



# Quantitative authenticity testing of buffalo mozzarella via $\alpha_{s1}$ -Casein using multiple reaction monitoring mass spectrometry

Yvonne Gunning, Louis K.W. Fong, Andrew D. Watson, Mark Philo, E. Kate Kemsley\*

Quadram Institute Bioscience, Colney Lane, Norwich, NR4 7UA, UK

## ARTICLE INFO

### Keywords:

Mozzarella  
Cheese  
Authenticity  
Mass spectrometry  
MRM  
Marker peptide

## ABSTRACT

We address the detection and quantitation of bovine milk in ‘buffalo’ mozzarella cheese using multiple reaction monitoring (MRM) mass spectrometry (MS). Focussing on the abundant protein  $\alpha_{s1}$ -casein, present in both species but with 10 amino acid sequence differences, we extract a list of marker peptides specific to each species. ‘Identical’ peptides, exactly the same in both species, are used for relative quantitation of  $\alpha_{s1}$ -casein in each milk type, whereas ‘similar’ peptides, present in both species but differing typically by one amino acid, are used to demonstrate relative quantitation in binary cheese mixtures. In addition, we report a pilot survey of UK supermarket and restaurant products labelled as ‘buffalo mozzarella’, finding that 2/3 of restaurant meals and supermarket pizzas are either mislabelled or adulterated.

## 1. Introduction

Mozzarella is a soft, unripened and normally white cheese, traditionally associated with Southern Italy. The production process involves inoculation of the milk base with lactic acid bacteria, coagulation with rennet to create a curd, separation of the curd from whey and kneading of the curd to produce the characteristic soft and typically rubbery cheese.

Though Italian and European regulations specify that *Mozzarella di Bufala Campana* (DOP) and *Mozzarella di latte di bufala* are made from Italian buffalo milk, in general a cheese labelled simply as ‘mozzarella’ need not be made with buffalo milk: the name mozzarella describes the cheese type rather than the base milk. However, any mozzarella labelled specifically as ‘buffalo mozzarella’ is expected to be made with buffalo milk only.

Since bovine milk is cheaper than buffalo milk, there exists the possibility of economic fraud in which some or even all of the buffalo milk in ‘buffalo’ mozzarella is substituted with bovine milk. This defrauds consumers, restaurants, wholesalers and retailers. The legitimate producers of genuine products are also disadvantaged, as their business is undercut. Difficult for consumers to detect, this type of fraud is known to occur. The presence of bovine milk in ‘buffalo’ mozzarella has been reported in small-scale surveys by several authors (Angeletti, Gioacchini, Seraglia, Piro, & Traldi, 1998; Czerwenka, Muller, & Lindner, 2010; Goncalves, Silva, Conceicao, do Egito, & Ferrao, 2017;

Goncalves et al., 2016; Locci et al., 2008; Lopparelli, Cardazzo, Balzan, Giaccone, & Novelli, 2007; Russo et al., 2012).

In Europe, buffalo milk products are covered by EU rules stating that the cow's milk casein content must not equal or exceed that of a reference sample containing 1% cow's milk (European Commission, 2018). The default analytical approach (European Reference Method, ERM) is based on casein isolation, proteolysis using plasmin, isoelectric focussing and gel evaluation, and has successfully revealed bovine content in mozzarella in commercial products (Locci et al., 2008). However, the method is perceived as laborious, species-restrictive and capable of yielding false positives (Caira et al., 2017; Cozzolino, Passalacqua, Salemi, & Garozzo, 2002; Cuollo et al., 2010; Russo et al., 2012).

Recently, DNA-based methods have been used to detect the presence of bovine milk in mozzarella (Di Domenico, Di Giuseppe, Rodriguez, & Camma, 2017; Drummond et al., 2013). However, there are challenges associated with DNA-based quantitation in milk products due to uncertainties in the DNA content of milk, and the potential impact of processing on DNA (Mayer, Burger, & Kaar, 2012).

An alternative approach is to directly interrogate the protein content of the cheese using mass spectrometry (MS). Several authors have employed matrix-assisted laser desorption/ionization (MALDI) MS, with some targeting intact proteins (Cozzolino et al., 2002; Czerwenka et al., 2010). Others used peptides, including Cuollo et al. (2010) who demonstrated quantitation using synthetic peptide internal standards,

\* Corresponding author.

E-mail addresses: [yvonne.gunning@quadram.ac.uk](mailto:yvonne.gunning@quadram.ac.uk) (Y. Gunning), [louis.fong@jic.ac.uk](mailto:louis.fong@jic.ac.uk) (L.K.W. Fong), [andrew.watson@quadram.ac.uk](mailto:andrew.watson@quadram.ac.uk) (A.D. Watson), [mark.philo@quadram.ac.uk](mailto:mark.philo@quadram.ac.uk) (M. Philo), [kate.kemsley@quadram.ac.uk](mailto:kate.kemsley@quadram.ac.uk) (E.K. Kemsley).

<https://doi.org/10.1016/j.foodcont.2019.02.029>

Received 11 December 2018; Received in revised form 19 February 2019; Accepted 20 February 2019

Available online 27 February 2019

0956-7135/ © 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

and Caira et al. (2017) in their investigation of false positives in the ERM.

MALDI MS, lacking a high-performance liquid chromatography stage, is relatively quick and simple, but peak identification and robust quantitation can be challenging. In contrast, multiple reaction monitoring (MRM) MS offers a greater level of specificity and sensitivity than MALDI. This is especially true when combined with an HPLC separation stage (LC-MS/MS) that provides an additional peptide separation step prior to the MS.

MRM MS has been used by Bernardi et al. (2015) to locate marker peptides for bovine and buffalo (also sheep and goat), although few of their markers were assigned to a parent protein, and previously published markers were overlooked. The authors demonstrated linearity between a bovine transition peak area and the percentage of bovine milk in a mozzarella cheese, but no quantitation. Russo et al. (2012) used untargeted proteomics to reveal a phosphorylated  $\beta$ -casein marker peptide for bovine, and a partner for buffalo that differs by a single amino acid. MRM MS was then used to monitor peak areas of a single transition for each peptide, leading to a cursory study of relative quantitation lacking a supporting calibration study or any consideration of the protein levels in the two cheeses themselves.

Other authors have used MRM MS in related systems, including Camerini et al. (2016) in a study of whey proteins in ricotta cheese, and Guarino et al. (2010) who investigated sheep's milk in goat and bovine cheese via casein peptides. There have also been studies of marker peptides in milk (Nardiello, Natale, Palermo, Quinto, & Centonze, 2018), and of milk allergens in food (Ansari, Stoppacher, Rudolf, Schuhmacher, & Baumgartner, 2011).

In the present work, we use MRM MS to study the adulteration of buffalo mozzarella cheese with bovine milk. The target protein is  $\alpha_{s1}$ -casein, a highly abundant milk protein that is heat stable and forms a key component of cheese products. Crucially, candidate marker peptides are predicted *in silico*, thereby avoiding a lengthy experimental marker discovery phase. A predictive approach is sufficient since our objective is to determine suitable markers rather than full peptide coverage. The presence of the candidate peptides after trypsin proteolysis is then confirmed by MRM MS.

The targeted protein  $\alpha_{s1}$ -casein is present in both bovine (P02662) and buffalo (O62823) milk; it is associated with the gene CSN1S1 but differs by 10 amino acids between species. Using MRM MS, peptides that differ by one or more amino acids which arise from these types of 'corresponding proteins' can be identified and used as quantitative markers for the components of a mixture. This generic quantitation approach, based on 'corresponding peptides' arising from 'corresponding proteins', forms the basis of our 'CPCP' strategy, which we have discussed previously in terms of myoglobin in red meats (Watson, Gunning, Rigby, Philo, & Kemsley, 2015). Using CPCP, we develop an explicit relative quantitation for mixtures of bovine and buffalo milk or cheese, based upon ratios of transition peak areas. This directly yields the relative amounts of each species in binary mixtures, in contrast to quantitative work reported elsewhere (Camerini et al., 2016; Czerwenka et al., 2010; Russo et al., 2012).

Additionally, we use identical peptides which have exactly the same sequence in both species to determine the relative level of the two  $\alpha_{s1}$ -casein proteins in the component cheeses. The relative level in the component materials provides a correction factor linking peak area ratios to relative amounts of the two cheeses in a mixture. This is needed to achieve the best accuracy in relative quantitation.

Finally, we report a limited survey of UK retail and restaurant products labelled as containing 'buffalo mozzarella'. Several products are found to contain undeclared bovine milk; some contain no buffalo at all. This strongly suggests an ongoing problem with integrity in this sector of the food supply chain.

## 2. Materials and methods

### 2.1. Software tools

Sequences P02662 (bovine, *Bos taurus*, abbreviated T in the present work) and O62823 (buffalo, *Bubalus bubalis*, abbreviated B) from the UniProtKB protein sequence database were entered into the opensource Skyline tool for creating MRM methods (MacLean et al., 2010). The enzyme was set to trypsin, with conventional cleavage rules [KR|P] and 0 missed cleavages. The minimum peptide length was set to 5, the maximum to 10, and the exclude N-terminal amino acids set to 1 to avoid initial methionine. Collision energy was set to 'ABI 4000 QTrap'. Method files were created both with and without serine phosphorylation set in Skyline, in recognition of the fact that  $\alpha_{s1}$ -casein is known to be phosphorylated.

Statistical analysis was implemented in Matlab with the Statistics and Machine Learning toolbox installed (The Mathworks, Cambridge, UK).

### 2.2. Materials

Methanol and acetonitrile were purchased from Fisher Scientific (Loughborough, UK). Urea and trypsin (from bovine pancreas, treated with L-(tosylamido-2-phenyl)ethylchloromethyl ketone (TPCK)) were purchased from Sigma-Aldrich (Gillingham, UK). Formic acid was purchased from BDH Chemicals (Poole, UK). Buffalo milk was purchased from Laverstoke Park Farm (Overton, Hants, UK), mozzarella and pizza samples were purchased from local supermarkets. Mozzarella cheese for calibration mixtures was Galbani (bovine) and Galbani di latte di bufala (buffalo), both sourced from a local supermarket. Phosphorylated and dephosphorylated  $\alpha_{s1}$ -casein from bovine milk were purchased from Sigma-Aldrich (Gillingham UK).

#### 2.2.1. Survey samples

A limited survey of locally available commercial mozzarella products was conducted using the relative quantitation approach outlined above. The products tested comprised four groups: 'supermarket mozzarella' sold as cheese only, 'supermarket pizza' carrying a 'buffalo mozzarella' claim (or in five cases merely 'mozzarella'), 'restaurant pizza' bearing a 'buffalo mozzarella' claim, and 'restaurant other' (salads and pastas) also bearing a 'buffalo mozzarella' claim.

### 2.3. Sample preparation

#### 2.3.1. Preparation of purified reference $\alpha_{s1}$ -Casein

Purified reference  $\alpha_{s1}$ -caseins were prepared from whole bovine and buffalo milk, as buffalo  $\alpha_{s1}$ -casein was unavailable for purchase. The purification method was initially tested using bovine milk and the resulting peptides compared to those from commercial bovine  $\alpha_{s1}$ -casein using LC-MS/MS. For both bovine and buffalo milk, approximately 100 mL of whole milk was defatted by centrifugation at 5000 rpm and 20 °C for 30 min and filtration through cotton wool. 45 mL of 3.3 M sodium acetate at pH 4.6 was added to the skimmed milk and mixed well then centrifuged at 7500 rpm and 4 °C for 1 h. The precipitate was collected and re-suspended in 100 mL of water then centrifuged at 7500 rpm and 4 °C for 5 min. The washed casein precipitate was collected and stored in a freezer overnight prior to freeze-drying for 3 days.

For the chromatographic separation of the caseins, two buffers were prepared: Buffer A containing 20 mM imidazole, 3.3 M urea and 0.2% v/v thioglycerol, pH 7. Buffer B containing 20 mM imidazole, 3.3 M urea, 0.2% v/v thioglycerol and 0.5M NaCl, pH 7.

Approximately 1.5 g of lyophilised casein, 4.3 g of urea, 60 mg of dithiothreitol and 50 mg of EDTA were dissolved in 12 mL of buffer A. The pH of the sample solution was adjusted to pH 7 by slowly adding 1 M NaOH, after which 500  $\mu$ L of thioglycerol was introduced. To

remove the insoluble particles of polymerised casein for the FPLC (Fast Protein Liquid Chromatography) process, the sample was centrifuged at 4000 g for 5 min, and the supernatant collected then filtered through a 0.2 µm syringe filter.

The crude extracts were loaded onto a HiLoad 16/10 Q Sepharose anion exchange column attached to an ÄKTA FPLC automated liquid chromatography system. The column was equilibrated with buffer A then the caseins were eluted using a linear gradient (0–100% buffer B) over a period of 40 min at a flow rate of 3 mL/min. The eluent was monitored for protein at an absorbance of 280 nm. The  $\alpha_{s1}$ -casein fractions were pooled and stored at 4 °C.

Before proteolysis and analysis by LC-MS, buffer exchange of the  $\alpha_{s1}$ -casein fractions was carried out using a Sartorius Vivaspin Turbo 15 device with molecular weight cut off at 3000. The device was activated by centrifugation at 4000g using water until 2 mL of water remained. The  $\alpha_{s1}$ -casein fractions were then added to the activated device and centrifuged at 4000 g until approximately 2 mL of liquid remained. The concentrated  $\alpha_{s1}$ -casein was diluted with 25 mM ammonium bicarbonate and centrifuged. The concentration and dilution steps were repeated four times until the original buffer was replaced with 25 mM ammonium bicarbonate. 1 mL of the  $\alpha_{s1}$ -casein in 25 mM ammonium bicarbonate solution was used for proteolysis.

### 2.3.2. Defatting and extraction of whole milk samples

Whole milk samples (100 µL) were extracted by heating with 100 mM ammonium bicarbonate (100 mL) at 60 °C for 15 min. The mixture was centrifuged at 6500 g at 4 °C for 20 min. 200 µL of the supernatant was transferred to an Eppendorf tube and made up to 1 mL with 25 mM ammonium bicarbonate.

### 2.3.3. Preparation and extraction of raw and cooked mozzarella

For single species mozzarella, 1 g of sample was finely sliced and placed into a 15 mL Corning tube. Mozzarella mixtures (0%, 1%, 3%, 5%, 10%, 20% and 100% bovine in buffalo) were prepared by finely slicing samples of single-species cheese and combining the appropriate accurately weighed amounts, totalling 1g, into 15 mL Corning tubes.

Cooked samples of single species mozzarella (10g) were prepared by placing the sample on a domestic metal baking tray in a pre-heated oven at 200 °C for 15 min. The samples were allowed to cool, and 1 g excised for subsequent analysis.

100 mM Ammonium bicarbonate (5 mL) was added to the mozzarella samples (single species, mixtures, and cooked samples; 1 g), mixed well, then extracted by heating at 60 °C for 60 min. After extraction, aliquots (1 mL) of the mixture were transferred to Eppendorf tubes and centrifuged at 6500 g and 4 °C for 20 min. 200 µL of the supernatant was transferred to an Eppendorf tube and made up to 1 mL with 25 mM ammonium bicarbonate.

### 2.3.4. Proteolysis of milk and cheese samples

The extracted samples were heated in a hot block at 95 °C for 30 min, then cooled to room temperature, and urea was added to a final concentration of 0.5 M. Trypsin solution (1 µg/µL) was then added in a ratio of 1:30 enzyme:substrate by weight. The sample was gently

vortexed and incubated at 37 °C overnight.

The digested sample was diluted 1:2 with water and desalted using a Strata-X 33µ polymeric reversed-phase cartridge filled with 30 mg/mL RP material (Phenomenex, Macclesfield, UK). The cartridge was washed and activated with 1 mL of methanol followed by 1 mL of 1% formic acid. The entire sample (approximately 3 ml) was loaded on to the cartridge which was washed with 1 mL of 5% methanol/1% formic acid in water. The peptides were eluted into an Eppendorf tube containing 5 µL DMSO with 1 mL acetonitrile/water (90:10; 0.1% formic acid). The desalted sample was dried by in a centrifugal evaporator (180 min at 50 °C) and then redissolved in 250 µL of acetonitrile/water (3:97; 0.1% formic acid) for subsequent LC-MS/MS analysis.

## 2.4. LC-MS/MS analyses

High-performance liquid chromatography/tandem mass spectrometric (HPLC-MS/MS) analyses of the digested proteins were performed using an Agilent 1200 rapid resolution LC system (Stockport, UK) coupled to an AB Sciex 4000 QTrap triple-quadrupole mass spectrometer (Warrington, UK). Chromatographic separations were conducted on an XB C18 reversed-phase capillary column (100 × 2.1 mm, 2.6 µm particle size) maintained at 40 °C (Phenomenex, Macclesfield, UK) with a flow rate of 300 µL/min. The gradient profile used consisted of a binary gradient from 100% A (water + 0.1% formic acid) and 0% B (acetonitrile + 0.1% formic acid) to 38% B over 38 min, increasing to 100% B at 39 min and held for 1 min. Column re-equilibration was for a further 8 min. The injection volume was 1 µL. A single MS/MS methods file containing parameters appropriate for  $\alpha_{s1}$ -casein from both species was constructed based on Skyline predictions and analysis of reference material. Mass data acquisitions were made by Analyst 1.6.2. Software (AB Sciex). Eluted peptides were detected by positive electrospray in scheduled dynamic monitoring mode with a scan time of 2 s and a nominal retention time window of ± 50 s. Each peptide was monitored by the four most intense MRM transitions as determined by peak height. Turbospray source settings were operated with a curtain gas of 25 psi, desolvation gas (GS1) of 50 psi, and sheath gas (GS2) of 20 psi. The source temperature was 550 °C.

## 3. Results

### 3.1. Proteolysis of $\alpha_{s1}$ -Casein

Proteolysis was performed using a conventional protocol based on heat with urea. Fig. 1 shows the two  $\alpha_{s1}$ -casein protein sequences P02662 (bovine) and O62823 (buffalo), with bovine above and buffalo below. The sequences are aligned, with the ten sequence differences between the two highlighted in yellow. By inspection there are several candidate peptides that could act as species markers able to differentiate between bovine and buffalo. Having predicted the tryptic cleavage sites, the strategy is to experimentally detect species-specific markers suitable for species determination, then determine which, if any, offer a route to relative quantitation.

```

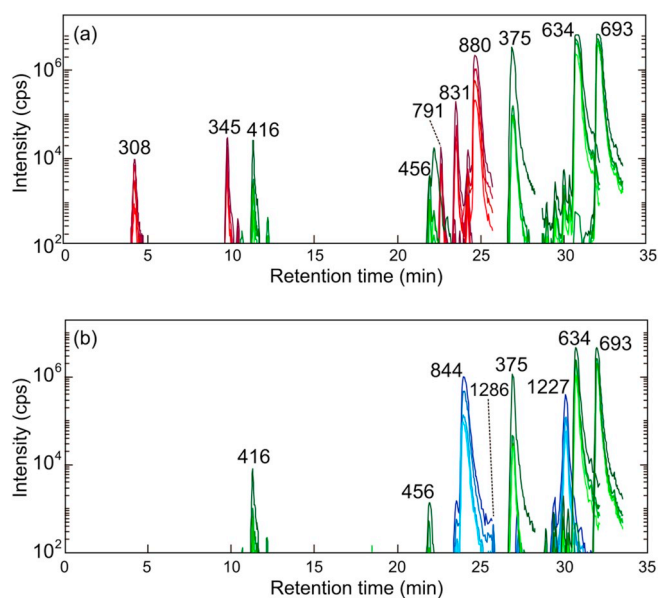
P02662 |MK*LLILTCLVAVALARPK*HP*IK*HQGLPQ*EVLNENLLR*FFVAPFPPEVFGK*EK*VNELSK*DIGESTEDQAMEDIK*
O62823 |MK*LLILTCLVAVALARPK*Q*PIK*HQGLPQ*EVLNENLLR*FFVAPFPPEVFGK*EK*VNELST*DIGESTEDQAMEDIK*

P02662 |QMEAESISSSEEIVPNSVEQK*HIQK*EDVPSER*YLGYLEQLLR*LK*K*YK*VPQLEIVPNSAEER*LHSMK*EGIHAQQK*
O62823 |QMEAESISSSEEIVPNSVEQK*HIQK*EDVPSER*YLGYLEQLLR*LK*K*YN*VPQLEIVPNLAEEQ*LHSMK*EGIHAQQK*

P02662 |EPMIGVNQELAYFYPELFR*QFYQLDAYPSGAWYYVPLGTQY*TDAPSFSDIPNPIGSENSEK*TTMPLW
O62823 |EPMIGVNQELAYFYPQLFR*QFYQLDAYPSGAWYYVPLGTQY*PDAPSFSDIPNPIGSENSGK*TTMPLW

```

Fig. 1.  $\alpha_{s1}$ -casein sequences, showing bovine (P02662) and buffalo (O62823) on respectively the upper and lower lines of each text block. Coloured sequence strings indicate the marker peptides listed in Table 1 (red for bovine, blue for buffalo, green for peptides identical in both). The red and blue dots indicate conventional tryptic cleavages. The initial methionine is included to give consistency with published protein libraries. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 2.** MRM transition intensities versus retention time for purified  $\alpha_{s1}$ -casein from (a) bovine milk and (b) buffalo milk. The data was acquired in dynamic mode. The numerical labels indicate precursor ion  $m/z$  values (daltons) for charge  $z = 2$ . Most ‘peaks’ are actually a cluster of peaks corresponding to four transitions from the same precursor ion. Transitions specific to bovine are shown in red, those specific to buffalo in blue, and those appearing in both species in green. Note that the intensity axis is  $\log_{10}$  (see text, also Supplementary Information). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

### 3.2. Mass spectrometry of reference $\alpha_{s1}$ -Casein

We performed dynamic MRM MS on reference  $\alpha_{s1}$ -casein purified from commercial bovine and buffalo milk samples to locate candidate species marker peptides. The resulting chromatograms are displayed in Fig. 2. These plots of transition intensities versus retention times are displayed using a  $\log_{10}$  intensity axis rather than the conventional linear intensity axis. This has the advantage of giving increased emphasis to less intense features, since a logarithmic scale better captures the dynamic range of the data. It also exposes imperfections in the data such as peak truncation due to retention time windowing. The use of logarithmic intensity scales is discussed in the Supplementary Information (section 1).

In Fig. 2, panel (a) shows bovine  $\alpha_{s1}$ -casein features and panel (b) those from buffalo. Red peaks are specific to bovine, blue specific to buffalo, and green appear in both species. Each feature is typically a set of four peaks corresponding to four different transitions preselected according to their superior intensity and signal quality from the full set of available transitions. This figure confirms that several species marker peptides can be experimentally detected.

Candidate marker peptides are summarised in Table 1. This lists  $m/z$  values (integer, in daltons) for the  $z = 2$  charge state, and the fragment masses of the four selected transitions. The fragments are listed according to transition intensity, with the most intense first. Three of the peptides contain one or two phosphorylated serines (indicated by S (Pho) in the sequences). The most useful species markers, in terms of signal quality, signal strength and number of detectable transitions, are 308, 345, 791, 831 and 880 for bovine, and 844, 1227 and 1286 for buffalo. The bovine-specific peptide 880 (HQGLPQEVLENLLR) and the buffalo-specific peptide 844 (HQGLPQGVLENLLR) form a CPCP pair, differing by a single amino acid (indicated in bold) and are promising candidates for relative quantitation. Both have been noted as species markers previously (Cuollo et al., 2010; Nardiello et al., 2018; Russo et al., 2012). Additionally, peptides 634 and 693, appearing in

both species, have been noted as milk casein markers in processed foods (Ansari et al., 2011).

### 3.3. Cheese samples: relative $\alpha_{s1}$ -Casein levels

The CPCP strategy uses ratios of transition peak areas to give relative levels of peptides. The ratio is used as a proxy for the relative level of protein (in the present work,  $\alpha_{s1}$ -casein) and hence the relative amounts of two cheeses in a mixture. However, if the  $\alpha_{s1}$ -casein levels of the two cheese types differ, it follows that identical amounts of the two cheeses will give a transition peak area ratio different from one. For this reason, the difference in  $\alpha_{s1}$ -casein levels in the two cheese types should be accounted for via a correction factor when converting CPCP peak area ratios into relative amounts of the two cheeses in a mixture.

We define  $r_{\alpha_{s1}} = (\text{bovine } \alpha_{s1})/(\text{buffalo } \alpha_{s1})$  to be the relative levels of  $\alpha_{s1}$ -casein in the bovine and buffalo mozzarellas. We have used MRM MS to measure  $r_{\alpha_{s1}}$  via peptides that share an identical sequence in  $\alpha_{s1}$ -casein from both species. These peptides are listed in Table 1, together with those that differ between bovine and buffalo.

To show how  $r_{\alpha_{s1}}$  can be extracted from the data for these identical peptides in a cheese mixture series, we first define the bovine cheese percent weight-for-weight (%w/w) as  $100M_T/(M_T + M_B)$ . Here,  $M_T$  and  $M_B$  are the masses of bovine and buffalo mozzarellas in a mixture. For an  $\alpha_{s1}$ -casein peptide having the same sequence in both species, a plot of a transition peak area from that peptide (on the vertical axis) versus %w/w of bovine mozzarella in a mixture (on the horizontal axis) would yield a horizontal line if the levels of  $\alpha_{s1}$ -casein were the same in both cheeses. This is the special case where  $r_{\alpha_{s1}} = 1$  (see Supplementary Information, Section 2, Fig. 3(a)).

However, if for example bovine mozzarella contained higher levels of  $\alpha_{s1}$ -casein than buffalo,  $r_{\alpha_{s1}} > 1$ , then the plot would show a positive slope. This is because an increasing proportion of bovine mozzarella would contribute more  $\alpha_{s1}$ -casein to the mix than the buffalo mozzarella it displaces. The ratio  $r_{\alpha_{s1}}$  is given by where  $m$  is the plot gradient and  $A_0$  is the vertical axis intercept. Note that lower abundance peptides exhibit lower values of  $A_0$  and therefore correspond to a smaller gradient  $m$ .

For each peptide with an identical sequence in  $\alpha_{s1}$ -casein from both species (Table 1), we recorded 4 transitions from 6 nominal bovine mozzarella levels: 0%, 1%, 3%, 5%, 10% and 20%. Each mixture was created, extracted and digested independently 3 times (biological replicates) and each of these in turn analysed 3 times by LC-MS/MS (technical replicates) to give 9 replicates in total.

To perform the data analysis, we first note that the measured peak intensities for each peptide and mixture series are correlated across the four transitions (data not shown). The four transitions may be viewed as the outputs of four ‘counters’ of different efficiencies that monitor the same parent peptide. To facilitate subsequent regression analysis embracing all four transitions per peptide, we first scale the data according to a scale factor  $f_j = \frac{\bar{x}_g}{\bar{x}_j}$ . Here,  $\bar{x}_g$  denotes the grand mean across all data for the four transitions for the current peptide, and  $\bar{x}_j$  denotes the mean for all data for transition  $j$ . For the peptide under consideration, the scale factor  $f_j$  is applied to all data for transition  $j$ . The scale factor preserves the ratio  $\frac{m}{A_0}$ , necessary because the value of  $r_{\alpha_{s1}}$  is an intrinsic property of the sample, which should therefore be the same for all transitions from that sample.

To capture the structure within the data, where that structure arises from different biological replicates and peptide transitions, we have used linear multilevel modelling, a statistical approach for handling grouped data. The different extracts and transitions are expressed as two categorical grouping variables, %w/w of bovine mozzarella the continuous predictor variable, and peak areas as the response variable. This regression approach yields estimates for  $m$  and  $A_0$ , giving  $r_{\alpha_{s1}} = 1.06 \pm 0.27$  for the samples used in this study based on the three most intense peptides, YLGYLEQLLR ( $m/z = 634$ ), TTMLPW ( $m/z$

**Table 1**

Table of  $\alpha_1$ -casein-derived marker peptides. Columns denote the precursor  $m/z$  (daltons); selected fragment masses (daltons); retention time  $R_t$  (minutes); species code (T = present in bovine, B = present in buffalo, and TB = present in both); and precursor peptide amino acid sequence. S(Pho) indicates phosphorylated serine. Bold underlined text indicates the markers used for relative quantitation. In the final two columns, dots indicate which of the peptides are expected to be present in sheep and goat's milk.

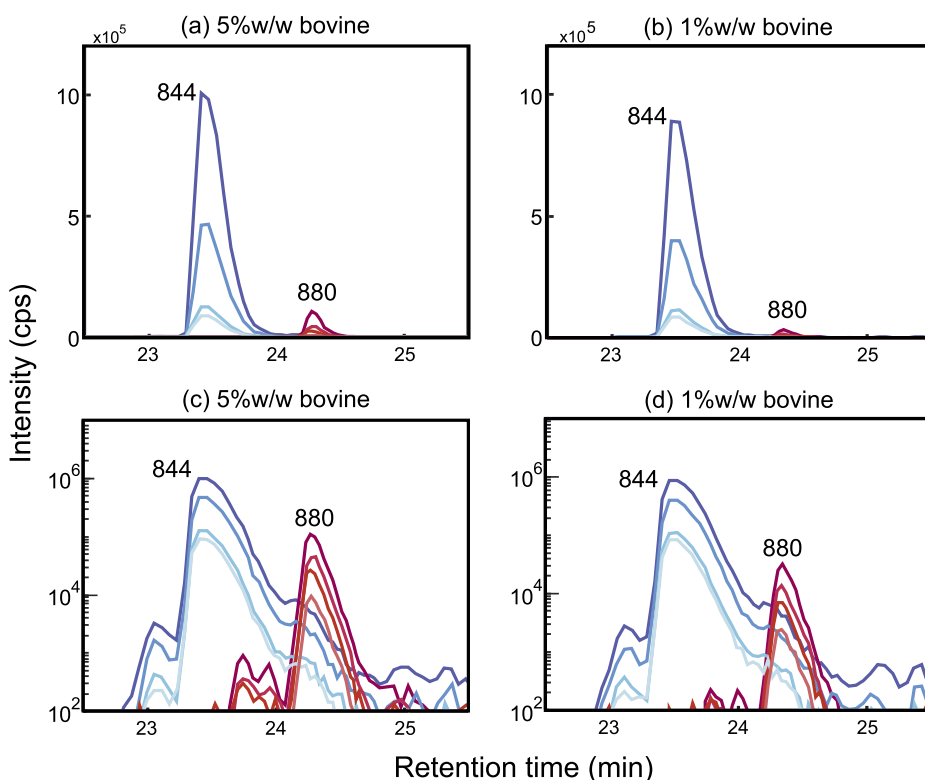
$m/z$	fragments	$R_t$	species	sequence	goat	sheep
308.2	(365, 502, 147, 278)	4.5	T	LHSMK		
345.2	(590, 476, 234, 347)	10.0	T	VNELSK		
374.7	(205, 415, 546, 318)	26.3	TB	TTMPLW	•	
416.2	(488, 391, 587, 175)	11.4	TB	EDVPSER	•	
456.2	(568, 471, 667, 304)	22.9	TB	EDVPS(Pho)ER	•	•
634.4	(992, 771, 658, 935)	29.8	TB	YLGYLEQLLR	•	•
692.9	(920, 992, 676, 1091)	31.1	TB	FFVAPFPEVFGK		
790.9	(802, 901, 1015, 1257)	22.0	T	VPQLEIVPNSAEER		
830.9	(882, 981, 1094, 785)	22.9	T	VPQLEIVPNS(Pho)AEER		
844.5	(1253, 1423, 1028, 872)	23.3	B	<b><u>HQGLPOGVLNENLLR</u></b>		
880.5	(1325, 1495, 872, 971)	24.0	T	<b><u>HQGLPOEVLNENLLR</u></b>		
1226.6	(1397, 278, 1496, 1609)	29.2	B	YNVPQLEIVPNLAEEQLHSMK		
1286.0	(375, 635, 260, 504)	25.0	B	VNELS(Pho)TDIGS(Pho)ESTEDQAMEDIK		

$z = 375$ ) and FFVAPFPEVFGK ( $m/z = 693$ ). The quoted error is the 95% confidence interval. The error in the estimation of  $r_{\text{cs1}}$  is dominated by the error in the estimation of the gradient  $m$ . Further, since  $r_{\text{cs1}}$  is close to unity, the value of  $\frac{m}{A_0}$  is close to zero, implying that the most reliable estimates of  $r_{\text{cs1}}$  come from those peptides which yield the overall strongest signals.

On a practical note, this shows that in a CPCP relative quantitation mixture series, a value for  $r_{\text{cs1}}$  can also be extracted without involving any extra experimental work. All that is required is to include the relevant transitions for identical peptides in the MRM MS method file. The relative levels of protein in the two component cheeses can be calculated using the same mixture samples in which the relative levels of the two cheeses are being determined.

### 3.4. Cheese samples: relative quantitation

The CPCP pair 844 (bovine) and 880 (buffalo) can be used for

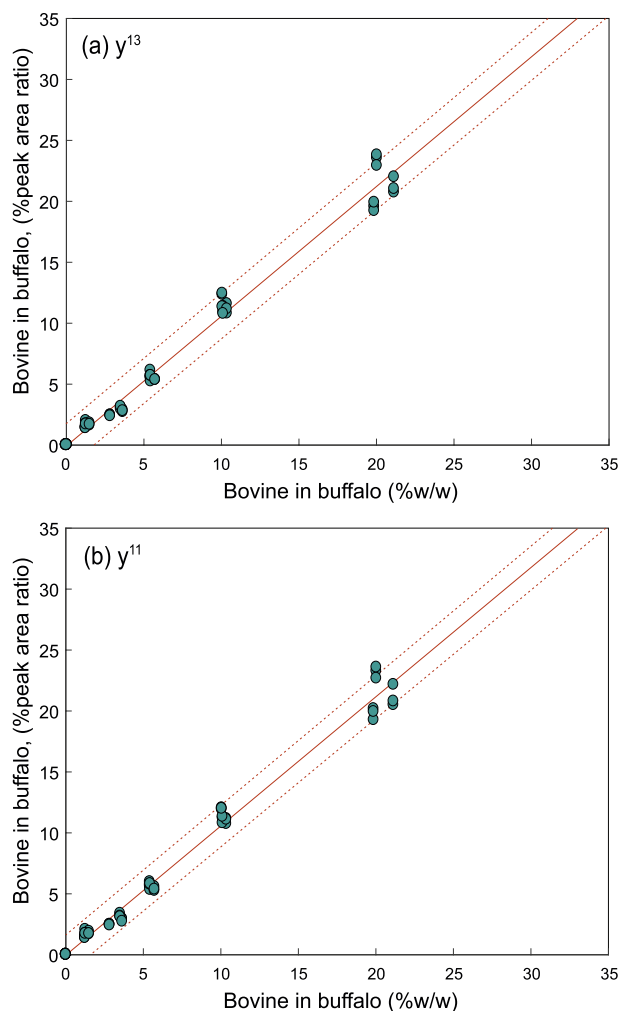


**Fig. 3.** MRM transition intensities for the buffalo marker peptide 844 (blue) and the bovine marker peptide 880 (red) showing four transitions for each. The data was acquired in dynamic mode from mixtures of commercial mozzarella cheeses. Panels (a) and (b) respectively show the 5%w/w and 1% w/w bovine data using linear intensity axes; there is a proportionate decrease in the size of the 880 bovine marker. Panels (c) and (d) show the same data but presented on log10 intensity axes. Here, the individual transitions and signal quality of the bovine 880 signal are better appreciated. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

relative quantitation. Fig. 3 shows the four recorded transitions for each peptide derived from the same cheese mixtures and experimental runs described above for the determination of  $r_{\text{cs1}}$ . Panels (a) and (b) display the data using linear intensity axes, whereas panels (c) and (d) display the same data using log10 intensity axes. The intensity reduction as the bovine level changes from 5% to 1% is apparent in both pairs of plots. Additionally, on the log10 axes, both 5% and 1% are clearly seen to comprise four transitions.

For relative quantitation, transitions involving peptide fragmentation as distant as possible from the region of sequence difference typically give the most accurate results. This is because the fragmentation probability may be affected by nearby sequence differences. This in turn would manifest as differences in peak area ratios that are unrelated to levels of the respective peptides and serve only to confuse the result.

Fig. 4 shows relative quantitation for (a) the  $y^{13}$  transition (T(880  $\rightarrow$  1495) and B(844  $\rightarrow$  1423)), and (b) the  $y^{11}$  transition (T(880  $\rightarrow$  1325) and B(844  $\rightarrow$  1253)). In both panels, the horizontal



**Fig. 4.** Calibration plots of bovine mozzarella in buffalo mozzarella, as percent by weight versus percent by transition peak area. The blue circles show data points (3 biological replicates, each with three technical replicates) with regression lines in red and inverse prediction intervals (95% confidence level) in dotted red. Panel (a) is for the  $y^{13}$  transition pair, with T(880  $\rightarrow$  1495) and B(844  $\rightarrow$  1423). Panel (b) is for the  $y^{11}$  transition pair, with T(880  $\rightarrow$  1325) and B(844  $\rightarrow$  1253). Slopes of the regression lines are both 1.06 with a standard error of 0.02. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

axis shows %w/w bovine mozzarella ( $100M_T/(M_T + M_B)$ , from above), and the vertical axis shows the peak area ratio corrected for the difference in  $\alpha_{s1}$ -casein levels in the two cheeses. This is calculated from  $100A_T/(A_T + r_{\alpha s1}A_B)$ , where  $A_T$  and  $A_B$  are bovine and buffalo transition peak areas respectively, and  $r_{\alpha s1}$  takes the value 1.06 as described above.

The plots show a linear and reproducible calibration over a test range of 0–20 %w/w bovine mozzarella in buffalo mozzarella, consistent with the assumptions of the CPCP approach. The regression line gradients are both  $1.06 \pm 0.02$  (standard error), close to the expected value of unity due to the use of the correction factor.

Note that achieving straight line calibration curves of gradient unity is predicated on the use of an appropriate correction factor for the samples at hand (here,  $1.06 \pm 0.27$ ). If a substantively incorrect  $r_{\alpha s1}$  value had been used, then the curvature and gradients of the regression lines in Fig. 4 would deviate from straight lines and unity respectively. More particularly, if the value of  $r_{\alpha s1}$  used is greater than the ‘true’ value, then the data points would fall on a curve arching below the straight line of gradient unity; inappropriately small values of  $r_{\alpha s1}$  would cause the curve to arch above (see Supplementary Information,

Section 2, Fig. 3(b)). Therefore, the fact that the data points as depicted in Fig. 4 fall on a straight line provides a consistency check for the value of  $r_{\alpha s1}$  (Czerwenka et al., 2010).

### 3.5. Pilot surveillance study of retail mozzarella products

Since some of the survey products involved cooked mozzarella, we first studied the impact of cooking (15 min at 200 °C, a level comparable with pizza cooking) on the presence of species marker peptides. There was no impairment to the quality of the data in cooked as opposed to uncooked mozzarella (see Supplementary Information, Section 3). The ‘supermarket pizza’ samples typically included grated bovine cheddar cheese in addition to ‘buffalo mozzarella’. Whereas the grated cheddar was distributed across the pizza, the mozzarella was present as a limited number of discrete pieces on the top of the pizza. These tended to merge with the grated cheddar upon cooking. The presence of the bovine cheddar made sampling of the mozzarella component challenging, especially in the cooked product, and great care had to be taken to avoid contamination of the mozzarella fraction with the bovine cheddar. Restaurant pizzas did not appear to contain any additional grated cheese.

The outcomes of the survey are listed in Table 2. In the table, a sample is noted as containing bovine milk if at least four bovine-specific marker peptides (from the set  $m/z = 308, 345, 791, 831$  and 880) were detected via at least three transitions and at the expected retention time. Fig. 5 shows an example of a restaurant product sold as containing only buffalo mozzarella, demonstrating how even at low levels (estimated at  $\sim 4\%$ w/w in the relative quantitation for this sample) the signal from bovine milk is clear and unambiguous.

All the survey products sold as mozzarella cheese were judged to be labelled correctly with regards to the declared species. Samples 1 to 8 inclusive carried a specific ‘buffalo mozzarella’ claim, whereas samples 9 to 13 inclusive carried simply a ‘mozzarella’ claim with ‘milk’ on the list of ingredients; we have interpreted the latter as meaning bovine milk. The results demonstrate that species determination using our identified markers is robust when confronted with real-world samples.

Of the other three product types, overall 65% were deemed suspicious: these cases are indicated by bold/underlining in Table 2. Sample 20 is the same restaurant menu item as sample 19 but purchased at least one week later. Likewise, 22 is the same menu item as 21 but purchased on a later occasion. In both cases, one sample was found to contain mixed species mozzarella, whereas the repeat purchase was as described on the menu. In addition, three samples declared as ‘buffalo mozzarella’ (24, 25 and 30), contained no detectable buffalo mozzarella at all, only cheese of bovine origin.

Using relative quantitation, it is possible to add more depth to the surveillance study. Fig. 6 shows an indicative quantitation plot of the surveillance samples using the  $y^{13}$  transitions, bovine T(880  $\rightarrow$  1495) and buffalo B(844  $\rightarrow$  1423). The samples are plotted in order of increasing percentage peak area  $100A_T/(A_T + A_B)$ , setting to unity the factor  $r_{\alpha s1}$  that corrects for different  $\alpha_{s1}$ -casein in the two cheeses. The dotted black line indicates the 1% threshold.

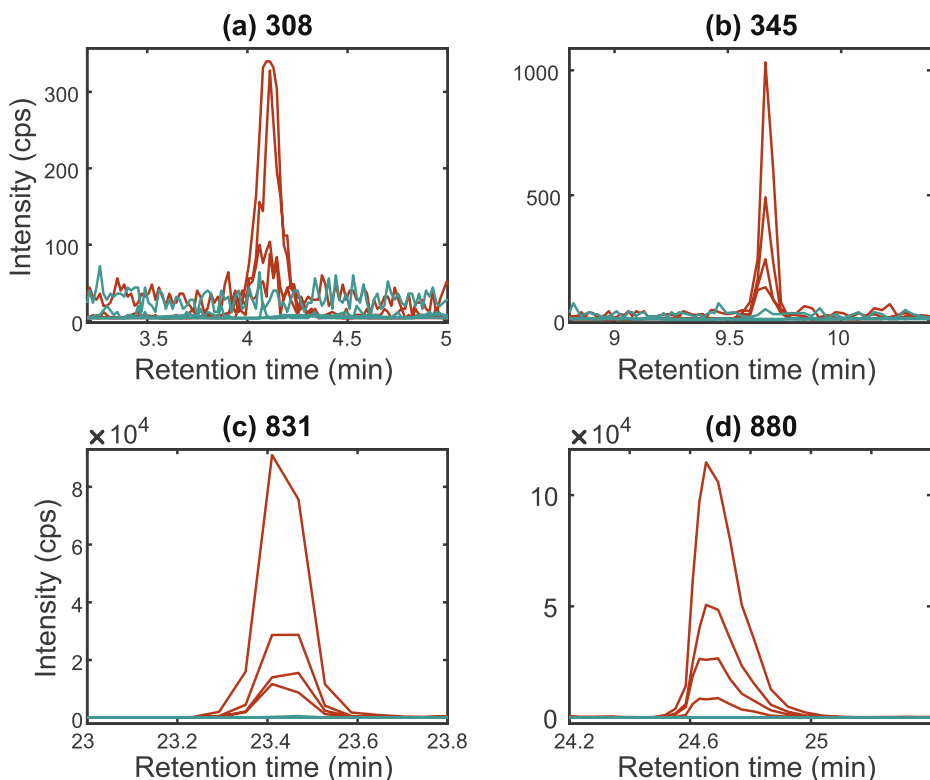
Some degree of relative quantitation provides more detail on the nature of the suspicious results. At the most extreme level, three restaurant products labelled as buffalo mozzarella contained no buffalo at all, only bovine mozzarella. This may indicate intentional fraudulent substitution, although it could also be due to accidental mislabelling by the restaurant supplier, or accidental use of the wrong product by the restaurateur.

There are five products for which the percentage peak area ratio ranges from approximately 4%–60%. Mozzarella tends to be added to pizzas and restaurant products as discrete pieces, rather than grated or powdered. It is difficult to imagine the restaurateur or the supermarket supplier blending bovine with buffalo mozzarella to create a mixed-species cheese. The inference is that this mixing of species is occurring in the form of milks, further up the supply chain, potentially an example

**Table 2**

Table of mozzarella surveillance survey results. The rightmost three columns indicate the results of testing according to the marker peptides described in the text. Sample 20 is a second purchase of the same menu item as sample 19; likewise, 22 is a repeat purchase of 21. Samples 14, 15 and 16 were tested both uncooked and cooked. Bold and underlined code numbers indicate samples where there is a conflict between the test results and the product description.

No.	Sample description	Declared buffalo mozzarella	Cooked?	Outcome		
				100% buffalo	100% bovine	Mixed species
<b>Supermarket mozzarella:</b>						
1	Mozzarella (Di Bufala Campana)	✓	×	✓		
2	Mozzarella (100% Latte di Bufala)	✓	×	✓		
3	Organic Buffalo Mozzarella	✓	×	✓		
4	Mozzarella (Di Bufala Campana)	✓	×	✓		
5	Mozzarella (Di Bufala Campana)	✓	×	✓		
6	Mozzarella (Di Bufala Campana)	✓	×	✓		
7	Mozzarella (Di Bufala Campana)	✓	×	✓		
8	Mozzarella (Di Bufala Campana)	✓	×	✓		
9	Mini Mozzarella	×	×		✓	
10	Mozzarella	×	×		✓	
11	Italian Mozzarella	×	×		✓	
12	Italian Mozzarella	×	×		✓	
13	Mozzarella	×	×		✓	
<b>Supermarket pizza:</b>						
<b>14</b>	Buffalo mozzarella & tomato pizza	✓	Both			✓
<b>15</b>	Buffalo mozzarella & tomato pizza	✓	Both			✓
<b>16</b>	Buff mozzarella & rocket pesto pizza	✓	Both			✓
<b>17</b>	Buff mozzarella & tomato pizza	✓	×			✓
<b>18</b>	Margherita pizza	✓	×			✓
<b>Restaurant Pizza:</b>						
<b>19</b>	Pizza	✓	✓			✓
20	Pizza (repeat purchase of 19)	✓	✓	✓		
<b>21</b>	Pizza	✓	✓			✓
22	Pizza (repeat purchase of 21)	✓	✓	✓		
23	Pizza	✓	✓	✓		
<b>24</b>	Pizza	✓	✓		✓	
<b>25</b>	Pizza	✓	✓		✓	
<b>Restaurant Other:</b>						
26	Salad	✓	×	✓		
27	Pasta dish	✓	×	✓		
<b>28</b>	Starter	✓	×			✓
29	Salad	✓	×	✓		
<b>30</b>	Starter	✓	×		✓	



**Fig. 5.** Four MRM transitions from each of four bovine-specific peptides, in red, arising from a restaurant pizza sold as containing buffalo mozzarella only. The *m/z* values of the four peptides are indicated above each panel. Quantitation using the  $y^{13}$  transitions (T(880 → 1495) with B(844 → 1423), buffalo data not shown) puts the bovine level estimate at ~4 %w/w (see Fig. 6 below). The cyan lines indicate the same signal regions acquired for quality control blanks run immediately before each restaurant sample. Note the linear intensity axes are scaled to the maximum peak height in each panel. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)





- (2017). Eventual limits of the current EU official method for evaluating milk adulteration of water buffalo dairy products and potential proteomic solutions. *Food Chemistry*, 230, 482–490. <https://doi.org/10.1016/j.foodchem.2017.03.055>.
- Camerini, S., Montepeloso, E., Casella, M., Crescenzi, M., Marianella, R. M., & Fuselli, F. (2016). Mass spectrometry detection of fraudulent use of cow whey in water buffalo, sheep, or goat Italian ricotta cheese. *Food Chemistry*, 197, 1240–1248. <https://doi.org/10.1016/j.foodchem.2015.11.073>.
- Cozzolino, R., Passalacqua, S., Salemi, S., & Garozzo, D. (2002). Identification of adulteration in water buffalo mozzarella and in Ewe cheese by using whey proteins as biomarkers and matrix-assisted laser desorption/ionization mass spectrometry. *Journal of Mass Spectrometry*, 37(9), 985–991. <https://doi.org/10.1002/jms.358>.
- Cuollo, M., Caira, S., Fierro, O., Pinto, G., Picariello, G., & Addeo, F. (2010). Toward milk speciation through the monitoring of casein proteotypic peptides. *Rapid Communications in Mass Spectrometry*, 24(11), 1687–1696. <https://doi.org/10.1002/rcm.4564>.
- Czerwenka, C., Muller, L., & Lindner, W. (2010). Detection of the adulteration of water buffalo milk and mozzarella with cow's milk by liquid chromatography-mass spectrometry analysis of beta-lactoglobulin variants. *Food Chemistry*, 122(3), 901–908. <https://doi.org/10.1016/j.foodchem.2010.03.034>.
- Di Domenico, M., Di Giuseppe, M., Rodriguez, J. D. W., & Camma, C. (2017). Validation of a fast real-time PCR method to detect fraud and mislabeling in milk and dairy products. *Journal of Dairy Science*, 100(1), 106–112. <https://doi.org/10.3168/jds.2016-11695>.
- Drummond, M. G., Brasil, B. S. A. F., Dalsecco, L. S., Brasil, R. S. A. F., Teixeira, L. V., & Oliveira, D. A. A. (2013). A versatile real-time PCR method to quantify bovine contamination in buffalo products. *Food Control*, 29(1), 131–137. <https://doi.org/10.1016/j.foodcont.2012.05.051>.
- European Commission (2018). Commission Implementing Regulation (EU) 2018/150 of 30 January 2018 amending Implementing Regulation (EU) 2016/1240 as regards methods for the analysis and quality evaluation of milk and milk products eligible for public intervention and aid for private storage. *Official Journal of the European Union*, L26, 14–47.
- Feligni, M., Bonizzi, I., Buffoni, J. N., Cosenza, G., & Ramunno, L. (2009). Identification and quantification of alpha(S1), alpha(S2), beta, and kappa-caseins in water buffalo milk by reverse phase-high performance liquid chromatography and mass spectrometry. *Journal of Agricultural and Food Chemistry*, 57(7), 2988–2992. <https://doi.org/10.1021/jf803653v>.
- Goncalves, B. R. F., Silva, G. D., Conceicao, D. G., do Egito, A. S., & Ferrao, S. P. B. (2017). Buffalo mozzarella chemical composition and authenticity assessment by electrophoretic profiling. *Semina: Ciências Agrárias*, 38(4), 1841–1851. <https://doi.org/10.5433/1679-0359.2017v38n4p1841>.
- Goncalves, B. R. F., Silva, G. D., Pontes, S. F. O., Fontan, R. D. I., do Egito, A. S., & Ferrao, S. P. B. (2016). Evaluation of the peptide profile with a view to authenticating buffalo mozzarella cheese. *International Journal of Food Science and Technology*, 51(7), 1586–1593. <https://doi.org/10.1111/ijfs.13128>.
- Guarino, C., Fuselli, F., La Mantia, A., Longo, L., Faberi, A., & Marianella, R. M. (2010). Peptidomic approach, based on liquid chromatography/electrospray ionization tandem mass spectrometry, for detecting sheep's milk in goat's and cow's cheeses. *Rapid Communications in Mass Spectrometry*, 24(6), 705–713. <https://doi.org/10.1002/rcm.4426>.
- Locci, F., Ghiglietti, R., Francolino, S., Iezzi, R., Oliviero, V., Garofalo, A., et al. (2008). Detection of cow milk in cooked buffalo Mozzarella used as Pizza topping. *Food Chemistry*, 107(3), 1337–1341. <https://doi.org/10.1016/j.foodchem.2007.09.040>.
- Lopparelli, R. M., Cardazzo, B., Balzan, S., Giaccone, V., & Novelli, E. (2007). Real-time TaqMan polymerase chain reaction detection and quantification of cow DNA in pure water buffalo mozzarella cheese: Method validation and its application on commercial samples. *Journal of Agricultural and Food Chemistry*, 55(9), 3429–3434. <https://doi.org/10.1021/jf0637271>.
- MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., ... MacCoss, M. J. (2010). Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, 26(7), 966–968. <https://doi.org/10.1093/bioinformatics/btq054>.
- Mayer, H. K., Burger, J., & Kaar, N. (2012). Quantification of cow's milk percentage in dairy products - a myth? *Analytical and Bioanalytical Chemistry*, 403(10), 3031–3040. <https://doi.org/10.1007/s00216-012-5805-1>.
- Nardiello, D., Natale, A., Palermo, C., Quinto, M., & Centonze, D. (2018). Milk authenticity by ion-trap proteomics following multi-enzyme digestion. *Food Chemistry*, 244, 317–323. <https://doi.org/10.1016/j.foodchem.2017.10.052>.
- Russo, R., Severino, V., Mendez, A., Lliberia, J., Parente, A., & Chambery, A. (2012). Detection of buffalo mozzarella adulteration by an ultra-high performance liquid chromatography tandem mass spectrometry methodology. *Journal of Mass Spectrometry*, 47(11), 1407–1414. <https://doi.org/10.1002/jms.3064>.
- Watson, A. D., Gunning, Y., Rigby, N. M., Philo, M., & Kemsley, E. K. (2015). Meat authentication via multiple reaction monitoring mass spectrometry of myoglobin peptides. *Analytical Chemistry*, 87(20), 10315–10322. <https://doi.org/10.1021/acs.analchem.5b02318>.