarchical models were constructed using ad-hoc MATLAB functions, while the PLS-DA models were implemented using the Classification Toolbox for MATLAB (Ballabio & Consonni, 2013).

2.7. Validation of classification models

The hierarchical classification approach was validated using a double-validation procedure (Ballabio, Robotti, et al., 2018; Varmuza & Filzmoser, 2016), which was repeated iteratively 1000 times. In each iteration, the dataset was first randomly divided into temporary test (20%) and training (80%) sets by maintaining the class proportion. The training set was used to build Model A to discriminate between "haymilk" and "non-haymilk" classes. Then, all "non-haymilk" samples were used to build Model B and to discriminate between MSM and GSM classes. The optimal number of latent variables (LVs) for each model was determined through internal cross-validation of the training samples. Importantly, the test-set samples were never used during model training or in the optimization of the number of LVs for either PLS-DA model. Afterwards, test samples were predicted with Model A and Model B and labelled according to the hierarchical classification approach such as haymilk, MSM or GSM.

Model discrimination capability was assessed using sensitivity and specificity as performance metrics (Ballabio, Grisoni, & Todeschini, 2018). Sensitivity for a given class is defined as the percentage of samples correctly classified as belonging to that class, while specificity refers to the percentage of samples from other classes that are correctly identified as not belonging to the class.

3. Results and discussion

3.1. Effect of season on the milk NMR metabolomic profile

Principal Component Analysis (PCA) was initially applied to the full ^{1}H NMR spectra information, in both aqueous and lipid fractions of milk samples, to check variability of milk composition across season (summer, $n = 130$ and winter, $n = 115$) and sampling period (2019–2020, 2020, 2020–2021 and 2021). PCA was carried out independently for haymilk, MSM, and GSM groups. The score plots for the first two principal components (PCs) based on the aqueous fraction of the milk samples revealed a clustering of the samples according to the season and collection period (Fig. 1). This trend is particularly evident for haymilk samples. Haymilk samples collected during the summer of 2020 (negative scores in PC1) are different to haymilk samples collected during summer 2021 (negative scores in PC2) and to samples collected during winter in the 2019–2020 and 2020–2021 campaigns (positive scores in PC2). These findings may suggest that milk aqueous composition is influenced by the season and the collection period of milk.

The effect of seasonality on the milk composition has been widely studied using NMR spectroscopy (O'Callaghan et al., 2016; Rojas-Gómez et al., 2025). Seasonal variability in the milk polar metabolomic profile is generally attributed to changes in the living environment, diet, and metabolic activity of the animals across seasons (Tenori et al., 2018). This variability could also be affected by the inclusion of grazing or fresh herbage in the diet, which contributes to changes in milk chemical composition (Cabiddu et al., 2022; Niero et al., 2022).

On the contrary, a clustering trend regarding season and period of collection was not observed in the analysis of milk samples based on lipid fractions (Fig. S1).

3.2. Feeding systems discrimination based on milk lipid fraction

Lipid fractions of milk samples have been investigated to differentiate between haymilk and non-haymilk samples. The original dataset consisted of 245 samples, of which 19 were considered outliers on the basis of a preliminary analysis with PCA and looking at Hotelling's T^2 and Q residuals (Bro & Smilde, 2014), as shown in Fig. S2. The remaining 226 milk samples were labelled as haymilk (44 samples),

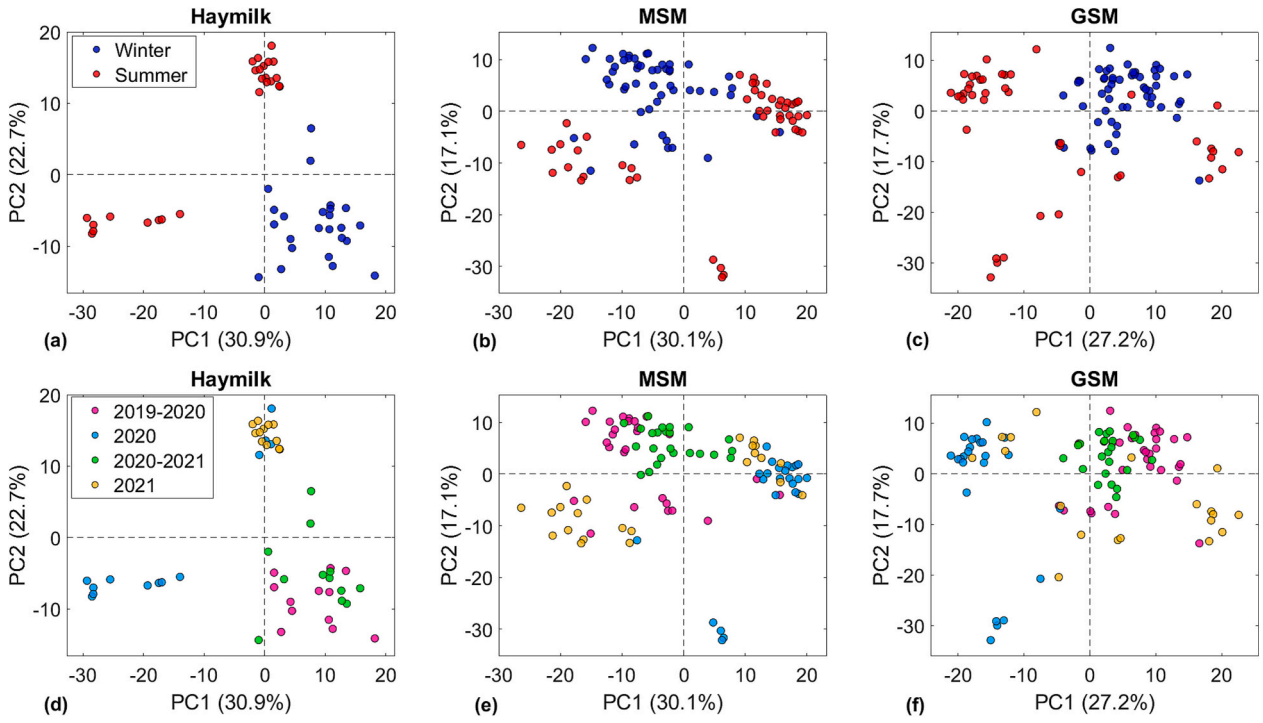

Fig. 1. Principal component analysis (PCA) score plots of the aqueous milk fraction for haymilk samples (a, d), MSM samples (b, e), and GSM samples (c, f). Samples are colored according to the season (winter: dark blue; summer: red) or according to the sampling campaign (2019–2020: pink; 2020: light blue; 2020–2021: green; 2021: yellow). Haymilk: production according to the EU regulation 2016/304; MSM: maize silage milk; GSM: grass silage milk. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

MSM (94 samples), and GSM (88 samples). The hierarchical PLS-DA modelling was validated through 1000 iterations following the double validation procedure previously described. Predictions of test samples were used to estimate the predictive capability of the hierarchical classification procedure. Sensitivity and specificity for haymilk, MSM, and GSM classes for test samples are reported in Table 1.

The hierarchical modelling resulted in a promising sensitivity and specificity for haymilk samples equal to 87% and 94%, respectively. These results indicate a good capability of discriminating haymilk against non-haymilk (silage-based feeding), as can be seen in the score plot of LV 1 and 3 (Fig. 2). Similarly, Tenori et al. (2018) distinguished milk from farms where cows were fed with silage and without silage using fatty acid ^{1}H NMR fingerprinting. Moreover, O'Callaghan et al. (2018) reported that NMR-based metabolomics of fatty acid profile of milk can be used to discriminate pasture-based feeding from silage feeding.

Sensitivity (74%, 71%) and specificity (80%, 82%) when discriminating MSM and GSM classes are lower than those observed for the haymilk class (Table 1); however, the classification performance can be considered satisfactory, given the greater overlap between MSM and GSM samples. This is evident from the latent variable score plot of Model A, which differentiates haymilk and non-haymilk samples (Fig. 2a). When considering Model B (Fig. 2b), the MSM and GSM classes exhibit a slightly higher degree of overlap, despite maintaining good discriminatory performance.

Table 1
Classification measures for test milk samples from three feeding groups (haymilk, maize silage milk, and grass silage milk) based on the ^{1}H NMR lipid profiles.

	Haymilk	Maize silage milk	Grass silage milk
Sensitivity (%)	87	74	71
Specificity (%)	94	80	82

3.3. NMR features for discrimination of samples based on lipid fraction

When modelling NMR data by means of PLS-DA, spectra were preprocessed by means of first derivative and autoscaling for row and column scaling, respectively. Therefore, the application of these preprocessing methods leads to the attenuation of information related to signal intensity, while enhancing the contribution of low-intensity signals that may otherwise be masked. This approach is particularly beneficial in cases where specific peaks are not available to discriminate between classes. For instance, the average spectra of haymilk, MSM, and GSM may be indistinguishable when examined solely based on the most intense peaks. However, after the first derivative transformation, previously masked differences due to low-intensity signals become more evident and can be leveraged for discrimination.

A potential limitation of this method is the increased susceptibility of the models to rely on signals associated with noise, highlighting the importance of considering the instrument's limit of detection (LOD). To address this, we identified the most influential variables (ppm ranges in the ^{1}H NMR spectra) used for class discrimination by looking at most important PLSDA coefficients combined with manual comparison of NMR spectra for discrimination among milk groups. Some differences in intensities between the three milk groups were observed within the following spectral region: $\delta = -0.34$ ppm, $\delta = 0.55$ ppm and $\delta = 0.64$ ppm, $\delta = 0.88$ ppm, $\delta = 2.01$ ppm, $\delta = 2.80$ ppm, $\delta = 5.65$ ppm, $\delta = 5.80$ ppm, $\delta = 5.93$ ppm and $\delta = 6.28$ ppm (Fig. 3, Fig. S3).

The signals at −0.34 ppm, 0.55 ppm and 0.64 ppm have been assigned to *cis*-methylene proton, *trans*-methylene proton, and to the two trans-methine protons of the cyclopropane ring of the cyclopropane fatty acids (CPFAs), respectively (Eltemur, Robatscher, Oberhuber, & Ceccon, 2023; Knothe, 2006; Lolli et al., 2018). Further confirmation of the above assignments was achieved with ^{1}H–^{13}C HSQC NMR experiments of the CPFA standard (Fig. S4). Milks from silage feeding (MSM and GSM) revealed to have more intensive signals of CPFAs compared to haymilk. Additionally, MSM samples showed higher levels

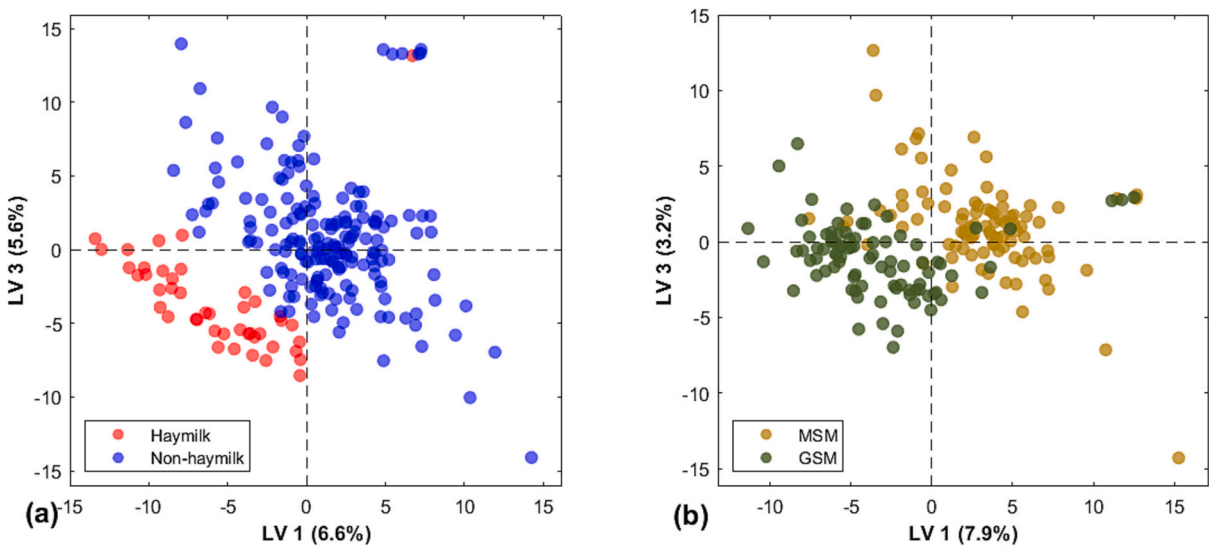

Fig. 2. ^{1}H NMR lipid fraction dataset: scores plot of the LVs 1 and 3 of: PLS-DA model A for differentiating haymilk and non-haymilk (a) and PLS-DA model B for differentiating MSM and GSM (b). Haymilk: production according to the EU regulation 2016/304; MSM: maize silage milk; GSM: grass silage milk.

of CPFAs with a range from 244.4 to 837.2 mg kg^{-1} of fat (mean: 301.0 mg kg^{-1} fat) compared to GSM samples, CPFAs content of which ranged widely from 0.0 to 516.8 mg kg^{-1} of fat (mean:198.1 mg kg^{-1} fat), with 77% of the samples resulting positive for CPFAs. Our research group investigated the authenticity of haymilk (silage-free) through the detection of cyclopropane fatty acid analysis using both ^{1}H NMR and GC–MS analyses performed on the same sample pool. The results demonstrated that haymilk samples do not showed detectable levels of CPFAs, whereas, MSM and GSM contained CPFAs (Fava et al., 2025; Iannone et al., 2024; Imperiale et al., 2021; Riccio et al., 2025). Indeed, CPFAs are known molecular markers for the detection of silage feeding in cow's diet, as they are primarily formed during silage fermentation through bacterial metabolism.

Moreover, the signals with important discriminate role among the milk groups at 0.88 ppm, at 2.01 ppm, at 2.80 ppm and 5.80 ppm are attributed to CH_3-ω1 group (all FA except n-3 and butyric), allyl methylene group of unsaturated FA, bis allyl methylene groups of α-linolenic acid and H9 of caproleic acid, respectively (Erich et al., 2015; Scano et al., 2011; Tsiafoulis et al., 2014; Tsiafoulis et al., 2019). Additionally, the signals at 5.65 ppm, 5.93 ppm and 6.28 ppm have been assigned to olefinic protons of conjugated (9-cis, 11-trans) 18:2 linoleic acid (CLA), which is the major conjugated linoleic acid present in milk lipid fraction (Tsiafoulis et al., 2014; Tsiafoulis et al., 2019). Further confirmation of the above assignments was achieved with ^{1}H NMR and 2D ^{1}H–^{13}C HSQC experiments performed on CLA standard (Fig. S5). In all cases, NMR spectra of milk samples from haymilk showed the highest signal intensities in the above-mentioned regions, followed by milk samples from GSM and MSM. Correspondingly, CLA contents were 6.26 ± 2.3 mg kg^{-1} fat in haymilk, 6.13 ± 2.6 mg kg^{-1} fat in GSM and 5.45 ± 2.3 mg kg^{-1} fat in MSM.

These findings are in accordance with Staszak (2005) who reported higher CLA contents in milk from cows with hay-based diet compared to milk from silage-fed (maize silage) cows. However, Collomb et al. (2008) found that hay feeding supplemented with grass silage (with 27% grass silage combined with 4% maize silage) increased CLA content in milk by 12.0% compared to silage-free feeding in milk from cows in mountain areas of Switzerland. Indeed, previous studies have linked variations in α-linolenic acid and CLA contents in milk to the inclusion of grazed herbage in the cow diet, which positively influence such fatty acids (O'Callaghan et al., 2016; Roda et al., 2015; Villeneuve et al., 2013).

3.4. *Feeding system discrimination based on milk aqueous fraction*

The seasonal variabilities observed on milk polar fraction revealed challenges to differentiate milk samples based on feeding systems. Following the identification and removal of outliers via a preliminary PCA (utilizing Hotelling's T^2 and Q residuals metrics, as shown in Fig. S2), the PLS-DA model was calibrated on a final set of 234 samples. Fig. 4 shows the PLS-DA score plots used to discriminate haymilk from non-haymilk, and MSM from GSM, in which a clear overlap between classes is observed.

This is also confirmed by the results achieved through the double validation procedure: sensitivity and specificity for haymilk, MSM, and GSM classes for test samples are reported in Table 2. The model showed a low specificity, with 64% of non-haymilk samples incorrectly classified as haymilk, as well as low specificity values also for MSM and GSM.

Consistence with these findings, a similar study by Lanza et al. (2021) reported that milk NMR polar fraction was less sensitive to differences among milk samples obtained from three feeding regimens (maize silage, grass-legume and maize silage, and hay) compared to the fatty acid profile, showing limited discriminant capacity in distinguishing between milk from hay-fed and silage-fed cows.

4. Conclusion

This study applied untargeted ^{1}H NMR metabolomics to differentiate milk samples from three different feeding systems: haymilk, maize silage milk (MSM), and grass silage milk (GSM). Samples were analyzed with both polar and lipid ^{1}H NMR metabolomic profiles. Milk polar metabolome showed strong seasonal and interannual variability, especially in haymilk samples. Its limited sensitivity to dietary differences restricted its use for feeding system discrimination. In contrast, ^{1}H NMR analysis of the milk lipid fraction, combined with multivariate analysis enabled robust classification between haymilk and non-haymilk (silage feeding), achieving correct classification rates of 87% for haymilk and 94% for silage-derived milk samples. Additionally, the hierarchical modelling approach further discriminated between MSM and GSM.

Conjugated 18:2 linoleic acid (CLA) and cyclopropane fatty acids (CPFAs) were found to have influential discriminant roles between the three milk groups. Haymilk samples showed higher CLA content compared to GSM and MSM. In contrast, CPFAs were detected only in MSM and GSM compared to haymilk samples. Therefore, CLA and CPFAs emerged as reliable biomarkers in this untargeted ^{1}H NMR

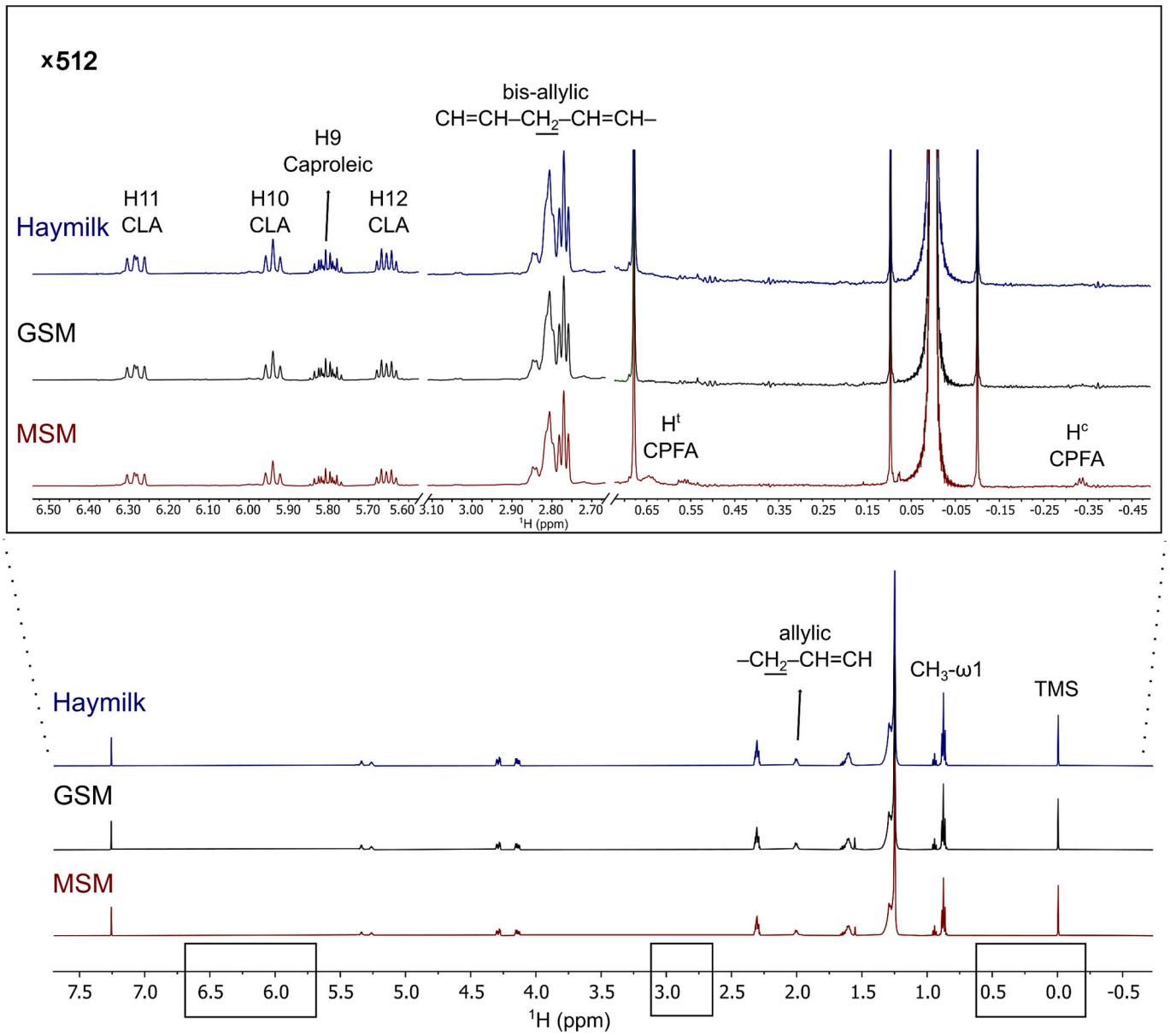

Fig. 3. 600 MHz ^{1}H NMR average spectra of the lipid fraction of lyophilized milk samples from haymilk (n = 49, blue color), GSM (n = 98, black color) and MSM (n = 98, red color) in $CDCl_3$. Haymilk: production according to the EU regulation 2016/304; MSM: maize silage milk; GSM: grass silage milk. The signals discriminating between the milk groups are denoted. The insert shows ×512 magnification of the spectra to display conjugated linoleic acid (CLA) and caproleic acid in the region between 5.60 and 6.50 ppm, bis allyl methylene groups of α-linolenic acid in the region between 3.10 and 2.70 ppm, and cyclopropane fatty acid (CPFA) in the region between 0.65 and −0.45 ppm. The spectra are normalized to TMS internal standard. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

metabolomics approach to discriminate between haymilk and milk from silage feeding. Indeed, CPFAs are already well-established targeted biomarkers for silage feeding in dairy products. Overall, these findings highlight the lipid ^{1}H NMR metabolomic profile as a reliable and informative medium for feeding system differentiation and support the use of ^{1}H NMR-based untargeted metabolomics combined with chemometrics as a promising tool for authenticity assessment of haymilk.

CRediT authorship contribution statement

Dilek Eltemur: Writing – original draft, Visualization, Investigation, Formal analysis, Data curation. **Enmanuel Cruz Muñoz:** Writing – review & editing, Visualization, Software, Formal analysis, Data curation. **Davide Ballabio:** Writing – review & editing, Visualization, Software, Formal analysis, Data curation. **Ksenia Morozova:** Writing – review & editing, Visualization, Supervision, Formal analysis, Data curation. **Demian Martini-Lösch:** Investigation, Data curation. **Giovanni Peratoner:** Writing – review & editing, Data curation, Conceptualization. **Daniela Eisenstecken:** Writing – review & editing, Project administration, Conceptualization. **Elena Venir:** Writing – review & editing, Project administration, Conceptualization. **Peter Robatscher:** Writing – review & editing, Supervision. **Michael Oberhuber:** Writing – review & editing, Funding acquisition. **Matteo Scampicchio:** Writing – review & editing, Visualization, Supervision, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

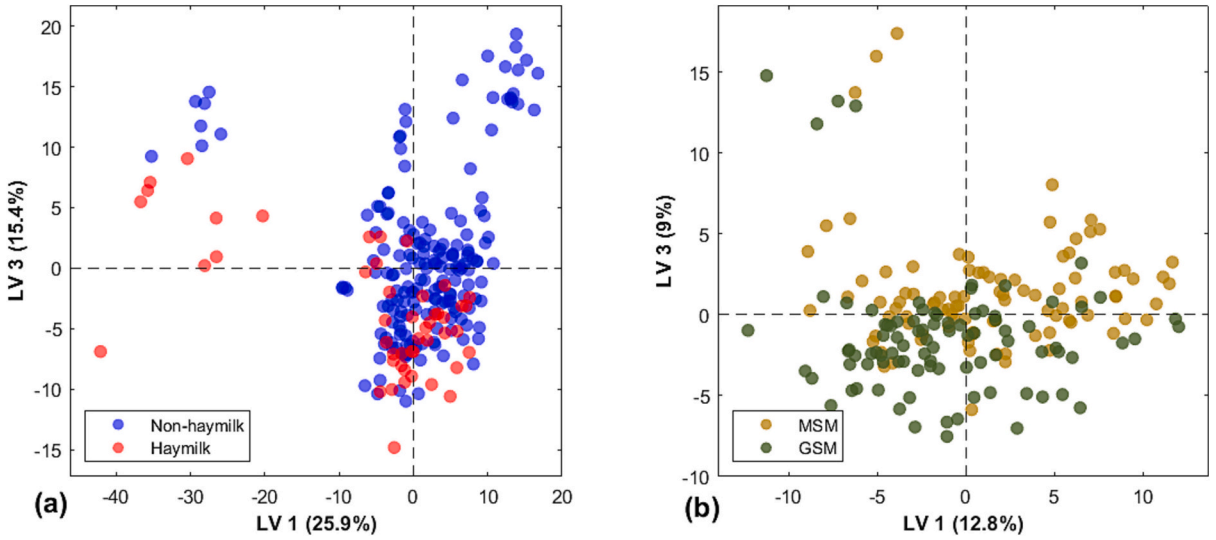

Fig. 4. NMR aqueous fraction dataset: scores plot of the LVs 1 and 3 of (a) PLS-DA model A for differentiating haymilk and non-haymilk and (b) PLS-DA model B for differentiating between MSM and GSM. Haymilk: production according to the EU regulation 2016/304; MSM: maize silage milk; GSM: grass silage milk.

Table 2
Classification measures for test milk samples from three feeding groups (haymilk, maize silage milk and grass silage milk) based on the ^{1}H NMR aqueous metabolomic profiles.

	Haymilk	Maize silage milk	Grass silage milk
Sensitivity (%)	86	70	73
Specificity (%)	64	58	52

the work reported in this paper.

Acknowledgments

The authors would like to thank the Sennereiverband Südtirol and the BRING Beratungsring Berglandwirtschaft for their valuable support during sampling.

This research is funded by the FESR-EFRE 2014–2020 "Investitionen in Wachstum und Beschäftigung", project Heumilch, FESR1129 CUP: H36H19000000007. The authors would like to thank the Autonomous Province of Bozen-Bolzano for its financial support through the NOI Capacity building I and II funding frame (Department of Innovation, Research and University, Decision n. 1472, 07.10.2013; Decision n. 864, 04.09.2018) and within the Action Plan 2016–2022 for Research and Training in the Fields of Mountain Agriculture and Food Science of the Autonomous Province of Bozen/Bolzano (Decision n. 1016, 01.09.2015). The PhD thesis of D.E. is financed by Laimburg Research Centre (LCH-nmr-23-1). Laimburg Research Centre is funded by the Autonomous Province of Bolzano.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2026.119495.

Data availability

Data will be made available on request.

References

Ballabio, D., & Consonni, V. (2013). Classification tools in chemistry. Part 1: Linear models. PLS-DA. *Analytical Methods, 5*(16), 3790. https://doi.org/10.1039/c3ay40582f

Ballabio, D., Grisoni, F., & Todeschini, R. (2018). Multivariate comparison of classification performance measures. *Chemometrics and Intelligent Laboratory Systems, 174*, 33–44. https://doi.org/10.1016/j.chemolab.2017.12.004

Ballabio, D., Robotti, E., Grisoni, F., Quasso, F., Bobba, M., Vercelli, S., Gosetti, F., Calabrese, G., Sangiorgi, E., Orlandi, M., & Marengo, E. (2018). Chemical profiling and multivariate data fusion methods for the identification of the botanical origin of honey. *Food Chemistry, 266*, 79–89. https://doi.org/10.1016/j.foodchem.2018.05.084

Bro, R., & Smilde, A. K. (2014). Principal component analysis. *Analytical Methods, 6*(9), 2812–2831. https://doi.org/10.1039/c3ay41907j

Butler, G., Stergiadis, S., Seal, C., Eyre, M., & Leifert, C. (2011). Fat composition of organic and conventional retail milk in Northeast England. *Journal of Dairy Science, 94*(1), 24–36. https://doi.org/10.3168/jds.2010-3331

Cabiddu, A., Peratoner, G., Valenti, B., Monteils, V., Martin, B., & Coppa, M. (2022). A quantitative review of on-farm feeding practices to enhance the quality of grassland-based ruminant dairy and meat products. *Animal, 16*, Article 100375. https://doi.org/10.1016/j.animal.2021.100375

Collomb, M., Bisig, W., Bütikhofer, U., Sieber, R., Bregy, M., & Etter, L. (2008). Influence of supplementing hay with grass silage on the fatty acid composition of mountain milk. *ALP Science, 526*, 1–17.

Commission Implementing Regulation (EU) 2016/304 of 2 March 2016. (2016). *Entering a name in the register of traditional specialties guaranteed (Heumilch/Haymilk/Latte fieno/Lait de foin/Leche de heno (TSG))*.

Cui, J., Zhu, D., Su, M., Tan, D., Zhang, X., Jia, M., & Chen, G. (2019). The combined use of ^{1}H and 2D NMR-based metabolomics and chemometrics for non-targeted screening of biomarkers and identification of reconstituted milk. *Journal of the Science of Food and Agriculture, 99*(14), 6455–6461. https://doi.org/10.1002/jsfa.9924

Eltemur, D., Robatscher, P., Oberhuber, M., & Ceccon, A. (2023). Improved detection and quantification of cyclopropane fatty acids via Homonuclear decoupling double irradiation NMR methods. *ACS Omega, 8*(44), 41835–41843. https://doi.org/10.1021/acsomega.3c06538

Eltemur, D., Robatscher, P., Oberhuber, M., Scampicchio, M., & Ceccon, A. (2023). Applications of solution NMR spectroscopy in quality assessment and authentication of bovine Milk. *Foods, 12*(17), 3240. https://doi.org/10.3390/foods12173240

Erich, S., Schill, S., Annweiler, E., Waiblinger, H.-U., Kuballa, T., Lachenmeier, D. W., & Monakhova, Y. B. (2015). Combined chemometric analysis of 1H NMR, 13C NMR and stable isotope data to differentiate organic and conventional milk. *Food Chemistry, 188*, 1–7. https://doi.org/10.1016/j.foodchem.2015.04.118

Fava, F., Martini-Lösch, D., Peratoner, G., Robatscher, P., Matteazzi, A., Soini, E., … Venir, E. (2025). Effect of diet on CPFAs used as markers in Milk for the detection of silage in the ration of dairy cows. *Foods, 14*(3), 476. https://doi.org/10.3390/foods14030476

Garcia, C., Lutz, N. W., Confort-Gouny, S., Cozzone, P. J., Armand, M., & Bernard, M. (2012). Phospholipid fingerprints of milk from different mammalians determined by 31P NMR: Towards specific interest in human health. *Food Chemistry, 135*(3), 1777–1783. https://doi.org/10.1016/j.foodchem.2012.05.111

Iannone, F., Eltemur, D., Morozova, K., Fava, F., Martini-Lösch, D., Robatscher, P., Ferrentino, G., Asma, U., Peratoner, G., Venir, E., Eisenstecken, D., Oberhuber, M., & Scampicchio, M. (2024). Establishing authenticity of hay milk: Detection of silage feeding through cyclopropane fatty acids analysis using 1H NMR spectroscopy. *Food Chemistry, 438*, Article 138048. https://doi.org/10.1016/j.foodchem.2023.138048

Imperiale, S., Kaneppele, E., Morozova, K., Fava, F., Martini-Lösch, D., Robatscher, P., Peratoner, G., Venir, E., Eisenstecken, D., & Scampicchio, M. (2021). Authenticity of Hay Milk vs. milk from maize or grass silage by lipid analysis. *Foods, 10*(12), Article 2926. https://doi.org/10.3390/foods10122926

Knothe, G. (2006). NMR characterization of dihydrosterculic acid and its methyl ester. *Lipids, 41*(4), 393–396. https://doi.org/10.1007/s11745-006-5110-x

Lachenmeier, D. W., Humpfer, E., Fang, F., Schütz, B., Dvortsak, P., Sproll, C., & Spraul, M. (2009). NMR-spectroscopy for nontargeted screening and simultaneous quantification of health-relevant compounds in foods: The example of melamine. *Journal of Agricultural and Food Chemistry, 57*(16), 7194–7199. https://doi.org/10.1021/jf902038j

Lanza, I., Lolli, V., Segato, S., Caligiani, A., Contiero, B., Lotto, A., Galaverna, G., Magrin, L., & Cozzi, G. (2021). Use of GC–MS and 1H NMR low-level data fusion as an advanced and comprehensive metabolomic approach to discriminate milk from dairy chains based on different types of forage. *International Dairy Journal, 123*, Article 105174. https://doi.org/10.1016/j.idairyj.2021.105174

Li, Q., Yu, Z., Zhu, D., Meng, X., Pang, X., Liu, Y., Frew, R., Chen, H., & Chen, G. (2017). The application of NMR-based milk metabolite analysis in milk authenticity identification: NMR-based milk metabolite analysis. *Journal of the Science of Food and Agriculture, 97*(9), 2875–2882. https://doi.org/10.1002/jsfa.8118

Lolli, V., Marseglia, A., Palla, G., Zanardi, E., & Caligiani, A. (2018). Determination of cyclopropane fatty acids in food of animal origin by^{1} H NMR. *Journal of Analytical Methods in Chemistry, 2018*, 1–8. https://doi.org/10.1155/2018/8034042

Magan, J. B., O'Callaghan, T. F., Kelly, A. L., & McCarthy, N. A. (2021). Compositional and functional properties of milk and dairy products derived from cows fed pasture or concentrate-based diets. *Comprehensive Reviews in Food Science and Food Safety, 20*(3), 2769–2800. https://doi.org/10.1111/1541-4337.12751

Marchi, L., Krylov, I., Roginski, R. T., Wise, B., Di Donato, F., Nieto-Ortega, S., … Bro, R. (2022). Automatic hierarchical model builder. *Journal of Chemometrics, 36*(12). https://doi.org/10.1002/cem.3455

Niero, G., Meoni, G., Tenori, L., Luchinat, C., Visentin, G., Callegaro, S., Visentin, E., Cassandro, M., De Marchi, M., & Penasa, M. (2022). Grazing affects metabolic pattern of individual cow milk. *Journal of Dairy Science, 105*(12), 9702–9712. https://doi.org/10.3168/jds.2022-22072

O'Callaghan, T. F., Hennessy, D., McAuliffe, S., Kilcawley, K. N., O'Donovan, M., Dillon, P., … Stanton, C. (2016). Effect of pasture versus indoor feeding systems on raw milk composition and quality over an entire lactation. *Journal of Dairy Science, 99*(12), 9424–9440. https://doi.org/10.3168/jds.2016-10985

O'Callaghan, T. F., Vázquez-Fresno, R., Serra-Cayuela, A., Dong, E., Mandal, R., Hennessy, D., … Ross, R. (2018). Pasture feeding changes the bovine rumen and Milk metabolome. *Metabolites, 8*(2), 27. https://doi.org/10.3390/metabo8020027

Oever, S. P., Haselmann, A., Schreiner, M., Fuerst-Waltl, B., Zebeli, Q., Mayer, H. K., & Knaus, W. (2021). Hay versus silage: Does hay feeding positively affect milk composition? *International Dairy Journal, 118*. https://doi.org/10.1016/j.idairyj.2021.105024

Olson, D. L., Peck, T. L., Webb, A. G., Magin, R. L., & Sweedler, J. V. (1995). High-resolution Microcoil1 H-NMR for mass-limited, nanoliter-volume samples. *Science, 270*(5244), 1967–1970. https://doi.org/10.1126/science.270.5244.1967

Pérez, N. F., Ferré, J., & Boqué, R. (2009). Calculation of the reliability of classification in discriminant partial least-squares binary classification. *Chemometrics and Intelligent Laboratory Systems, 95*(2), 122–128. https://doi.org/10.1016/j.chemolab.2008.09.005

Phuenpong, T., Kongboonkird, M., Duangmal, K., Lerdvorasap, W., Suksawwawimon, M., Mekboonsonglarp, W., Nuamchit, J., Chantaprasarn, N., & Settachaimongkon, S. (2021). Molecular discrimination between organic and conventional liquid milk products in Thailand using 1H-NMR metabolomics approach. *Tropical Animal Science Journal, 44*(4), Article 4. https://doi.org/10.5398/tasj.2021.44.4.478

Renou, J.-P., Deponge, C., Gachon, P., Bonnefoy, J.-C., Coulon, J.-B., Garel, J.-P., Vérité, R., & Ritz, P. (2004). Characterization of animal products according to geographic origin and feeding diet using nuclear magnetic resonance and isotope ratio mass spectrometry: Cow milk. *Food Chemistry, 85*(1), Article 1. https://doi.org/10.1016/j.foodchem.2003.06.003

Riccio, G. M., Eltemur, D., Fava, F., Martini-Lösch, D., Peratoner, G., Venir, E., … Ceccon, A. (2025). Differentiating cyclopropane fatty acids to support milk authenticity through GC–MS and NMR spectroscopy. *Food Chemistry: X, 31*, Article 103033. https://doi.org/10.1016/j.fochx.2025.103033

Roda, G., Fialà, S., Vittorini, M., & Secundo, F. (2015). Fatty acid composition and fat content in milk from cows grazing in the Alpine region. *European Food Research and Technology, 241*(3), 413–418. https://doi.org/10.1007/s00217-015-2473-3

Rojas-Gómez, P., Pariyani, R., Bateman, L. M., Lynch, D., Timlin, M., Dineen, M., … O'Callaghan, T. F. (2025). Effect of proportion of pasture in the cow diet and seasonality on the milk metabolome as determined by 1H-NMR. *Journal of Dairy Science, 108*(5), 4659–4673. https://doi.org/10.3168/jds.2024-26168

Sacco, D., Brescia, M. A., Sgaramella, A., Casiello, G., Buccolieri, A., Ogrinc, N., & Sacco, A. (2009). Discrimination between Southern Italy and foreign milk samples using spectroscopic and analytical data. *Food Chemistry, 114*(4), Article 4. https://doi.org/10.1016/j.foodchem.2008.11.056

Scano, P., Anedda, R., Melis, M. P., Dessi', M. A., Lai, A., & Roggio, T. (2011). 1 H- and13 C-NMR characterization of the molecular components of the lipid fraction of Pecorino Sardo cheese. *Journal of the American Oil Chemists' Society, 88*(9), 1305–1316. https://doi.org/10.1007/s11746-011-1797-9

Sobolev, A. P., Circi, S., Capitani, D., Ingallina, C., & Mannina, L. (2017). Molecular fingerprinting of food authenticity. *Current Opinion in Food Science, 16*, 59–66. https://doi.org/10.1016/j.cofs.2017.08.002

Staszak, E. (2005). Conjugated linoleic acid (CLA) content of milk from cows fed different diets. *Folia Biologica, 53*(4), 103–106. https://doi.org/10.3409/173491605775789290

Sundekilde, U., Larsen, L., & Bertram, H. (2013). NMR-based milk metabolomics. *Metabolites, 3*(2), 204–222. https://doi.org/10.3390/metabo3020204

Tata, A., Massaro, A., Riuzzi, G., Lanza, I., Bragolusi, M., Negro, A., Novelli, E., Piro, R., Gottardo, F., & Segato, S. (2022). Ambient mass spectrometry for rapid authentication of milk from Alpine or lowland forage. *Scientific Reports, 12*(1), 7360. https://doi.org/10.1038/s41598-022-11178-9

Tenori, L., Santucci, C., Meoni, G., Morrocchi, V., Matteucci, G., & Luchinat, C. (2018). NMR metabolomic fingerprinting distinguishes milk from different farms. *Food Research International, 113*, 131–139. https://doi.org/10.1016/j.foodres.2018.06.066

Tsiafoulis, C. G., Papaemmanouil, C., Alivertis, D., Tzamaloukas, O., Miltiadou, D., Balayssac, S., … Gerothanassis, I. (2019). NMR-based metabolomics of the lipid fraction of organic and conventional bovine milk. *Molecules, 24*(6), Article 6. https://doi.org/10.3390/molecules24061067

Tsiafoulis, C. G., Skarlas, T., Tzamaloukas, O., Miltiadou, D., & Gerothanassis, I. P. (2014). Direct nuclear magnetic resonance identification and quantification of geometric isomers of conjugated linoleic acid in milk lipid fraction without derivatization steps: Overcoming sensitivity and resolution barriers. *Analytica Chimica Acta, 821*, 62–71. https://doi.org/10.1016/j.aca.2014.03.010

Varmuza, K., & Filzmoser, P. (2016). *Introduction to multivariate statistical analysis in Chemometrics* (0 ed.). CRC Press. https://doi.org/10.1201/9781420059496

Villeneuve, M.-P., Lebeuf, Y., Gervais, R., Tremblay, G. F., Vuillemard, J. C., Fortin, J., & Chouinard, P. Y. (2013). Milk volatile organic compounds and fatty acid profile in cows fed timothy as hay, pasture, or silage. *Journal of Dairy Science, 96*(11), 7181–7194. https://doi.org/10.3168/jds.2013-6785

Wei, W., Li, D., Jiang, C., Zhang, X., Zhang, X., Jin, Q., Zhang, X., & Wang, X. (2022). Phospholipid composition and fat globule structure II: Comparison of mammalian milk from five different species. *Food Chemistry, 388*, Article 132939. https://doi.org/10.1016/j.foodchem.2022.132939

Wishart, D. S. (2019). NMR metabolomics: A look ahead. *Journal of Magnetic Resonance, 306*, 155–161. https://doi.org/10.1016/j.jmr.2019.07.013

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*(2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

Wölk, M., Milkovska-Stamenova, S., Schröter, T., & Hoffmann, R. (2021). Influence of seasonal variation and processing on protein glycation and oxidation in regular and hay milk. *Food Chemistry, 337*, Article 127690. https://doi.org/10.1016/j.foodchem.2020.127690