Tracing the identity of Parmigiano Reggiano "Prodotto di Montagna - Progetto Territorio" cheese using NMR spectroscopy and multivariate data analysis

Analytica Chimica Acta 1278 (2023) 341761

Contents lists available at ScienceDirect

Analytica Chimica Acta

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

Tracing the identity of Parmigiano Reggiano "*Prodotto di Montagna - Progetto Territorio*" cheese using NMR spectroscopy and multivariate data analysis

N. Cavallini [a], L. Strani [b], P.P. Becchi [b,c], V. Pizzamiglio [d], S. Michelini [d], F. Savorani [a], M. Cocchi [b,*], C. Durante [b]

[a] *Department of Applied Science and Technology, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129, Torino, Italy*
[b] *Department of Chemical and Geological Sciences, Università di Modena e Reggio Emilia, Via Campi 103, 41125, Modena, Italy*
[c] *Department for Sustainable Food Process, Università Cattolica del Sacro Cuore, 29122, Piacenza, Italy*
[d] *Consorzio Formaggio Parmigiano Reggiano, via Kennedy 18, 42124, Reggio Emilia, Italy*

HIGHLIGHTS

- For the first time a NMR metabolomic characterization of Parmigiano Reggiano "Prodotto di Montagna - Progetto Territorio" was obtained.
- A MATLAB toolbox was developed allowing pure metabolites profiles determination and putative identification.
- It has been possible to highlight metabolites capable to differentiate Mountain Parmigiano Reggiano from conventional Parmigiano Reggiano PDO cheese.

ARTICLE INFO

Handling Editor: Prof. L. Buydens

Keywords:
Parmigiano Reggiano PDO cheese
Mountain identity
NMR
Chemometrics
Foodomics

ABSTRACT

Background: Nuclear magnetic resonance (NMR) spectroscopy is one of the well-established tools for food metabolomic analysis, as it proved to be very effective in authenticity and quality control of dairy products, as well as to follow product evolution during processing and storage. The analytical assessment of the EU mountain denomination label, specifically for Parmigiano Reggiano "Prodotto di Montagna - Progetto Territorio" (Mountain-CQ) cheese, has received limited attention. Although it was established in 2012 the EU mountain denomination label has not been much studied from an analytical point of view. Nonetheless, tracing a specific profile for the mountain products is essential to support the value chain of this specialty.

Results: The aim of the study was to produce an identity profile for Parmigiano Reggiano "Prodotto di Montagna - Progetto Territorio" (Mountain-CQ) cheese, and to differentiate it from Parmigiano Reggiano PDO samples (conventional-PDO) using ^{1}H NMR spectroscopy coupled with multivariate data analysis. Three different approaches were applied and compared. First, the spectra-as-such were analysed after proper preprocessing. For the other two approaches, Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) was used for signals resolution and features extraction, either individually on manually-defined spectral intervals or by reapplying MCR-ALS on the whole spectra with selectivity constraints using the reconstructed "pure profiles" as initial estimates and targets. All approaches provided comparable information regarding the samples' distribution, as in all three cases the separation between the two product categories conventional-PDO and Mountain-CQ could be highlighted. Moreover, a novel MATLAB toolbox for features extraction via MCR-ALS was developed and used in synergy with the Chenomx library, allowing for a putative identification of the selected features.

Significance: A first identity profile for Parmigiano Reggiano "Prodotto di Montagna - Progetto Territorio" obtained by interpreting the metabolites signals in NMR spectroscopy was obtained. Our workflow and toolbox for generating the features dataset allows a more straightforward interpretation of the results, to overcome the limitations due to dimensionality and to peaks overlapping, but also to include the signals assignment and matching since the early stages of the data processing and analysis.

* Corresponding author. Department of Chemical and Geological Sciences, University of Modena and Reggio Emilia, Via Campi 103, 41125, Modena, Italy.
E-mail address: marina.cocchi@unimore.it (M. Cocchi).

https://doi.org/10.1016/j.aca.2023.341761
Received 28 April 2023; Received in revised form 4 August 2023; Accepted 28 August 2023
Available online 28 August 2023

1. Introduction

EU Regulation 1152/2012 [1] introduced the indication "Mountain Product" to enhance the value of food products from mountain areas by making them immediately recognizable to consumers. Normally specific valence is attributed to these products, which are very often rewarded with greater willingness to buy; this may allow mountain producers an adequate income that could contribute to the permanence of agricultural activity in these disadvantaged areas and to the overall vitality of mountain areas. Set against these potential benefits is the issue of falsification and adulteration with respect to product identity and authenticity. Thus, the need to evaluate analytical methods capable of objectively identifying the identity characteristics of mountain products to ensure their origin and traceability. However, studies focusing on the mountain denomination are still limited [2–4].

In this study, a NMR-based foodomic approach was used, since NMR spectroscopy can detect low-molecular weight metabolites which may bring valuable information useful to reveal identity traits linked to the mountain denomination. Parmigiano Reggiano "Prodotto di Montagna - Progetto Territorio" [5] was selected as a Case of Study.

Parmigiano Reggiano cheese, with its unique and inimitable taste, was recognized as Protected Designation of Origin (PDO) food and it is appreciated all over the world. This name is usually associated with a single product idea, however different varieties can be found. In particular, the product "Prodotto di Montagna - Progetto Territorio" (Mountain-CQ in the text) represents a quality denomination for Parmigiano Reggiano cheese that in addition to the PDO also respect the mountain denomination [1] and must comply with additional rules, as detailed in the following, established by the producer consortium [5]. In particular, Mountain-CQ Parmigiano Reggiano cheese production insists in mountain areas, with over 30 dairies, located in the Apennines, that everyday help to strengthen the national economy and preserve the uniqueness of the mountain areas of Parma, Reggio Emilia, Modena and Bologna districts. The mountain certification must meet the following requirements: (i) 100% stables milk from mountain areas; (ii) more than 60% of the cows feed coming from the mountain area, while the additional specific requirements are: (iii) dairy and maturing up to 12 months; (iv) qualitative selection at 24 months with "hammer" evaluation of the Consortium's experts and (v) it should pass the sensory evaluation.

In this scenario, considering the production protocol and the difficulties in terms of production costs, lower yields, etc., there has been an increasing demand, from both dairy farmers and consortia, to safeguard the authenticity of Mountain-CQ Parmigiano Reggiano cheese from the analogues. In fact, this high added value product plays an important role in supporting the sustainability of the mountain areas where it is produced, offering revenue opportunities for the local economy. Furthermore, to promote the Mountain-CQ Parmigiano Reggiano cheese as a higher quality product (of high added value) and to protect its uniqueness require a great knowledge of the product itself.

Comprehensive analytical techniques can support this target by furnishing an objective assessment of quality and identity by extended characterization, as well as distinctive criteria. Among these techniques, high-resolution proton nuclear magnetic resonance (^{1}H NMR) spectroscopy has been widely used for a metabolic characterization [6–8]. The potentiality of this technique relies on obtaining a broad amount of information related to metabolites with a single analytical run [8] therefore resulting perfect for untargeted metabolomics. As a matter of fact, ^{1}H NMR can provide detailed information about the molecular structure, connectivity, and conformation of organic compounds, including the relative positions of atoms, presence of functional groups, being a valuable tool for biomarkers identification. With respect to other -omics analytical techniques (i.e., LC-MS, GC-MS) NMR spectroscopy is a more stable and reproducible technique, resulting more suitable for quantitative analysis, as it can determine the concentration of specific compounds within a mixture, even though, compared with MS, its sensitivity is definitively lower. In the food context, NMR spectroscopy has been widely used [9], just to cite a few applications concerned: the detection of water loss, uptake or migration of nutrients, protein denaturation and starch crystallisation [9–12], as well as identification of sugars, small organic acids, vitamins, nucleotides, and aromatic compounds [7,13–19]. Furthermore, NMR coupled with multivariate analysis have provided optimal results in food characterization and authenticity [20–22].

Concerning the dairy production chain, several NMR-based studies were performed to characterise and discriminate different cow, goat and sheep cheese through lipid biomarkers and metabolite profiles [6,14,23–26], also highlighting the ability of ^{1}H NMR to differentiate cheese samples as function of their ripeness, brine composition and adjunct cultures [17,27,28].

As regards Parmigiano Reggiano cheese, several studies were focused on the development of analytical methodologies able to discriminate Parmigiano Reggiano from non-authentic products. Very interesting results were obtained by using both un-targeted approaches, such as liquid chromatography coupled with high resolution mass spectrometry [29] (or quadrupole time-of-flight mass spectrometry [30]), Raman spectroscopy [31], ^{1}H NMR [13] and targeted approaches, such as stable isotope ratio analysis [32]. In most of these studies chemometrics tools were applied and allowed obtaining interpretable models with promising results in terms of classification capabilities. Notably, most of the studies on the chemical profile of Parmigiano Reggiano cheese have been focused on its ripening, authentication, and geographical characterization, while, to the best of the authors' knowledge, none of the used analytical techniques (including the NMR-based approaches) has been used to study Mountain CQ Parmigiano Reggiano cheese or products coming from a restricted geographical area.

In general, the mountain denomination label has been less investigated and specifically the evaluation of potential different metabolic profiles with respect to the same food commodity from plane area has been limited to few types of cheese [3,4,33].

In this study, the metabolic profiles of Mountain-CQ Parmigiano Reggiano cheese and conventional PDO Parmigiano Reggiano (conventional-PDO) cheese were evaluated by using NMR spectroscopy. Notwithstanding the richness of information held in NMR signals, the NMR profiles are rather complex to be analysed, mainly due to their high dimensionality, the presence of shifts among peaks and overlaps among signals. Furthermore, the extracted information needs to be examined altogether considering the uniqueness of the NMR spectrum as well as the non-triviality in the assignment of resonance to a given compound. Therefore, in this study, three different chemometrics strategies were used and compared. In the first, the spectra-as-such dataset was analysed by means of principal component analysis (PCA), without any compression and after a proper preprocessing.

In the latter two approaches, Multivariate Curve Resolution–Alternating Least Squares (MCR-ALS) was applied to obtain a satisfactory resolution of the individual peaks in ^{1}H NMR spectra [34] and to putatively identify the different metabolites. In both strategies, a resolution by intervals approach is applied [20], however, whilst in the first, the features obtained from each defined spectral interval were kept distinct, in the second, the resolved features corresponding to the same metabolite were combined to obtain whole pure spectral profiles. Finally, the latter were used as initial guess, and to define selectivity constraints in MCR-ALS then applied to the spectra-as-such NMR signals, according to a modified version of the method described in Puig et al. [35]. The selected features were also compared with specific spectral reference libraries (Chenomx software) for a putative identification of the selected compounds/metabolites.

2. Materials and methods

2.1. Samples collection

A total of 39 cheese samples was collected from Parmigiano Reggiano producers affiliated to "Consorzio Parmigiano Reggiano" (the Parmigiano Reggiano consortium): 20 were Italian Parmigiano Reggiano cheese with "Prodotto di Montagna - Progetto Territorio" denomination (Mountain-CQ), and 19 were conventional PDO Parmigiano Reggiano cheese (Conventional-PDO). The conventional-PDO samples were collected from dairies located in plain areas excluding those production sites located in the mountain areas (Figure S1 shows the sampled locations over the PDO production region map). The ripening time of the collected samples was about 24 months, with variability in the range from 23 to 35 months (Table S1). All cheese samples, soon after collection, were grated at 4500 rpm for 15 s (Grindomix GM200, Retsch, Germany) and stored in airtight bags at −20 °C until further analysis. The pH was measured (in aqueous environment) for each sample and resulted in the range 5.53–5.68, consistently with the study by Consonni et al. [13].

2.2. Samples preparation

Regarding sample preparation as related to the NMR untargeted metabolomics workflow, the extraction procedure was carried out as previously reported in literature by Consonni et al. [13]. Briefly, after thawing samples at 5 °C overnight, 100 mg of grated Parmigiano Reggiano cheese were dissolved into 600 μL of D_2O, mixed by vortexing (Vibration Mix, Falc, Italy) for 15 min and then subjected to sonication (Branson 3510 Mt, Electron Microscopy Sciences, Hatfield, PA, USA) for 15 min. After centrifugation at 12.000 rpm for 10 min, 540 μL of supernatant were transferred into the NMR tube. All NMR tubes were filled up to the required volume of 600 μL, by adding 60 μL of sodium-3-(-trimethylsilyl)propionate-d_4 (TSP-d_4). The samples were then analysed by NMR spectroscopy following a predetermined random order. To check the variability and the analytic reproducibility of the extraction procedure, one Mountain-CQ sample and one conventional-PDO sample were used as 'control samples' and replicated in each preparatory session. The list reporting the sample number, the ripening date, ageing in months, and the corresponding classification according to the different denominations (i.e., Mountain-CQ vs conventional-PDO samples) can be found in the supplementary material (Table S1).

2.3. ^{1}H NMR data acquisition

The ^{1}H NMR spectra were acquired on a Bruker Avance III 600 spectrometer (Bruker Biospin GmbH, Rheinstetten, Germany) operating at the Larmor frequency of 600.13 MHz for protons, equipped with a double-tuned cryoprobe (TCl) set for 5 mm sample tube and a cooled autosampler (SampleCase, at 5 °C). The spectra were acquired with TOPSPIN 2.1 (Bruker Biospin GmbH, Rheinstetten, Germany), using the NOESYGPPR1D sequence [36,37]. Presaturation of the water signal (4.77 ppm) [17,24,27,38] was employed. All experiments were performed at 298 K with a fixed receiver gain. Each free induction decay (FID) was collected using a total of 64 scans plus four dummy scans. Acquisition time was set to 7.8 s with a spectral width of 8403.361 Hz. Prior to Fourier transformation, the FIDs were zero-filled to 64 k points, and a 0.3 Hz Lorentzian line broadening was applied. The spectra were baseline- and phase-corrected using TOPSPIN built-in processing tools. This correction was performed automatically for all spectra and then, depending on the obtained results, further manual adjustments were performed when strictly necessary. For all spectra, the ppm scale was referenced to the TSP peak (0.00 ppm). The spectral window was set to 20.5 ppm (−5 to 15 ppm).

2.4. Data analysis

The raw NMR spectra were imported and processed under the MATLAB environment. The uninformative areas at the extremes of the spectrum (i.e., above 9 ppm and at negative ppm values, with no signals of appreciable intensity) together with the area corresponding to the water peak (around 4.77 ppm, with a broken residual water signal after presaturation) were removed from the spectral data, leading to a dataset with 66,200 experimental data points. The data were then downsampled (for ease of computation) and arranged into a numerical matrix of dimensions 49 × 33,100 (samples × variables), which also included the replicates of the both control samples.

Data preprocessing is a critical and case-dependent step in any multivariate data analysis workflow. Since in this study three different approaches were tested and compared, the employed preprocessing methods will be individually described for each approach.

The analytical workflow is depicted in Fig. 1. The aligned NMR spectra on the left-hand side, were firstly analysed "as such" by means of principal component analysis (PCA, Fig. 1a). These spectra then underwent interval-wise integration by means of multivariate curve resolution (MCR, Fig. 1b), which led to the creation of the "features dataset", also explored using PCA. Finally, the MCR-resolved signals corresponding to individual metabolites were combined to obtain a set of "reconstructed pure profiles" (Fig. 1c), which were in turn used as initial estimates for another run of MCR, this time on the "full spectra dataset", to try to unravel the mixture using the whole spectral information.

2.4.1. Spectra alignment

Spectra alignment was performed using the *icoshift* algorithm [39]. Signal alignment is needed since even small changes of experimental conditions may determine horizontal peak shifts, which, despite being reduced in scale, may introduce variability not imputable to real differences among the samples.

The alignment was directly done by intervals, as a preliminary global alignment (an option allowed by *icoshift* [40]) did not provide any clear alignment improvement. Therefore, a set of manually chosen small intervals was defined so that each one of them would contain individual peaks or small groups of signals, to allow better alignment. The average spectrum multiplied by a factor of 3 (option "average2" in *icoshift* [40]) was used as the alignment target. The result of the alignment process led to the definition of 80 intervals. The aligned dataset was the input for all multivariate data analysis methods as described in Sections 2.4.2 and 2.4.3 and is referred to as the "spectra-as-such dataset".

2.4.2. Multivariate curve resolution (MCR) for features extraction

Multivariate Curve Resolution (MCR [41,42]) was applied to extract features, i.e., peak areas of resolved metabolites, from the NMR spectra [20]. Eighty intervals were defined and MCR was automatically applied to each one of them to build four models of increasing complexity, from 2 to 5 components. To this aim a set of in-house MATLAB routines (a first version of a toolbox) was assembled. The toolbox allows: i) to easily visualize the modelled interval (Fig. 2, with the raw data plotted in 2a) and the resolved spectral profiles for the MCR models of increasing complexity models (Fig. 2b–e show the pure spectral profiles for the 2 to 5 components models, respectively); ii) automatically matching the resolved spectral profiles provided by MCR with a list of metabolite assignments obtained from literature sources (light blue bars in Fig. 2b–e) and, thus iii) selecting the components to retain and generating the features data set holding the resolved peak areas corresponding to the selected components. This setup also allows a quick and simple comparison between the resolved spectral profiles and the Chenomx reference library (version 9.02, Chenomx Inc., Edmonton, Alberta, Canada, https://www.chenomx.com, last access: March 06, 2023), as detailed in Section 2.4.2.1.

The joint use of literature sources and Chenomx was aimed at making the putative identification and assignment procedure more robust.

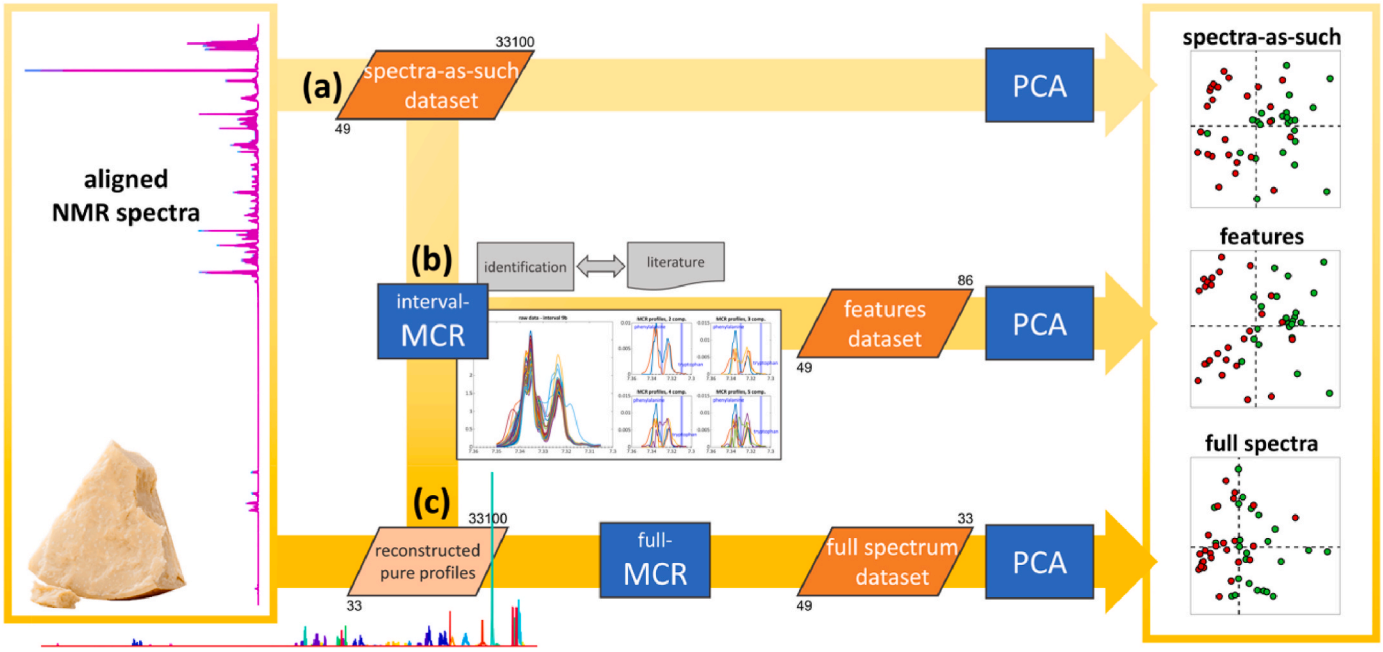

Fig. 1. The whole data analysis workflow. Starting from the aligned NMR data, three routes were taken: a) PCA analysis of the "spectra-as-such" dataset; b) interval-MCR resolution procedure to obtain and analyse the "features" dataset; c) full spectra MCR decomposition using the recombined signals as initial estimates.

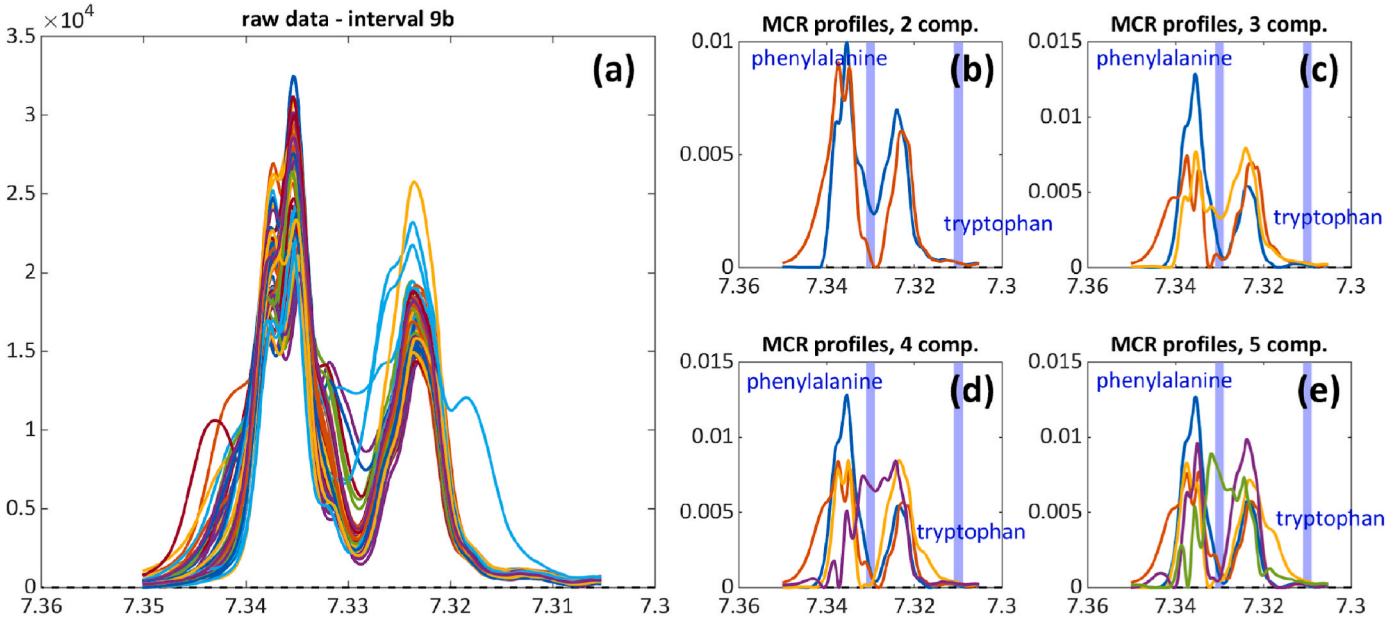

Fig. 2. Example of the graphical output of the *ad hoc* developed in-house MATLAB toolbox. The raw data in the selected interval are shown on the left side (a), while on the right side there are as many subplots as the computed MCR models (b–e). The resolved spectral profiles (one for each MCR component) are plotted in different colours, and the positions of the putative metabolites that could be present in the interval, based on literature references, are reported as light blue vertical lines, together with the corresponding metabolite's name. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

However, the level of confidence of the metabolite assignments can be more objectively assessed using the framework defined by the Chemical Analysis Working Group of the Metabolomics Standards Initiative [43,44]. In this framework, our work would approximately correspond to level 2 (i.e., "Putatively annotated compounds"). All metabolites' assignments are reported in Table 1: for each putative compound chemical shift, signal multiplicity and literature references that support each identification are provided.

2.4.2.1. MCR interval MATLAB toolbox. As introduced above, we developed an *ad hoc* toolbox (this novel toolbox will be made available on request from the authors, and a GUI development is in progress), to assist and automatize the interval processing procedure and visualize the MCR modelling results (Fig. 2). In particular, to guide the choice of the correct number of components, four MCR models are automatically built for each defined spectral interval (intervals are manually defined and provided in input on an excel file as a list indicating starting and ending point for each interval). Regarding the default settings for MCR

Table 1
List of resolved compounds with tentative names, chemical shifts (δ, ppm), signal multiplicity, signal assignment and references supporting the signal identifications. ([a] reported without information about chemical shift and multiplicity;[b] pure MCR-resolved feature obtained;[c] no MCR-resolved pure feature obtained, only mixed;[d] signals combined as an individual compound-related feature in the final "features dataset").

	Tentative compound name	Chemical shift (δ, ppm)	Multiplicity and assignment	References with NMR information	References mentioning the metabolite[a]
1	Acetate (acetic acid)	1.94[b]	s, α-CH_3	[13–17]; *Chenomx*	[7]
2	Alanine[c]	3.76[c]	q, α-CH	[13,14, 16–18]; *Chenomx*	[7,19]
		1.48[c]	d, β-CH_3	[13–18]; *Chenomx*	[6,19]
3	Arginine	3.75[c]	t, α-CH	[13,18]; *Chenomx*	[7]
		3.25[b]	t, δ-CH_2	[13,18]; *Chenomx*	[7]
		1.9[c]	m, β-CH_2	[13,15,18]; *Chenomx*	[7]
		1.65[b]	m, γ-CH_2	*Chenomx*	[7]
4	Asparagine[d]	4.0[b]	m, α-CH	[13,14,18, 19]; *Chenomx*	[7]
		2.96[b]	m, β-CH_2	[13,14, 17–19]; *Chenomx*	[7]
		2.86[b]	m, β′-CH_2	[13,19]*, [14,17,18]; *Chenomx*	[7]
5	Aspartate (aspartic acid)	3.9[b]	m, α-CH	[13,14]; *Chenomx*	[7]
		2.82[b]	m, β-CH_2	[13,14,17, 19]; *Chenomx*	[7]
		2.69[b,c]	m, β′-CH_2	[14,17]; *Chenomx*	[7]
6	Butyrate (butyric acid)	2.15[c]	t, α-CH_2	[17]; *Chenomx*	[7]
		1.57[b]	sext, β-CH_2	[17]; *Chenomx*	[7]
		0.9[b]	t, γ-CH_3	[17]; *Chenomx*	[7]
7	Cadaverine[a]	3.03[c]	t	*Chenomx*	[17]
8	Choline	4.1[c]	m, CH_2	[14,16]; *Chenomx[*	[7]
		3.2[b]	s, CH_3	[14–16]; *Chenomx*	[7]
9	Citrate (citric acid)	2.73[b,c]	d, CH_2	[14,17]; *Chenomx*	[7]
		2.55[b]	d, CH_2	[14,15,17]; *Chenomx*	[7]
10	Creatinine	4.04[b]	s, CH_2	[16]; *Chenomx*	[7,15]
11	Ethanol	1.18[b]	t, CH_3	[15]; *Chenomx*	[7]
12	Glutamate (glutamic acid)[c]	2.36[c]	m, γ-CH_2	[13–15,17, 18]; *Chenomx*	[7,16]
		2.13[c]	m, β-CH	[14,17]; *Chenomx*	[7,16]
		2.05[c]	m, β′-CH	[17–19]; *Chenomx*	[7,16]
13	Glycerol	3.8[c]	m, CH	[14,17,18]; *Chenomx*	[7]
		3.65[b]	m, CH_2	[14,17,18]; *Chenomx*	[7]
		3.56[b]	q, CH_2	[14,17,18]; *Chenomx*	[7]
14	Glycine	3.56[b]	s, α-CH_2	[13–19]; *Chenomx*	[7]

Table 1 (*continued*)

	Tentative compound name	Chemical shift (δ, ppm)	Multiplicity and assignment	References with NMR information	References mentioning the metabolite[a]
15	Isoleucine[d]	3.67[b]	d, α-CH	[13,14,18, 19]; *Chenomx*	[7,16]
		1.46[c]	m, γ-CH_2	[13,14,18]; *Chenomx*	[7,16]
		1.27[b]	m, γ′-CH_2	[13,14]; *Chenomx*	[7,16]
		1.01[b]	d, γ-CH_3	[13–15,18, 19]; *Chenomx*	[7,16]
		0.94[c]	t, δ-CH_3	[13,14,18]; *Chenomx*	[7,16]
16	Lactate (lactic acid)	4.13[b,c]	q, α-CH	[13,14, 16–19]; *Chenomx*	[7]
		1.34[b]	d, β-CH_3	[13–19]; *Chenomx*	[7]
17	Leucine[c]	3.72[c]	m, α-CH	[13,14,18]; *Chenomx*	[7,16]
		1.71[c]	m, β-CH_2	[13–15, 17–19]; *Chenomx*	[7,16]
		0.97[c]	d, δ-CH_3	[13–15, 17–19]; *Chenomx*	[7,16]
		0.96[c]	d, δ′-CH_3	[13–15, 17–19]; *Chenomx*	[7,16]
18	Lysine	3.03[c]	t, ε-CH_2	[13–15,17, 18]; *Chenomx*	[7,16]
		1.9[c]	m, β-CH_2	[13,14,17, 18]; *Chenomx*	[7,16]
		1.73[c]	m, δ-CH_2	[14,17–19]; *Chenomx*	[7,16]
		1.5[c]	m, γ-CH_2	[14,18,19]; *Chenomx*	[7,16]
		1.45	m, γ′-CH_2	[14,18]; *Chenomx*	[7,16]
19	Methionine	3.87[b]	m, α-CH	[13,14,18]; *Chenomx*	[7,16]
		2.64[b]	t, γ-CH_2	[13,14, 17–19]; *Chenomx*	[7,16]
		2.14[c]	m, S-CH_3, β-CH_2	[13,14, 17–19]; *Chenomx*	[7,16]
20	N-acetyltyrosine[a]	4.39[b]	m	*Chenomx*	/
21	Ornithine	3.06[b]	t, δ-CH_2	[18]; *Chenomx*	[7,16]
22	Phenylalanine[d]	7.42[b]	m, 3,5-CH	[13,14,17, 18]; *Chenomx*	[7,16]
		7.37	m, 4-CH	[13,14,17, 18]; *Chenomx*	[7,16]
		7.33[b]	m, 2,6-CH	[13,14,17, 18]; *Chenomx*	[7,16]
		3.98[c]	m, α-CH	[13,14,18]; *Chenomx*	[7,16]
		3.28[b]	d, β-CH_2	[13,14, 17–19]; *Chenomx*	[7,16]
		3.13[b]	m, β′-CH_2	[13,14, 17–19]; *Chenomx*	[7,16]
23	Proline	3.42[b]	m, δ-CH_2	[13,14,17, 18]; *Chenomx*	[7,16]

(*continued on next page*)

Table 1 (*continued*)

	Tentative compound name	Chemical shift (δ, ppm)	Multiplicity and assignment	References with NMR information	References mentioning the metabolite[a]
		3.33[b]	m, δ'-CH_2	[13–15, 17–19]; *Chenomx*	[7,16]
		2.36[c]	m, β-CH_2	[13,14,17, 18]; *Chenomx*	[7,16]
		2.04[c]	m, β′-CH_2	[13,17–19]; *Chenomx*	[7,16]
		1.99	m, γ-CH_2	[13,14,17, 18]; *Chenomx*	[7,16]
24	Pyroglutamate (pyroglutamic acid)	4.18[b]	m, α-CH	[13,18]; *Chenomx*	[7]
		2.5[b]	m, β-CH_2	[13,18]; *Chenomx*	[7]
		2.4[b]	m, β′-CH_2	[13,18]; *Chenomx*	[7]
25	Serine	3.96[b]	m, β-CH_2	[13,14,18, 19]; *Chenomx*	[7]
		3.90[c]	m, β′-CH_2	[13,14,18, 19]; *Chenomx*	[7]
		3.85[b]	m, α-CH	[13,14,18]; *Chenomx*	[7]
26	Threonine	4.26[b]	m, β-CH	[13,14,18, 19]; *Chenomx*	[7,16]
		3.58[b]	d, α-CH	[13,14, 17–19]; *Chenomx*	[7,16]
		1.33	d, γ-CH_3	[13,14,18]; *Chenomx*	[7,16]
27	Tryptophan[a]	7.74[b]	m	*Chenomx*	[7,16]
		7.54[b]	m	*Chenomx*	[7,16]
		7.28[b]	m	*Chenomx*	[7,16]
		7.19[c]	m	*Chenomx*	[7,16]
		4.06[b]	m	*Chenomx*	[7,16]
		3.48[b]	m	*Chenomx*	[7,16]
28	Tyrosine	7.18[b,c]	m, 3,5-CH	[13,14,17, 18]; *Chenomx*	[7,16,19]
		6.9[b]	m, 2,6-CH	[13–15,17, 18]; *Chenomx*	[7,16,19]
		3.2[b]	m, β′-CH	[14,17,18]; *Chenomx*	[7,16,19]
		3.06[b]	m, β-CH	[14,17,18]; *Chenomx*	[7,16,19]
29	Uracil[a]	5.8[b]	d	*Chenomx*	[7]
30	Valine	3.61[b]	d, α-CH	[13,14,17, 18]; *Chenomx*	[7,16]
		1.04[b]	d, γ-CH_3	[13,14, 17–19]; *Chenomx*	[7,16]
		0.99[b]	d, γ′-CH_3	[13–15,18, 19]; *Chenomx*	[7,16]
31	Xanthine[a]	7.92[b]	s	*Chenomx*	/

modelling, the initial spectral profile matrix is estimated by using the SIMPLISMA algorithm [45]. Then, the non-negativity constraints are applied to both the concentrations and the profiles directions, using in both cases the fast non-negative least squares algorithm. The maximum number of iterations is set to 1000. The resolved spectral profiles are normalised using the Euclidean norm. Finally, through the plot, as in Fig. 2, the results for each interval can be inspected to identify and select the chemically meaningful components, while background and baseline-like profiles could be discarded. An example of discarded components can be seen in Figure S2c, where one citrate signal was obtained from a 3 components MCR model, while the other two components represented baseline-like behaviour and spurious signals.

An example of the resolved profile selection, matching with reference information and identification process is given in Fig. 3, in which the amino acid phenylalanine is investigated and identified thanks to its very distinctive NMR profile (Fig. 3a). The presence of phenylalanine in cheese has been reported in literature [13,14,17–19] and in some cases also its chemical shifts and assignments have been provided. In this example, the resolved profiles of the four MCR models are shown in Fig. 2b–e: this is a simplified visualisation of the one provided by the developed toolbox, and the real look of the MATLAB interface is shown in Fig. 2. This representation of the models allowed us to easily compare and match the resolved profiles with the NMR profile of phenylalanine that can be found in the reference library of Chenomx (Fig. 3f). It is possible to recognize the signal related to phenylalanine (plotted in pink and red) in all four MCR models, while at the same time it is possible to make considerations on how models with different numbers of components result in different profile extraction performances. The first model (Fig. 3b, fitted with two components), and more specifically the first resolved component highlighted in red, provides the best performance in resolving the correct spectral profile, while at the same time filtering out peculiar signals (e.g., the shoulder signal between 7.315 and 7.32 ppm belonging to one sample alone) and unravelling the contribution due to another chemical compound (which has a peak different from phenylalanine at about 7.34 ppm). The other three models (Fig. 3c–e), fitted using from three to five components, show very similar performances, but with increasing model complexity some artefacts start appearing in the multiplet profile of phenylalanine. Based on these considerations, the two-components model was selected, and the resolved profile of the first component was identified as phenylalanine and therefore included in the features dataset.

2.4.2.2. Construction of the "features dataset". The result of the interval-based resolution process consisted of 93 resolved chemical components, whose relative concentrations (corresponding to the values in the concentrations matrix obtained by MCR decomposition) obtained by MCR were merged column-wise to generate a new dataset, hereinafter referred to as the "features dataset" of dimensions 49 × 86 (8 of the features, signals of the same compounds, were combined as detailed in Table 1). The lack of fit (l.o.f.) value ranged between 0.03 and 1.62% considering all selected MCR models. Most models however show l.o.f. values below 0.5%. Of the 86 extracted features 48 were tentatively assigned to many different signals of individual chemical compounds and 13 features to mixtures of different metabolites signals. The remaining 25 extracted features were labelled as "unknown", as reported in Section 3.1. A list reporting the 31 individual metabolites of which all or most signals were assigned is also provided in Table 1.

2.4.2.3. The "full spectrum-MCR" approach: reconstruction of the pure profiles from the resolved features. The MCR-resolved spectral profiles originating from different intervals but referring to the same metabolite were joined to recreate the compound's pure profile, including as many real characteristic signals as possible. Following this procedure, a set of reconstructed "pure" and complete profiles was obtained. This spectral matrix was used as the initial estimates input for an additional MCR modelling step, this time running on the spectra-as-such dataset following a modified version of the approach originally suggested by Puig et al. [35].

The key aspect of this approach is to resolve the individual metabolites profiles directly from the NMR spectra-as-such, thus operating on the whole spectral width. Due to the complexity of the NMR signals, this cannot be achieved in a single step. Puig et al. [35] proposed to proceed first by applying MCR on small intervals (as we did for obtaining the NMR features data set), then to recursively use Pearson's correlation coefficient among the resolved features to highlight the ones most likely

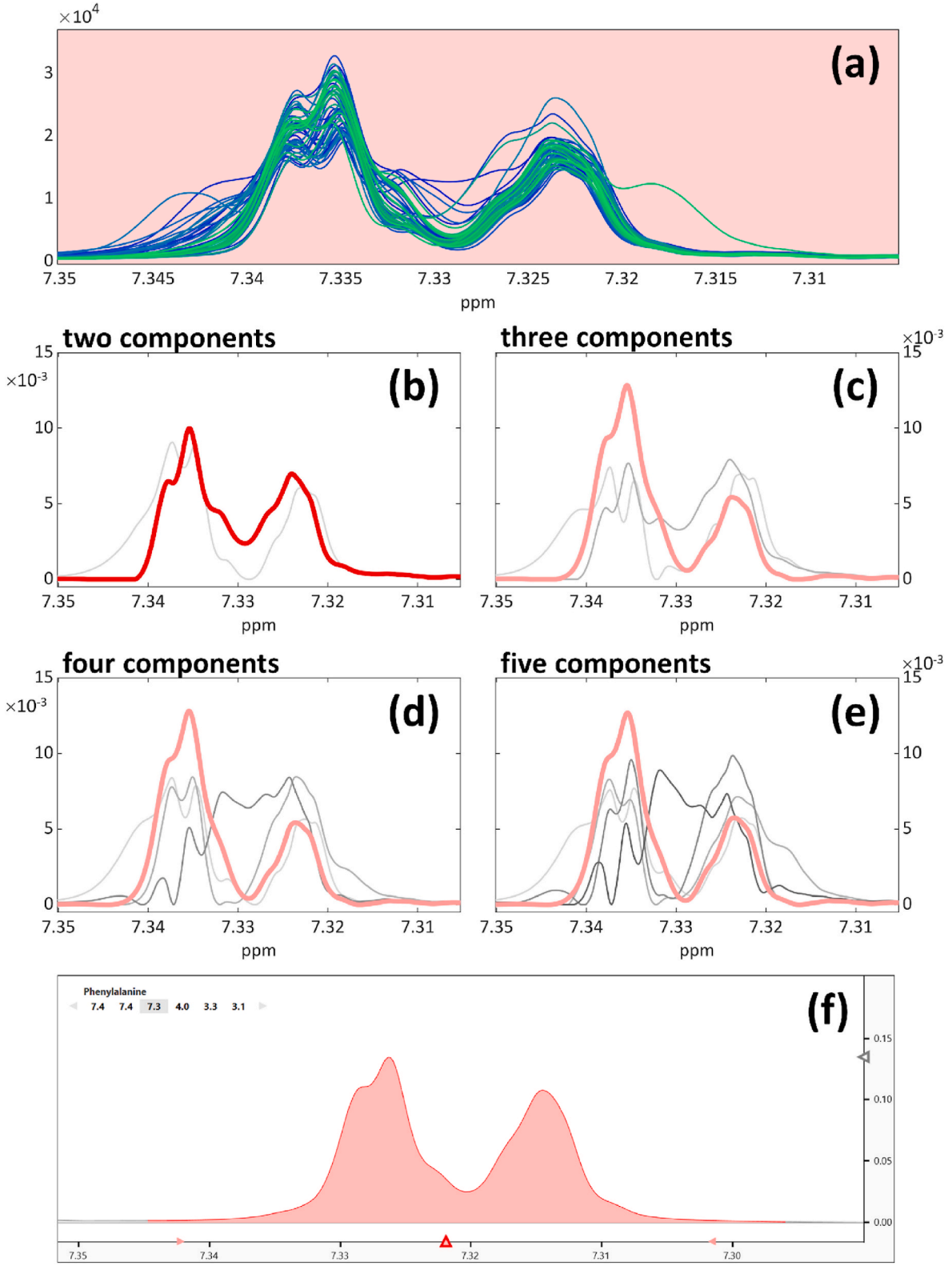

Fig. 3. Identification procedure of the amino acid phenylalanine. The aligned raw data are shown in (a). Four models (b–e) were built using from two to five components: the resolved profiles were compared with the real profile of phenylalanine (f), provided here by the reference library Chenomx, reported here as a screenshot from the software interface. The profiles matching with phenylalanine are represented in pink, while the one selected for building the features dataset is plotted in red (b, model with 2 components). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

to be associated to a single metabolite. Finally, the resolved "recombined" pure profiles, which are in principle directly linked to the assigned metabolites, are used as the initial guess of the pure spectra matrix, with the aim of helping the MCR algorithm to converge to a feasible solution. In this way, each metabolite can be modelled as one individual component.

Our approach is similar, but the profiles to be recombined, differently from Puig's, were selected based on putative identification through the literature and Chenomx library search instead of correlation analysis of the resolved features. Starting from the 86 resolved features (Section 2.4.2.2), a total of 33 pure profiles were obtained, of which 26 corresponding to individual metabolites, 5 to overlapped signals of two metabolites, and 2 to overlapped signals of two non-identified metabolites. After using these 33 profiles as initial estimates for MCR on the NMR spectra-as-such dataset (the "full spectrum MCR" approach) a new dataset of resolved MCR components was obtained and will be hereafter referred to as the "full-MCR dataset".

2.4.3. Exploratory data analysis

Exploratory analysis was performed using principal component analysis (PCA [46,47]) on all datasets (i.e., the NMR *spectra-as-such* dataset, the *features* dataset and the *full-spectrum* MCR dataset). The obtained PCA results were compared in terms of extracted information, interpretability, and workload.

Blockscaling to unit block variance was applied to the aligned NMR signals in the case of the spectra-as-such dataset. The same intervals as defined for the alignment step with *icoshift* were used. Autoscaling was used on the features dataset and the *full-spectrum* MCR dataset.

2.5. Software

Data analysis was carried out on MATLAB 2019a and 2021a (Mathworks, Natick, MA, USA) using a combination of functions, toolboxes, and in-house written routines. Data preprocessing, MCR and PCA modelling were performed using the functions included in the PLS_Toolbox (version 9.0, Eigenvector Research Inc., Manson, WA, USA). Peak alignment was performed using the *icoshift* algorithm [39], downloadable from https://ucphchemometrics.com/186-2/algorithms/, last accessed April 20, 2023). Peak identification and assignment of the resolved MCR components were based on the reference library of Chenomx NMR Suite (version 8.3, Chenomx Inc., Edmonton, Alberta, Canada).

3. Results and discussion

3.1. Extracted chemical features by MCR

The list of identified compounds is reported in Table 1. Note that some signals listed on the table do not correspond to an individual resolved chemical feature, but all of them are represented in the features dataset. Together with the literature references that support the assignments of each identified metabolite, their presence on *Chenomx* is also reported, as well as possible mentions of the metabolite in references about cheese (last column of Table 1). All unassigned signals (marked as "unknown") are reported in Table S2 (Supplementary material).

The Parmigiano Reggiano cheese spectra resulted rich in protein and non-protein amino acid spin systems, namely proline, aspartic acid, leucine, valine, methionine, tyrosine, phenylalanine and serine and, to a lesser extent, amines, alcohols, organic acids and nucleotides. The last two chemical classes present contributions characterised by low intensities which are indeed rather difficult to resolve. As reported in previous studies [13], most of the detectable metabolites in Parmigiano Reggiano cheeses could be the product of enzymes and/or microorganism activity during manufacturing or coming from biochemical reactions during cheese ripening. As regards amino acids, they can also result from proteolysis concerning the casein matrix degradation during the ripening time and their contents could change as a function of time [13].

A representation of a typical spectrum of Parmigiano Reggiano is provided in Fig. 4, together with the metabolites reported on Table 1. Metabolites with very weak signals may not be reported in the figure, as well as some of the unknown signals.

3.2. Analysis of spectra-as-such vs MCR extracted features

The discussion of the results is focused on the grouping related to the different denominations defined by PCA analysis for both datasets. In Fig. 5 are reported the scores and loadings plots for the most relevant principal components for spectra-as-such (Fig. 5a and b) and MCR-extracted features (Fig. 5c, d and 5e) datasets.

The Mountain-CQ and conventional-PDO replicated samples (described in section 2.2) lie very close in the scores plot (encircled by a dashed line in Fig. 5a and c) testifying good reproducibility.

Starting from the information about the different denominations (i.e., Mountain-CQ vs conventional-PDO), the PCA analysis on the NMR *spectra-as-such* dataset provided a clear grouping trend related to the two different denominations (Fig. 5a), mainly visible on the first principal component, where Mountain-CQ samples are generally located at negative values and conventional-PDO samples at positive ones. Few samples are mislocated and are the same for the two data sets, in particular two conventional-PDO samples, i.e., number 34 and 19 (which looking at the single features show a higher citrate and lactate content w.r.t. the others of the same category); samples number 5, 8 and 27 of Mountain-CQ which show a combination of higher pyroglutamate, phenylalanine and lower citrate w.r.t. to their category. These samples are worthy of further investigation also considering that samples 5 and 27 (as well as 30) are produced in a very close specific mountain area.

Fig. 5b shows the loadings of PC1, as a line plot coloured according to the variance accounted for by each individual spectral variable on PC1. The spectral regions that have been considered most relevant are regions characterised by an absolute loading value greater than 0.01 and attaining an explained variance higher than 70% (as highlighted by the black squares on the x-axis in Fig. 5b), and those regions showing the most negative loadings values. In particular, the most influential spectral regions are mainly related to amino acids, as well as spectral regions which can be related to organic acids and esters. These are summarised in Table 2, along with their chemical shift values, distinctly for each denomination (based on the sign of PC1 loadings, those metabolites with positive sign are directly related to conventional-PDO and those with negative sign to Mountain-CQ). In particular, for conventional-PDO samples we detected spectral regions that could be associated with pyroglutamic and glutamic acids, lysine, valine, glycine, phenylalanine, and nucleosides.

However, the interpretation of loading information is laborious because the different peaks are strongly overlapped and only for a few of them was possible to attempt a putative identification.

On the contrary, in the case of the features dataset, the application of MCR provided clear insights since the overlapped signals could be resolved, making the interpretation of PCA loadings more straightforward, since the chemical information can be directly inspected in terms of specific metabolites. It can be noticed that, also in this case, the first PC is the one describing the separation between the two denominations (Fig. 5c): the Mountain-CQ samples are generally located at negative values of PC1 while the conventional-PDO samples are located at positive values. As expected, free amino acids are the most influent metabolites as it can be deduced from the corresponding loadings plot, that for ease of interpretation is split into two distinct subplots: in Fig. 5d the features corresponding to the amino acids are depicted, while in Fig. 5e the features corresponding to other metabolites, such as organic acids, amine, alcohols, and purine bases are shown. The two plots share the same scale. Thus, by applying the interval MCR approach the different

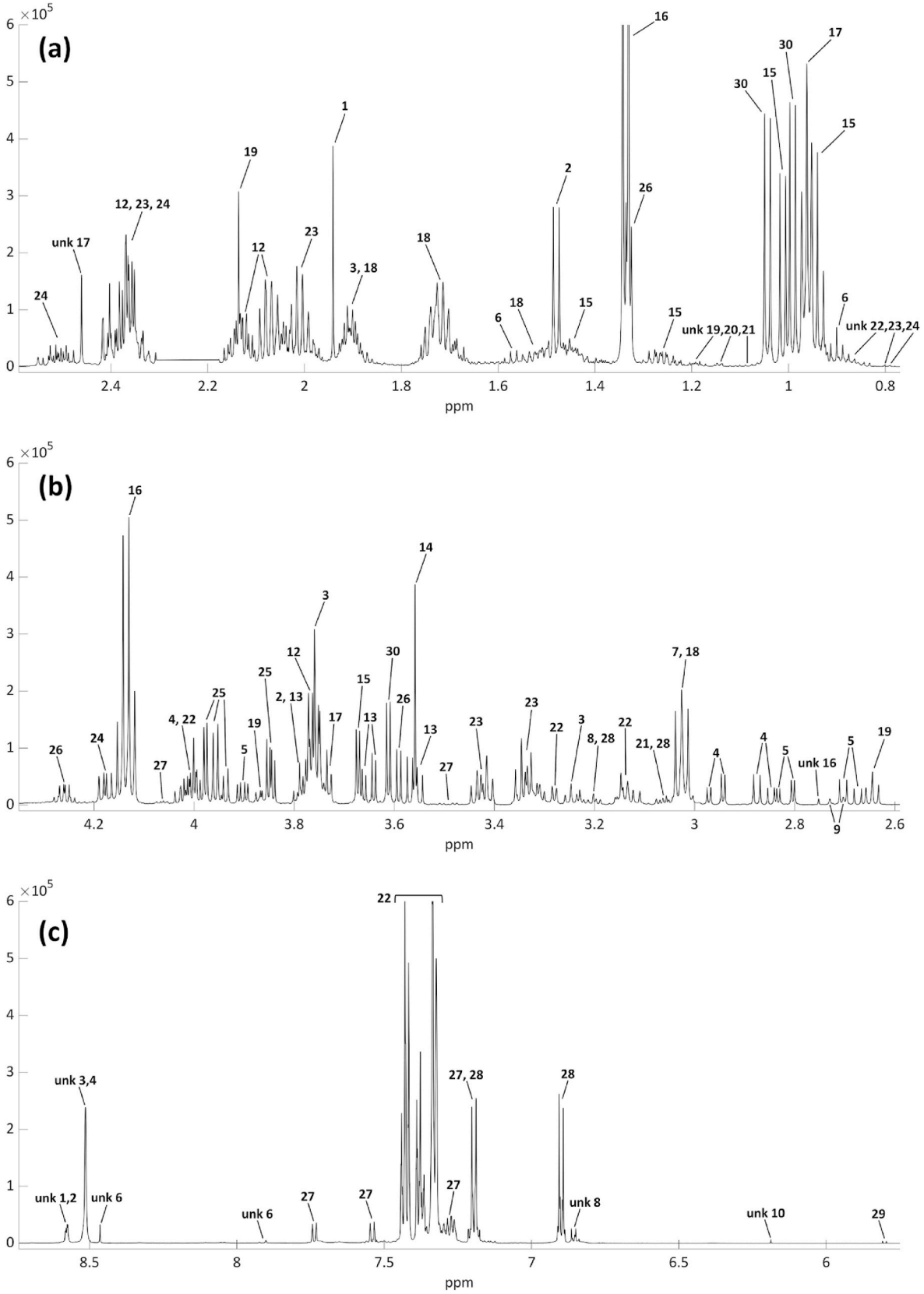

Fig. 4. A representation of a typical spectrum of Parmigiano Reggiano with all the identified signals (full assignments are reported in Table 1). Metabolites with very weak signals may not be reported in the figure, as well as some of the unknown signals.

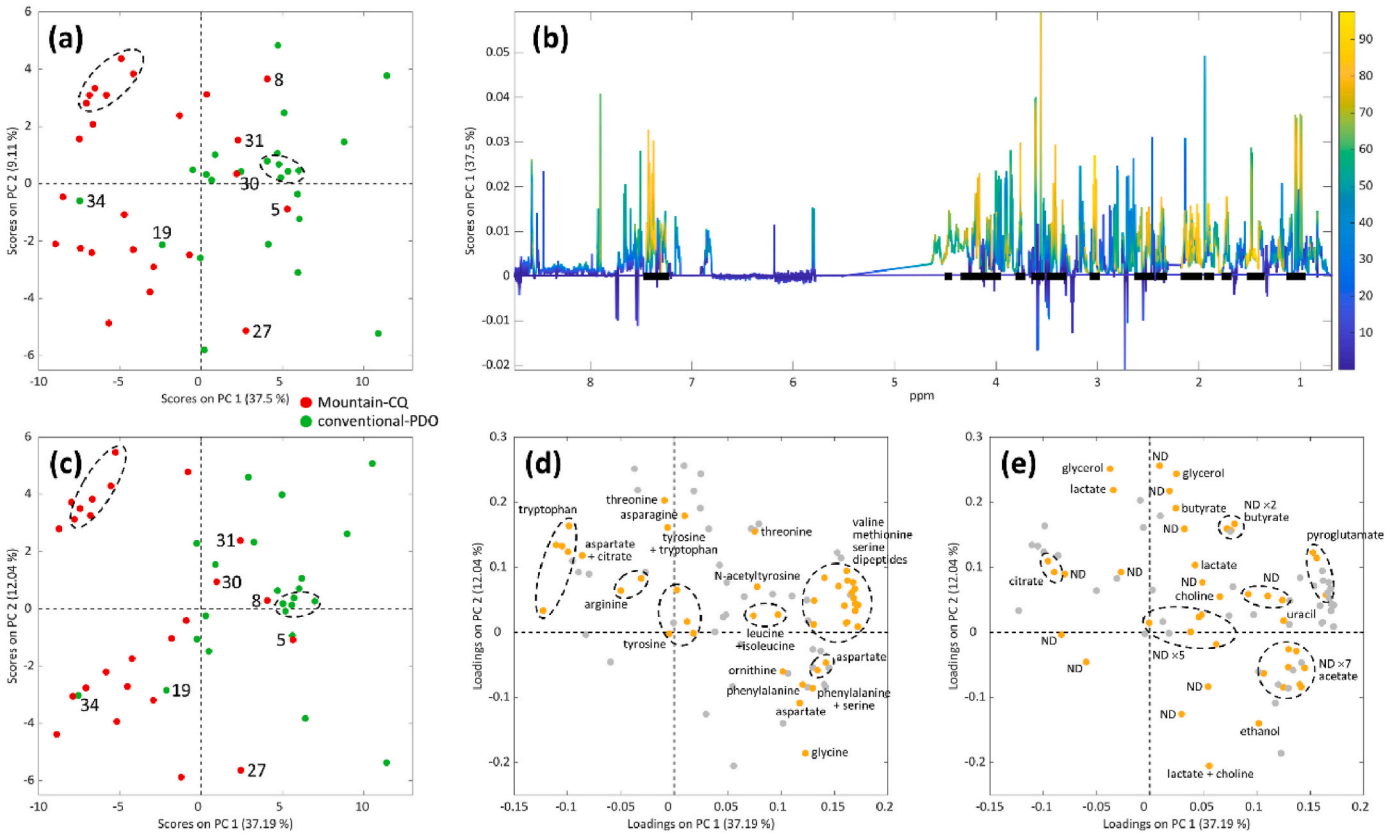

Fig. 5. Comparison of the PCA results of the *spectra-as-such* and *features* data sets. PCA of spectra as such dataset: (a) - PC1 vs PC2 scores plot; (b) - PC1 loadings plot. The colour scale refers to the explained variance of each variable (i.e., a single wavelength) by PC1. The black squares on the x-axis correspond to wavelengths which have an absolute value greater than 0.01 and attaining an explained variance higher than 70%. PCA of features dataset: (c) - PC1 vs PC2 scores plot, (d) - PC1 vs PC2 loadings plot highlighting the features corresponding to amino acids (orange colour); (e) - PC1 vs PC2 loadings plot with other classes of compounds highlighted (orange colour). The samples which are encircled in (a) and (c) are replicates. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 2
List of putative metabolites corresponding to spectral regions depicted as salient by PC1 loadings (PCA on the full spectra) inspection.

Chemical shift (δ, ppm)	Mountain-CQ PR	conventional-PDO PR
8.52		unknowns 1 and 2
8.46	\	formic acid
7.92	\	xanthine
7.90	\	unknown 6
7.74; 7.54; 3.49	tryptophan	\
7.6; 1.55	\	nucleotides
7.42; 7.37; 3.98		phenylalanine
4.18	\	pyroglutamic acid
3.76; 1.48	\	alanine
3.61; 1.04; 0.99	\	valine
3.58	threonine	\
3.56	\	glycine
3.42; 2.04; 1.99	\	proline
3.25	arginine	\
3.03; 1.9; 1.73	\	lysine
2.73	citrate	\
2.36; 2.13; 2.05	\	glutamic acid
2.15	\	butyric acid

denominations under investigation appear to be distinguished by the amino acids profile, as well as by several organic acids, which seem more abundant in conventional-PDO samples.

The amino acids have been identified in previous studies based on NMR metabolomics analysis of Parmigiano Reggiano cheese [13,36,48]. In particular, our results show that Mountain-CQ samples can be characterised by few markers as tryptophan (which is associated with relatively negative loadings values on PC1 of the *features* dataset of all its three resolved signals from different intervals). Tryptophan in Parmigiano Reggiano PDO cheese has been reported in literature as a possible marker of ripening age [49]. However, in our study the ripening age varies moderately, and in any case PC1 is not related to this difference (as shown in Figure S3, which is coloured according to ripening months).

Arginine, citrate, lactate, and threonine are other metabolites which seem characteristic of the Mountain-CQ category, as highlighted by PCA analysis on both datasets. Arginine was reported in previous NMR metabolomics studies on Parmigiano Reggiano [13,36] as related to casein degradation occurring during ripening. Valine and serine were also reported as possible markers of conventional-PDO samples. In particular, valine was associated with seasonal variability of Parmigiano Reggiano production [48], while serine was reported to play a role in the cheese maturation process [13].

Several esters emerge as relevant in our results, more specifically acetate and butyrate, which seem more abundant in conventional-PDO samples. These findings seem consistent, since the most relevant esters (present in higher concentration) are short-chain even-numbered fatty acids [13,30]. In fact, these metabolites are usually associated with specific fruity and floral notes that characterise the organoleptic properties of this cheese [50]. Other compounds like choline [7,14,16], glycerol [7,14,17,18] and uracil [7] were detected in cheese with NMR spectroscopy, but no indication could be found about their relationship with production zones or cow feeding regimens.

3.3. Analysis of the full-MCR dataset

The 31 metabolites putatively identified (Table 1) and two couples of unknowns whose areas were found to be strongly correlated (i.e., unknown 1 + 3 and unknown 6 + 11), constituted the set of 33 recombined spectral profiles which could be used as an initial guess (the spectral matrix S) to the MCR algorithm operating on the spectra-as-such data. For these 33 components the selectivity constraint was also applied, based on their spectral profile. After MCR resolution the relative concentration values of these 33 components were used as the "full-MCR dataset", and autoscaled prior to PCA. The results are reported in Fig. 6.

The separation of the two denomination is also observable on the first PC (Fig. 6a), where Mountain-CQ samples are mostly located at negative PC1 values, the metabolites which attain negative PC1 loadings (Fig. 6b), hence characterised them, are: citrate, threonine, arginine, (tyrosine, lactate to a minor extent). These are mostly in agreement with the "MCR-extracted feature" model, the only exception being tryptophan and glycerol which are only slightly contributing to PC1 but with positive signs and tyrosine which was not contributing to PC1 in the "MCR-extracted feature" model. On positive PC1 are found the conventional-PDO samples which are characterised by the same metabolites previously discussed for the "features data set". The mislocated samples are less in this case (34, 5 and 27 are in common with the other two data sets).

Overall, the results are similar to the "MCR-extracted feature" model and thus this extra step seems not necessary in this case. The observed differences could be due to the fact that the extracted features, which we were not able to putatively assign (labelled with "ND" in Fig. 5c), were also not correlated among them, so only two couples were merged with the Puig approach to obtain a whole spectrum (the 2 additional components with respect to the 31 identified metabolites reported above) and the other were not used in the analysis of the full-MCR data set.

4. Conclusions

All datasets provided comparable information regarding the sample distribution, as in all three cases the separation between the two product categories conventional-PDO and Mountain-CQ could be highlighted. It is noteworthy that this separation is seen as the main data variance source (i.e., along PC1 in all models). This is the first study in which very closely related products, respecting the same PDO requirements, namely mountain denomination with additional quality requisites, i.e., Parmigiano Reggiano "Prodotto di Montagna - Progetto Territorio" (Mountain-CQ), and conventional-PDO Parmigiano Reggiano, selecting samples only produced in flatland areas, are compared. The obtained results are very promising notwithstanding some limitations which can be overcome by further investigations.

The metabolites found relevant need to be confirmed enlarging the sampling as well as the analytical characterization by other techniques. In the metabolomic field both MS and NMR are widespread techniques, and they are generally seen as complementary, thus considering the higher sensitivity of MS additional information may come from high resolution LC-MS and analysis of volatile fraction by GC-MS. In addition to these it could also be interesting to evaluate other spectroscopic techniques, e.g., NIR and Raman spectroscopies. Work is in progress in these directions.

In any case, this study provides a first indication that it could be possible to define identity traits of the Mountain-CQ.

From another point of view, the datasets of the study are denoted by different ease of interpretation of the variables, i.e., of the metabolites causing the samples separation into the two categories. In other words, the chemical information was more or less available depending on the dataset (and the analysts' NMR interpretation skills): the features dataset surely proved to be the easiest to interpret even if it contains, by definition, less information than its parent spectra-as-such dataset (which contains all signals). In this specific case, the full-MCR dataset did not provide any additional insights, nonetheless it is useful to have the whole NMR profile of each putative identified metabolite resolved, to ease interpretation and comparison with literature databases and/or standards.

Obtaining the three datasets required different efforts, and the use of specific toolboxes and in-house MATLAB functions. From this point of view, we believe that the time and dedication toward the construction of the features dataset are worth the effort, since throughout the process we were able to efficiently gather a lot of information about the chemical content of our samples, digging deeper into the structure of the data. Moreover, the integration of a literature list of previously documented metabolites (together with the reference library of Chenomx) into the analytical workflow allowed us to profitably make decisions on the data modelling side, while combining our chemical and spectroscopic knowledge about the system under examination.

The analytical workflow from raw data import to obtaining the MCR-resolved features dataset is quite demanding and it often prevents the inexperienced user from adopting it. Thus, to ease and improve this step

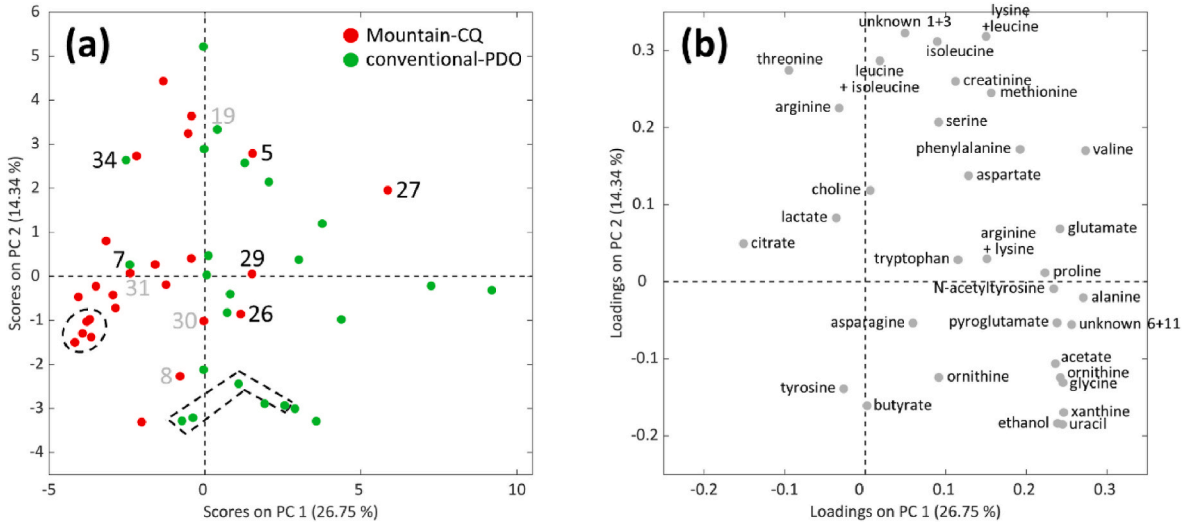

Fig. 6. PCA results of the full-MCR dataset: (a) PC1 vs PC2 scores plot and (b) PC1 vs PC2 loadings plot. Black labelled samples in (a) are those falling among the opposite class of belonging. The gray labelled samples are those which were found among the opposite class of belonging in the previous PCA models (Fig. 5a, c). The samples which are encircled in (a) are replicates.

we developed the in-house MATLAB functions and routines for both processing the data efficiently, but also to make the data and modelling results visualisation as straightforward as possible. Easy visualisation allows direct comparison with the Chenomx database, speeding up the metabolites identification and results interpretation. Moreover, the developed code reads the signal assignments from literature organised as a simple Excel table, and it allows plotting them automatically onto the raw data as well as onto the resolved MCR spectral profiles. In this perspective, the further development of the in-house routines could result in an easy-to-use MATLAB toolbox, to help reducing the workload required to obtain both the starting aligned spectra-as-such dataset and the extracted features dataset, while at the same time making the approach available to a wider audience of NMR data analysts.

CRediT authorship contribution statement

N. Cavallini: Conceptualization, Investigation, Methodology, Writing – review & editing, Software, Formal analysis, Data curation, Visualization, Writing – review & editing. **L. Strani:** Conceptualization, Methodology, Formal analysis, Data curation, Validation, Writing – original draft, Writing – review & editing. **P.P. Becchi:** Investigation, Formal analysis, Validation, Visualization. **V. Pizzamiglio:** Resources, Data curation, Writing – review & editing. **S. Michelini:** Investigation, Data curation, Writing – review & editing. **F. Savorani:** Investigation, Formal analysis, Supervision, Validation, Writing – review & editing. **M. Cocchi:** Conceptualization, Project administration, Formal analysis, Supervision, Writing – review & editing. **C. Durante:** Methodology, Writing – original draft, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Marina Cocchi reports financial support was provided by Consorzio Formaggio Parmigiano Reggiano. Marina Cocchi reports a relationship with Fondazione di Modena that includes: funding grants.

Data availability

The data that has been used is confidential.

Acknowledgments

This research has been supported by Consorzio Formaggio Parmigiano Reggiano, and by University of Modena and Reggio Emilia - Fondazione Modena through FAR Mission Oriented 2021 funds (project: MOUNTAIN-ID). The authors wish to thank Prof. Anna De Juan for useful advices and guidance concerning the approach derived from Puig et al., used to process the full-MCR dataset.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.aca.2023.341761.

References

[1] EUR-Lex - 32014R0665 - EN - EUR-Lex. https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=celex:32014R0665. (Accessed 20 April 2023).
[2] L. Zhang, P. Wang, S. Li, D. Wu, Y. Zhong, W. Li, H. Xu, L. Huang, Differentiation of Mountain- and garden-cultivated ginseng with different growth years using HS-SPME-GC-MS coupled with chemometrics, Molecules 28 (2023) (2016), https://doi.org/10.3390/molecules28052016.
[3] L. Moran, A. Aldezabal, N. Aldai, L.J.R. Barron, Terpenoid traceability of commercial sheep cheeses produced in mountain and valley farms: from pasture to mature cheeses, Food Res. Int. 126 (2019), 108669, https://doi.org/10.1016/j.foodres.2019.108669.
[4] S. Segato, G. Galaverna, B. Contiero, P. Berzaghi, A. Caligiani, A. Marseglia, G. Cozzi, Identification of lipid biomarkers to discriminate between the different production systems for Asiago PDO cheese, J. Agric. Food Chem. 65 (2017) 9887–9892, https://doi.org/10.1021/acs.jafc.7b03629.
[5] The biodiversity - Parmigiano Reggiano. https://www.parmigianoreggiano.com/product-biodiversity (accessed April 20, 2023).
[6] C.F. Balthazar, J.T. Guimarães, R.S. Rocha, T.C. Pimentel, R.P.C. Neto, M.I. B. Tavares, J.S. Graça, E.G. Alves Filho, M.Q. Freitas, E.A. Esmerino, D. Granato, S. Rodrigues, R.S.L. Raices, M.C. Silva, A.S. Sant'Ana, A.G. Cruz, Nuclear magnetic resonance as an analytical tool for monitoring the quality and authenticity of dairy foods, Trends Food Sci. Technol. 108 (2021) 84–91, https://doi.org/10.1016/j.tifs.2020.12.011.
[7] P. Scano, E. Cusano, P. Caboni, R. Consonni, NMR metabolite profiles of dairy: a review, Int. Dairy J. 90 (2019) 56–67, https://doi.org/10.1016/j.idairyj.2018.11.004.
[8] L. Laghi, G. Picone, F. Capozzi, Nuclear magnetic resonance for foodomics beyond food analysis, TrAC, Trends Anal. Chem. 59 (2014) 93–102, https://doi.org/10.1016/j.trac.2014.04.009.
[9] L. Mannina, A.P. Sobolev, V. Aru, A. Bellomaria, F. Bertocchi, B. Botta, L. R. Cagliani, A. Caligiani, F. Capozzi, D. Çela, F.C. Marincola, A. Ciampa, L. Del Cocoo, R. Consonni, C. Corsaro, M. Delfini, V. Di Tullio, F.P. Fanizzio, V. Gallo, F. Ghirga, R. Gianferri, C.R. Girellio, C. Ingallina, L. Laghi, M. Latronico, F. Longobardi, C. Luchinat, D. Mallamace, S. Mammi, W. Mandaliti, F. Marini, P. Mastrorilli, P. Mazzei, A. Miccheli, A. Micozzio, S. Miloneo, A. Mucci, R. Nepravishta, M. Paci, A. Palisi, A. Piccolo, G. Picone, N. Proietti, A. Randazzo, V. Righi, A. Rotondo, A. Salvo, F. Savorani, P. Scano, F. Sciubba, A. Trimigno, L. Tenori, E. Schievano, P. Turano, S. Vasi, D. Capitani, Chapter 5 - NMR methodologies in food analysis, in: Anal. Chem. Dev. Appl. Challenges Food Sci., Nova Science Publishers, Inc., 2017, pp. 103–156.
[10] J. Han, Y. Sun, R. Sun, T. Zhang, C. Wang, N. Jiang, Effects of freeze-thaw cycles on physicochemical properties and structure of cooked crayfish (Procambarus clarkii), Food Prod. Process. Nutr. 4 (2022) 1–14, https://doi.org/10.1186/S43014-022-00102-3.
[11] Y. Li, C. Li, M. Chen, Z. Liu, M. Zeng, Y. Hu, Sous-vide cooking endows a better microstructure for hairtail (Trichiurus lepturus) than traditional cooking: mechanisms of moisture migration, J. Food Sci. 87 (2022) 3953–3964, https://doi.org/10.1111/1750-3841.16260.
[12] P. Sun, J. Lin, X. Ren, B. Zhang, J. Liu, Y. Zhao, D. Li, Effect of heating on protein denaturation, water state, microstructure, and textural properties of Antarctic Krill (Euphausia superba) meat, Food Bioprocess Technol. 15 (2022) 2313–2326, https://doi.org/10.1007/S11947-022-02881-6.
[13] R. Consonni, L.R. Cagliani, Ripening and geographical characterization of Parmigiano Reggiano cheese by 1H NMR spectroscopy, Talanta 76 (2008) 200–205, https://doi.org/10.1016/j.talanta.2008.02.022.
[14] S. Kandasamy, J. Yoo, J. Yun, H.B. Kang, K.H. Seol, J.S. Ham, 1H HRMAS-NMR based metabolic fingerprints for discrimination of cheeses based on sensory qualities, Saudi J. Biol. Sci. 27 (2020) 1446–1461, https://doi.org/10.1016/j.sjbs.2020.04.043.
[15] C. Zhu, K. Tang, X. Lu, J. Tang, L. Laghi, An untargeted metabolomics investigation of milk from dairy cows with clinical mastitis by 1H-NMR, Foods 10 (2021) 1707, https://doi.org/10.3390/foods10081707, 1707. 10 (2021).
[16] M.S. Klein, M.F. Almstetter, G. Schlamberger, N. Nürnberger, K. Dettmer, P. J. Oefner, H.H.D. Meyer, S. Wiedemann, W. Gronwald, Nuclear magnetic resonance and mass spectrometry-based milk metabolomics in dairy cows during early and late lactation, J. Dairy Sci. 93 (2010) 1539–1550, https://doi.org/10.3168/jds.2009-2563.
[17] P. Scano, L.R. Cagliani, R. Consonni, 1H NMR characterisation of the lipid fraction and the metabolite profiles of Fossa (pit) cheese, Int. Dairy J. 90 (2019) 39–44, https://doi.org/10.1016/j.idairyj.2018.10.007.
[18] Y. Chen, W. MacNaughtan, P. Jones, Q. Yang, H. Williams, T. Foster, Selection of potential molecular markers for cheese ripening and quality prediction by NMR spectroscopy, LWT (Lebensm.-Wiss. & Technol.) 136 (2021), 110306, https://doi.org/10.1016/j.lwt.2020.110306.
[19] S. De Angelis Curtis, R. Curini, M. Delfini, E. Brosio, F. D'Ascenzo, B. Bocca, Amino acid profile in the ripening of Grana Padano cheese: a NMR study, Food Chem. 71 (2000) 495–502, https://doi.org/10.1016/S0308-8146(00)00192-8.
[20] N. Cavallini, F. Savorani, R. Bro, M. Cocchi, A metabolomic approach to beer characterization, Mol 26 (2021) 1472, https://doi.org/10.3390/molecules26051472, 26 (2021) 1472.
[21] V. Gallo, R. Ragone, B. Musio, S. Todisco, A. Rizzuti, P. Mastrorilli, S. Pontrelli, N. Intini, P. Scapicchio, M. Triggiani, A. Pascazio, C. Cobas, S. Mari, C. Garino, M. Arlorio, D. Acquotti, C. Airoldi, F. Arnesano, M. Assfalg, A. Barison, F. Benevelli, A. Borioni, L.R. Cagliani, L. Casadei, F.C. Marincola, K. Colson, R. Consonni, G. Costantino, M.A. Cremonini, S. Davalli, I. Duarte, S. Guyader, E. Hamon, M. Hegmanns, R. Lamanna, F. Longobardi, D. Mallamace, S. Mammi, M. Markus, L. R.A. Menezes, S. Milone, D. Molero-Vilchez, A. Mucci, C. Napoli, M.C. Rossi, E. Sáez-Barajas, F. Savorani, E. Schievano, F. Sciubba, A. Sobolev, P.G. Takis, F. Thomas, P. Villa-Valverde, M. Latronico, A contribution to the harmonization of non-targeted NMR methods for data-driven food authenticity assessment, Food Anal. Methods (2019), https://doi.org/10.1007/s12161-019-01664-8.
[22] M.C. Ezeanaka, J. Nsor-Atindana, M. Zhang, Online low-field nuclear magnetic resonance (LF-NMR) and magnetic resonance imaging (MRI) for food quality optimization in food processing, Food Bioprocess Technol. 129 (12) (2019) 1435–1451, https://doi.org/10.1007/S11947-019-02296-W, 2019.
[23] L. Haddad, J. Francis, T. Rizk, S. Akoka, G.S. Remaud, J. Bejjani, Cheese characterization and authentication through lipid biomarkers obtained by high-

resolution 1H NMR profiling, Food Chem. 383 (2022), 132434, https://doi.org/10.1016/j.foodchem.2022.132434.
[24] P. Mazzei, A. Piccolo, 1H HRMAS-NMR metabolomic to assess quality and traceability of mozzarella cheese from Campania buffalo milk, Food Chem. 132 (2012) 1620–1627, https://doi.org/10.1016/j.foodchem.2011.11.142.
[25] L. Shintu, F. Ziarelli, S. Caldarelli, Is high-resolution magic angle spinning NMR a practical speciation tool for cheese samples? Parmigiano Reggiano as a case study, Magn. Reson. Chem. 42 (2004) 396–401, https://doi.org/10.1002/mrc.1359.
[26] L. Shintu, S. Caldarelli, High-resolution MAS NMR and chemometrics: characterlzation of the ripening of parmigiano reggiano cheese, J. Agric. Food Chem. 53 (2005) 4026–4031, https://doi.org/10.1021/JF048141Y.
[27] C. Piras, F.C. Marincola, F. Savorani, S.B. Engelsen, S. Cosentino, S. Viale, M. B. Pisano, A NMR metabolomics study of the ripening process of the Fiore Sardo cheese produced with autochthonous adjunct cultures, Food Chem. 141 (2013) 2137–2147, https://doi.org/10.1016/j.foodchem.2013.04.108.
[28] T. Ruyssen, M. Janssens, B. Van Gasse, D. Van Laere, N. Van der Eecken, M. De Meerleer, L. Vermeiren, K. Van Hoorde, J.C. Martins, M. Uyttendaele, L. De Vuyst, Characterisation of Gouda cheeses based on sensory, analytical and high-field 1H nuclear magnetic resonance spectroscopy determinations: effect of adjunct cultures and brine composition on sodium-reduced Gouda cheese, Int. Dairy J. 33 (2013) 142–152, https://doi.org/10.1016/j.idairyj.2013.04.009.
[29] B. Popping, E. De Dominicis, M. Dante, M. Nocetti, Identification of the geographic origin of parmigiano reggiano (P.D.O.) cheeses deploying non-targeted mass spectrometry and chemometrics, Foods 6 (2017) 13, https://doi.org/10.3390/foods6020013, 13. 16 (2017).
[30] G. Rocchetti, L. Lucini, A. Gallo, F. Masoero, M. Trevisan, G. Giuberti, Untargeted metabolomics reveals differences in chemical fingerprints between PDO and non-PDO Grana Padano cheeses, Food Res. Int. 113 (2018) 407–413, https://doi.org/10.1016/j.foodres.2018.07.029.
[31] M. Li Vigni, C. Durante, S. Michelini, M. Nocetti, M. Cocchi, Preliminary assessment of parmigiano reggiano authenticity by handheld Raman spectroscopy, Foods 9 (2020) 1563, https://doi.org/10.3390/foods9111563.
[32] F. Camin, R. Wehrens, D. Bertoldi, L. Bontempo, L. Ziller, M. Perini, G. Nicolini, M. Nocetti, R. Larcher, C. H, N and S stable isotopes and mineral profiles to objectively guarantee the authenticity of grated hard cheeses, Anal. Chim. Acta 711 (2012) 54–59, https://doi.org/10.1016/j.aca.2011.10.047.
[33] E. Schievano, G. Pasini, G. Cozzi, S. Mammi, Identification of the production chain of Asiago d'Allevo cheese by nuclear magnetic resonance spectroscopy and principal component analysis, J. Agric. Food Chem. 56 (2008) 7208–7214, https://doi.org/10.1021/JF801391W.
[34] P. Ebrahimi, N. Viereck, R. Bro, S.B. Engelsen, Chemometric analysis of NMR spectra, Mod. Magn. Reson. (2017) 1–20, https://doi.org/10.1007/978-3-319-28275-6_20-1.
[35] F. Puig-Castellví, I. Alfonso, R. Tauler, Untargeted assignment and automatic integration of 1H NMR metabolomic datasets using a multivariate curve resolution approach, Anal. Chim. Acta 964 (2017) 55–66, https://doi.org/10.1016/j.aca.2017.02.010.
[36] A. Bordoni, G. Picone, E. Babini, M. Vignali, F. Danesi, V. Valli, M. Di Nunzio, L. Laghi, F. Capozzi, NMR comparison of in vitro digestion of Parmigiano Reggiano cheese aged 15 and 30 months, Magn. Reson. Chem. 49 (2011) S61–S70, https://doi.org/10.1002/mrc.2847.
[37] P. Scano, R. Anedda, M.P. Melis, M.A. Dessi, A. Lai, T. Roggio, 1H- and 13C-NMR characterization of the molecular components of the lipid fraction of pecorino sardo cheese, J. Am. Oil Chem. Soc. 88 (2011) 1305–1316, https://doi.org/10.1007/S11746-011-1797-9.
[38] L. Shintu, S. Caldarelli, Toward the determination of the geographical origin of Emmental(er) cheese via high resolution MAS NMR: a preliminary investigation, J. Agric. Food Chem. 54 (2006) 4148–4154, https://doi.org/10.1021/JF060532K.
[39] F. Savorani, G. Tomasi, S.B. Engelsen, icoshift: a versatile tool for the rapid alignment of 1D NMR spectra, J. Magn. Reson. 202 (2010) 190–202, https://doi.org/10.1016/j.jmr.2009.11.012.
[40] F. Savorani, G. Tomasi, S.B. Engelsen, Alignment of 1D NMR data using the iCoshift tool: a tutorial, in: J. van Duynhoven, P.S. Belton, G.A. Webb, H. van As (Eds.), Magn. Reson. Food Sci. Food Thought, Royal Society of Chemistry, 2013, pp. 14–24, https://doi.org/10.1039/9781849737531-00014.
[41] J. Jaumot, A. de Juan, R. Tauler, MCR-ALS GUI 2.0: new features and applications, Chemometr. Intell. Lab. Syst. 140 (2015) 1–12, https://doi.org/10.1016/j.chemolab.2014.10.003.
[42] A. de Juan, R. Tauler, Multivariate curve resolution (MCR) from 2000: progress in concepts and applications, Crit. Rev. Anal. Chem. 36 (2006) 163–176, https://doi.org/10.1080/10408340600970005.
[43] M.R. Viant, I.J. Kurland, M.R. Jones, W.B. Dunn, How close are we to complete annotation of metabolomes? Curr. Opin. Chem. Biol. 36 (2017) 64–69, https://doi.org/10.1016/j.cbpa.2017.01.001.
[44] L.W. Sumner, A. Amberg, D. Barrett, M.H. Beale, R. Beger, C.A. Daykin, T.W. M. Fan, O. Fiehn, R. Goodacre, J.L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A.N. Lane, J.C. Lindon, P. Marriott, A.W. Nicholls, M.D. Reily, J.J. Thaden, M.R. Viant, Proposed minimum reporting standards for chemical analysis: chemical analysis working group (CAWG) metabolomics standards initiative (MSI), Metabolomics 3 (2007) 211–221, https://doi.org/10.1007/s11306-007-0082-2.
[45] W. Windig, C.E. Heckler, F.A. Agblevor, R.J. Evans, Self-modeling mixture analysis of categorized pyrolysis mass spectral data with the SIMPLISMA approach, Chemometr. Intell. Lab. Syst. 14 (1992) 195–207, https://doi.org/10.1016/0169-7439(92)80104-C.
[46] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemometr. Intell. Lab. Syst. 2 (1987) 37–52, https://doi.org/10.1016/0169-7439(87)80084-9.
[47] R. Bro, A.K. Smilde, Principal component analysis, Anal. Methods 6 (2014) 2812–2831, https://doi.org/10.1039/C3AY41907J.
[48] M. Careri, S. Spagnoli, G. Panari, M. Zannoni, G. Barbieri, Chemical parameters of the non-volatile fraction of ripened Parmigiano-Reggiano cheese, Int. Dairy J. 6 (1996) 147–155, https://doi.org/10.1016/0958-6946(94)00056-5.
[49] G. Rocchetti, S. Michelini, V. Pizzamiglio, F. Masoero, L. Lucini, A combined metabolomics and peptidomics approach to discriminate anomalous rind inclusion levels in Parmigiano Reggiano PDO grated hard cheese from different ripening stages, Food Res. Int. 149 (2021), 110654, https://doi.org/10.1016/j.foodres.2021.110654.
[50] B.C. Weimer, Improving the Flavour of Cheese, Elsevier Inc., 2007, https://doi.org/10.1533/9781845693053.