Food Chemistry xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Food Chemistry

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

Principal component analysis of proteolytic profiles as markers of authenticity of PDO cheeses

Joana Santos Guerreiro [a,*], Mário Barros [a], Paulo Fernandes [a], Preciosa Pires [a], Ronald Bardsley [b]

[a] Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Viana do Castelo, Avenida do Atlântico, 4900-348 Viana do Castelo, Portugal
[b] School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD, United Kingdom

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
PCA
PDO cheeses
RP-HPLC
Urea-PAGE
Authenticity

ABSTRACT

The casein fraction of 13 Portuguese PDO cheeses were analysed using Urea-PAGE and reverse phase-high performance liquid chromatography (RP-HPLC) and then subjected to chemometric evaluation. The chemometric techniques of cluster analysis (CA) and principal component analysis (PCA) were used for the classification studies. Peptide mapping using Urea-PAGE followed by CA revealed two major clusters according to the similarity of the proteolytic profile of the cheeses. PCA results were in accordance with the grouping performed using CA.

CA of RP-HPLC results of the matured cheeses revealed the presence of one major cluster comprising samples manufactured with only ovine milk or milk admixtures. When the results of CA technique were compared with the two PCA approaches performed, it was found that the grouping of the samples was similar.

Both approaches, revealed the potential of proteolytic profiles (which is an essential aspect of cheese maturation) as markers of authenticity of PDO cheeses in terms of ripening time and milk admixtures not mentioned on the label.

1. Introduction

The broad objective in food authentication is to identify unique markers or groups of markers to characterise the authenticity of food or their potential adulterants/contaminants and use them to resolve authenticity issues. Proteins have been widely used as species markers in dairy products and for that purpose different technical approaches like electrophoresis, immunological and chromatographic techniques have been used (Mafra, Ferreira, & Oliveira, 2008).

Protected Denomination of Origin cheeses (PDO) are products of high commercial value confined according to legislative and proper labelling rules. Their popularity is growing and therefore it is crucial to protect the interests of the producers and to protect the consumers. Portuguese PDO cheeses have distinct characteristics. Their authenticity is associated with several factors such as a geographical area of production, materials and technology used. It is necessary to have in focus, that in the situation of the PDO products the authenticity issue has a wide range since each kind can be considered a distinct entity due to the microbiological, proteolytic and lipolytic profiles in the finished product.

Cheese ripening is the last step before commercialisation and it is an outcome of several biochemical and metabolic processes. Proteolysis is one of the most important biochemical events for the development of flavour and texture during ripening of most cheese varieties (McSweeney, 2004).

Any diagnostic procedure based on protein analysis must take into consideration that proteolytic activity is a key aspect of cheese manufacture. The rate, extent, and nature of proteolysis during cheese ripening, as well as the amount and nature of the degradation products, varies according to: (1) proteinases indigenous to milk (plasmin and cathepsin D for instance), (2) proteinases included in the coagulant formulation, (4) proteinases and peptidases from lactic acid bacteria, (3) the type of cheese and (4) the environmental conditions of ripening (McSweeney, 2004; Pereira, Gomes, Gomes, & Malcata, 2008).

Improved methods of studying cheese proteins, by various types of electrophoresis and chromatography, in conjunction with multivariate analysis, suggest that it may be possible to authenticate PDO cheeses considering not only the milk of origin, but also a geographical origin and method of production since all these factors influence the proteolytic patterns. The purpose of multivariate analysis is to develop a model that accurately characterises some property that is difficult to measure directly (Arvanitoyannis & Tzouros, 2005). Chemometrics can be visualised as the implementation of mathematical and statistical tools to the interpretation of

* Corresponding author. Tel.: +351 258 819700; fax: +351 258 827636.
E-mail address: joana@estg.ipvc.pt (J.S. Guerreiro).

0308-8146/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodchem.2012.02.066

patterns in multivariate data (Arvanitoyannis, Tsitsika, & Panagiotaki, 2005).

The implementation of novel and accurate quality methods in conjunction with chemometrics is a very powerful approach toward authentication of dairy products (Arvanitoyannis & Tzouros, 2005), Honey (Arvanitotoyannis, Chalhoub, Gotsiou, Lydakis-Simantiris, & Kefalas, 2005), classification of foods according to their variety or geographical (Tzouros & Arvanitoyannis, 2001) and fish authenticity (Arvanitoyannis et al., 2005).

The objective of this work was to evaluate the potential of proteolytic profiles when used in conjunction with chemometrics, as markers of authenticity of PDO cheeses in terms of ripening time and milk admixtures. The analytical results obtained due to the modest number of cheese samples can only give trends which further needs to be confirmed.

2. Materials and methods

2.1. Samples

Milk samples from sheep ($n = 3$), goats ($n = 3$) and cows ($n = 3$) were collected in sterile falcon tubes, filled to the top, and stored at −20 °C. A total of 13 kinds of PDO Portuguese cheeses, that have met the regulatory standards with the EU symbol and the label of certification, were obtained on the same occasion from the local market.

Cheese samples (approximately 100 g) were obtained from the interior part of cheese using sterilised instruments to avoid contamination. They were then cut into small pieces, distributed in falcon tubes and stored at −20 °C.

2.2. Caseins precipitation

Caseins from milk and cheese were extracted by precipitation at pH of 4.3 by the addition of 1 M ammonium-acetate buffer as described by Veloso, Teixeira, and Ferreira (2002).

2.3. Urea-PAGE

Polyacrylamide gel electrophoresis of casein samples from milk and cheeses was undertaken according to the method of Andrews (1983) with the modifications carried out by Veloso et al. (2002) regarding buffers composition and acrylamide concentration. The casein samples were dissolved in a diluted NaOH solution (0.1 M) at pH of 9.0. The samples were vigorously stirred and then placed in ultrasound bath (Transsonic 820/H Elma, Singen, Germany) until total dissolution. The casein concentration of casein samples was determined using the Lowry method (Lowry, Rosebrough, Farr, & Randall, 1951) with Bovine Serum Albumin (BSA) (Fluka, St. Louis, MO, USA) as standard. Casein samples were frozen at −20 °C until analysis.

Urea polyacrylamide gel electrophoresis of samples was undertaken in a vertical electrophoresis system (Hoefer Mighty Small II SE 260 Mini vertical, Amersham Biosciences, San Francisco, USA) using a E865 power supply (Consort, Turnhout, Belgium).

2.4. Caseins analysis by reversed-phase HPLC

The HPLC equipment consisted of a Hewlett Packard Series 1100 (Agilent Technologies, Waldbronn, Germany) with a quaternary pump and a UV detection type with double-beam photometer, the light source was a deuterium lamp, with a wavelength range between 190 and 600 nm. A reversed-phase Chrompack 300 RP column that contains a rigid macroporous polystyrene divinylbenzene polymer with a particle size of 8 μm and a pore size of 300 Å (150 × 4.6 mm I.D.) was used. The column was placed in a water bath (Memerrt W22, Schwabach, Germany) with a temperature of 45 ± 0.1 °C. The solvents that were used were of analytical purity grade. Solvent A consisted of 0.1% trifluoroacetic acid (TFA) in water and solvent B was acetonitrile: water: TFA (95:5:0.1, v/v).

The casein samples and the standards (Bovine milk casein and α-, β- and κ-caseins) were dissolved in a mixture of the mobile phase solvents A and B (70:30, v/v), they were vigorously stirred, placed on an ultrasound bath until dissolution and stored at −20 °C until use. Just before injection the solution was filtered through a 0.2 μm Acrodisic LC filter (Pall Life Sciences, Port Washington, NY, USA). For each analysis 20 μl (loop) of the solution was applied to the HPLC column.

Proteins were eluted with a series of linear gradients increasing the proportion of solvent B from 29% to 100% over 35 min: 1–5 min, 29% B; 5–10 min, 29–37% B; 10–12 min, 37–41% B; 12–14 min, 41–42.5% B; 14–16 min, 42.5% B; 16–17 min, 42.5–43% B; 17–19 min, 43% B; 19–21 min, 43–47% B; 21–23 min, 47% B; 23–25 min, 47–54% B; 25–27 min, 54% B; 27–28 min, 54–100% B; 28–30 min, 100–29% B; 30–35 min, 29% B. The flow-rate was 1 ml/min and the detection was made at a wavelength of 280 nm.

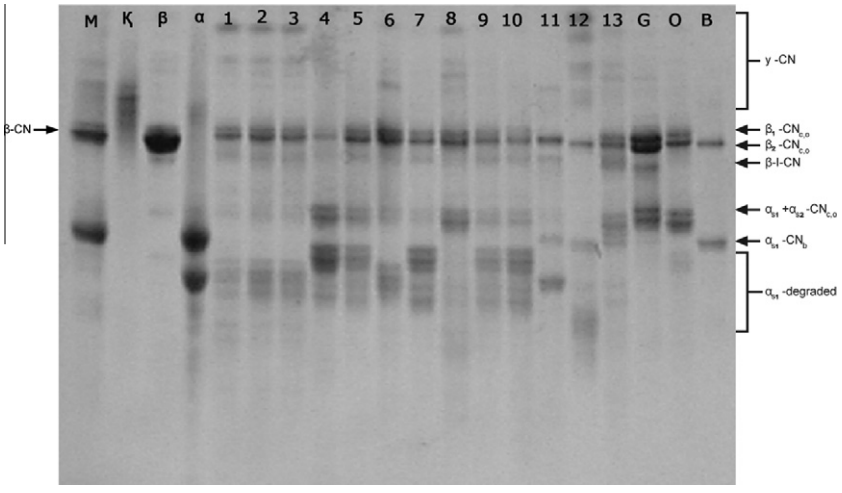

Fig. 1. Representative Urea-PAGE gel of PDO cheese casein samples, Lane M bovine casein standard; lane κ bovine κ-casein standard; lane β bovine β-casein standard; lane α bovine α-casein standard, lane 1–13 PDO cheeses, lane G raw caprine milk; lane O raw ovine milk and lane B raw bovine milk. Band identity based on standards or literature, CN, casein and subscripts indicate species; c = caprine, o = ovine and b = bovine.

Please cite this article in press as: Guerreiro, J. S., et al. Principal component analysis of proteolytic profiles as markers of authenticity of PDO cheeses. *Food Chemistry* (2012), doi:10.1016/j.foodchem.2012.02.066

Table 1
Casein fractions using Urea-PAGE electrophoresis in the 13 PDO cheeses (peak area %).

Fraction	1		2		3		4		5		6		7		8		9		10		11		12		13	
	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$
X1	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	14.53	0.63	n.d	–
X2	14.17	1.71	15.24	2.02	11.92	1.48	n.d	–	n.d	–	12.52	1.19	n.d	–	15.91	3.94	n.d	–	n.d	–	n.d	–	n.d	–	15.14	0.86
X3	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	12.22	1.35	n.d	–
X4	11.50	3.32	10.95	0.33	12.65	0.56	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–
X5	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	11.02	1.34	n.d	–
X6	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	16.00	2.80	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–
X7	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	10.04	1.29	n.d	–	n.d	–	n.d	–	n.d	–	14.12	0.44	n.d	–	n.d	–
X8	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	16.19	2.98	n.d	–
β-CN	20.51	1.81	19.99	0.35	24.52	1.89	21.78	1.03	22.07	0.42	22.52	3.25	26.37	1.63	36.79	1.80	25.54	0.47	17.70	1.25	20.58	0.05	9.64	3.29	20.41	4.99
X9	13.15	0.42	10.23	2.37	10.17	1.00	n.d	–	12.55	0.50	n.d	–	10.91	3.13	n.d	–	10.53	0.55	12.83	0.63	19.00	2.08	n.d	–	16.13	4.68
α_{s1}-Casein c,o	14.80	0.26	10.77	1.49	11.49	1.03	24.09	1.82	20.01	0.26	14.58	1.06	15.59	2.55	31.31	0.66	17.41	5.09	17.37	0.46	n.d	–	n.d	–	16.27	5.99
X10	n.d	–	n.d	–	n.d	–	15.86	1.14	16.33	2.89	n.d	–	11.93	3.35	n.d	–	13.48	2.42	13.68	0.58	20.40	1.15	13.63	0.65	15.53	1.01
X11	15.70	1.53	19.13	1.79	14.47	0.03	16.65	3.22	15.50	1.39	13.61	3.77	15.90	0.88	n.d	–	15.73	0.85	13.10	1.49	n.d	–	n.d	–	n.d	–
X12	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	25.90	2.74	n.d	–	5.09	7.20
X13	n.d	–	12.20	0.42	11.79	0.75	n.d	–	n.d	–	26.73	0.64	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–
X14	n.d	–	n.d	–	n.d	–	18.64	1.55	11.04	0.41	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–
X15	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	19.30	4.67	n.d	–	17.32	1.91	25.31	0.75	n.d	–	n.d	–	11.43	0.75
X16	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	5.82	0.26	n.d	–
X17	9.19	1.03	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	n.d	–	16.95	0.07	n.d	–

X = mean value; $s_{\bar{X}}$ = standard deviation ($n = 2$); sample 1 Terrincho cheese; sample 2 old Terrincho cheese; sample 3 Serra da Estrela Cheese; sample 4 Azeitão cheese; sample 5 Nisa cheese; sample 6 Évora cheese; sample 7 Serpa cheese; sample 8 Picante da Beira Baixa cheese; sample 9 Amarelo da Beira Baixa cheese; sample 10 Castelo Branco cheese; sample 11 Pico cheese; sample 12 S. Jorge cheese; sample 13 Cabra Transmontano. n.d., not detected.

2.5. Statistical analysis

Electrophoretic gel images from Urea-gels were analysed with 1D Image Analysis Software "KODAK" (Kodak, Rochester, NY, USA) and the peak areas, as percentage of total protein intensity, of the casein fractions of each sample were subjected to descriptive statistics, cluster analysis, utilising cluster procedures with single linkage amalgamation rule and Euclidean distances, performed with the software "Statistica 6.0" (StatSoft, Inc., Tulsa, OK, USA). Principal component analysis (PCA) was performed using the statistical software "ANDAD 7.12" (Sousa & Sousa, 2000) (CVRM, IST, Lisbon, Portugal). For HPLC analysis descriptive statistics and analysis of variance (ANOVA) were also performed with the "Statistica 6.0" (StatSoft, Inc., Tulsa, OK, USA).

3. Results and discussion

3.1. Urea-PAGE analysis of caseins

The study and evaluation of caseins and their proteolytic fragments by Urea-PAGE is a widely-used method and considered to be one of the oldest techniques for detection of milk mixtures of different species (Mayer, 2005).

The electrophoretic method allowed the visualisation of the caseins and high molecular weight peptides from 13 PDO cheeses, and analysis indicated that different patterns could be observed. The intensity of each band reflects the amount of each protein or peptide present in the extract. Thirteen PDO cheeses analysed are made with raw milk and it was possible to detect the extent of proteolysis in all of them.

For better identification of the casein bands and their breakdown products and the comparison of the cheese profiles, all 13 PDO cheeses were analysed simultaneously with the milk samples from the cows', goats' and ewes' milk. Fig. 1 shows a representative electropherogram of caseins from 13 PDO cheeses and milk resolved in Urea-PAGE at pH of 8.9 in the presence of 4.5 M Urea.

Using numerous Urea-PAGE analyses, the casein fractions and their breakdown products from 13 PDO cheeses and milk were identified by comparison with literature data (Farkye, Kiely, Allshouse, & Kindstedt, 1991; Freitas & Malcata, 1999; Furtado, 1983; Lane & Fox, 1996; Molina, Ramos, & Amigo, 2002; Tavaria et al., 1997; Van Hekken & Thompson, 1992) and with standards run alongside. While analysing different patterns from different species, it is possible to foresee that the presence of cowś milk should be easily detected in goat and sheep milk, based on αs1-casein mobility, as reported by Ferreira et al. (2006). Similar bands were tentatively identified visually and were matched in the Urea-PAGE gels. The proteolytic profiles of two gels were analysed by image analysis software and the peak area as percentage of total protein intensity was used for multivariate statistical analysis used as reported by Pripp, Stepaniak, and Sorhaug (2000). A total of 19 peaks were recognised between 13 PDO cheeses; some of them were recognised and matched as equivalent in some of the proteolytic profiles taking into account the mobility of the gel. Each peak represents a variable and each protein profile represents a sample in the multivariate data set. The data are listed in Table 1, with the peaks coded with "*X*" being unidentified and numbered sequentially according to increased migration. Good separation between β-casein and αs-casein was observed. Bands *X*1–*X*8 migrate more slowly than β-casein and may represent γ-casein resulting from the β-casein degradation attributable in part at least, to the activity of the endogenous milk proteinase, plasmin, with hydrolyses β-casein to γ-casein and protease–peptones (Eigel et al., 1984; Farkye et al., 1991; Ginger & Grigor, 1999; Molina et al., 2002). In the case of the band corresponding to the whole β-casein, there has been no significant decrease of intensity. However some pep-

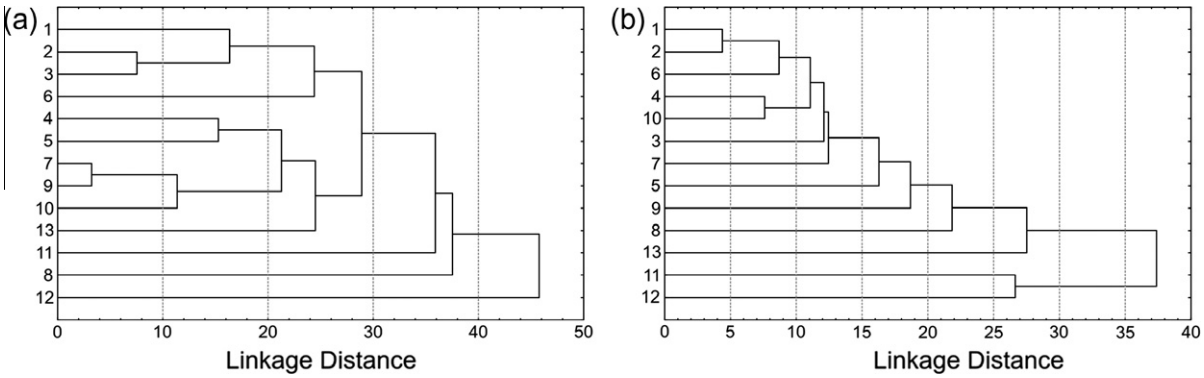

Fig. 2. (a) Dendrogram resulting from cluster analysis of the gel in Fig. 1. The numbers near the ends of the dendogram branches correspond to those of the gel lanes in Fig. 1. (b) Dendrogram resulting from cluster analysis of casein cheese samples analysed by RP-HPLC. The numbers at the end of branches correspond to cheeses identifiers.

tides derived from the hydrolysis of β-casein can be seen in some profiles. These observations are in tune with the literature review where β-casein has been claimed to be more resistant to proteolysis than αs1-casein in cheese manufactured with caprine, bovine and ovine milk and coagulated with animal and plant rennet (Freitas et al., 1999; Marcos, Esteban, Leon, & Fernandezsalguero, 1979). The pattern of γ-casein is quite complex, so that differentiation between bovine, ovine and caprine γ-casein is difficult (Fig. 1). Two types of cheese made with bovine milk (samples 11 and 12) have different patterns of degradation of β-casein due to different degrees of hydrolysis of β-casein. Cheese 12 (S. Jorge) presents a more extensive degradation of β-casein which corresponds to the appearance of γ-casein bands. Samples 4, 5, 7, 9 and 10 do not show fractions that result from β-casein hydrolysis.

In cheeses manufactured either with ovine and/or caprine milk, two bands could be observed in the electrophoretic region of β-casein. This result is consistent with the ones obtained by Freitas, Fresno, Prieto, Malcata, and Carballo (1997) using the same type of cheese as sample 8 studied in this work (Picante da Beira Baixa). From Fig. 1, it can be seen that in cheese samples 1, 2, 3, 5, 7, 9, 10, 11 and 13, two bands could be observed in the electrophoretic region of β-casein, one corresponding to a major peptide *X*9 (Table 1) that has been identified by some authors as β-I peptide (Carmona, Sanjuan, Gómez, & Fernández-Salguero, 1999; McSweeney, Fox, & Olson, 1995; Silva, Barros, & Malcata, 2002; Tavaria et al., 1997; Trujillo, Guamis, & Carretero, 1997). During cheese ripening αs1-casein is the first to be hydrolysed, although β-casein is also degraded. This pattern of αs1-casein hydrolysis was reported by several authors and makes the identification of milk admixtures in cheeses after a long ripening period quite difficult. From the casein profiles of the cheeses, it is possible to see that αs1-casein is hydrolysed forming the corresponding peptides, although a faint band in the region of αs1-casein from ovine and caprine milk appears in all samples with exception of samples 11 and 12, which were made with bovine milk (Fig. 1). The larger peptide α_{s1}I-casein, representing amino acids 24–199 (f24–199), from ovine and caprine α_{s1}-casein migrates in the same region as α_{s1}-casein bovine (*X*10) and it is further hydrolysed to smaller fragments. Low intensity bands *X*10–*X*17 with greater mobility than α_{s1}I-casein are present in all the cheese samples studied. The same was reported for cheddar cheese where degradation of α_{s1}-casein was extensive. It was initially hydrolysed to α_{s1}I-casein peptide (f24–199) and subsequently to α_{s1}-casein (f102–109) and other peptides with faster mobility than α_{s1}I-casein peptide (Lane & Fox, 1996).

Using the data obtained (Table 1), computer-assisted cluster analysis of proteolytic profiles was performed, matrix 13 × 19, and the best dendogram was obtained with the Euclidean distance using raw data. The purpose of this algorithm was to join the profiles together into successively larger clusters, using some measure of similarity. The distance between two clusters is determined by the distance of the two closest objects (nearest neighbours, in the different clusters) and Euclidean distances (the most straightforward way of computing distances between objects in a multi-dimensional space), it is simply the geometric distance in the multidimensional space) (StatSoft Inc., Tulsa, OK, USA). The hierarchical tree observed in Fig. 2a reveals proteolytic profiles were quite different. The cheeses were grouped according to the similarity of their proteolytic profile, and it was possible to identify the presence of two major clusters (samples 1, 2, 3 and 6 grouped together and samples 4, 5, 7, 9, 10 and 13 grouped together). The proteolytic profile of samples 8, 11 and 12 was different from others.

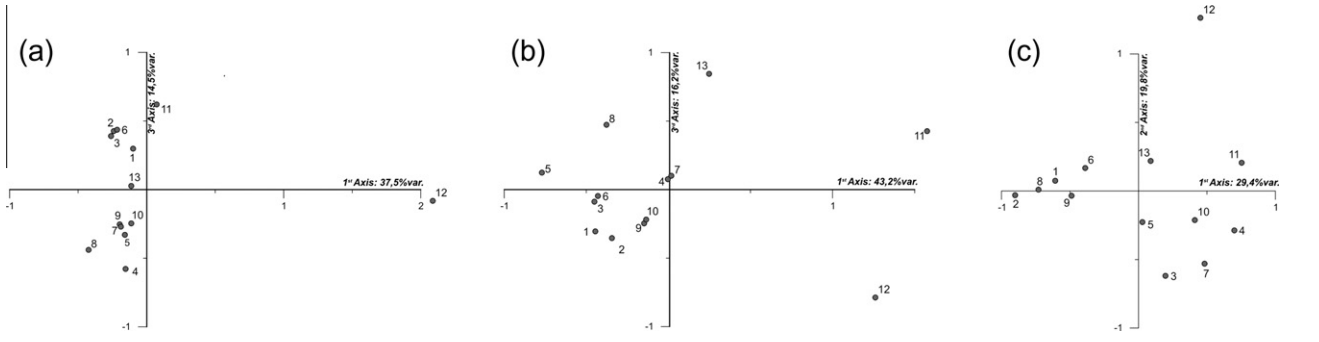

Fig. 3. (a) Secondary factorial plane of PCA for cheese casein analysed by Urea-PAGE. The numbers in the plot correspond to those of the gel lanes in Fig. 1. (b) Primary factorial plane of PCA for cheese caseins analysed by RP-HPLC (% peak area). (c) Primary factorial plane of PCA of PDO cheese chromatographic profiles analysed by RP-HPLC.

Please cite this article in press as: Guerreiro, J. S., et al. Principal component analysis of proteolytic profiles as markers of authenticity of PDO cheeses. *Food Chemistry* (2012), doi:10.1016/j.foodchem.2012.02.066

Table 2
Peptide pattern of the casein samples of the 13 PDO cheeses analysed by RP-HPLC (peak area %).

Retention time	Anova	1		2		3		4		5		6		7		8		9		10		11		12		13	
		$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$	$\bar{X}$	$s_{\bar{X}}$
18.00	***	1.12	0.32	1.69	0.22	9.71	0.15	10.09	0.67	4.42	0.45	2.06	0.10	13.47	0.22	6.97	1.34	5.74	0.67	5.59	0.20	2.92	0.11	1.82	1.83	21.14	2.95
19.80	***	23.70	2.58	25.85	0.41	20.97	0.22	18.24	0.66	23.78	1.08	28.67	0.97	20.72	0.87	22.86	0.46	26.58	1.13	22.20	1.11	7.47	0.93	8.11	3.04	10.87	1.91
21.60	***	n.d.	–	n.d.	–	n.d.	–	n.d.	–	n.d.	–	n.d.	–	n.d.	–	n.d.	–	n.d.	–	n.d.	–	12.25	0.60	n.d.	0.00	14.93	1.62
22.00	***	2.42	0.11	3.15	0.16	1.51	0.19	n.d.	0.00	n.d.	–	2.70	0.06	n.d.	0.00	n.d.	–	2.59	0.19	n.d.	–	14.67	1.01	11.40	2.67	n.d.	–
22.50	***	4.21	0.17	4.46	0.05	n.d.	–	7.56	0.73	n.d.	–	4.34	0.06	4.67	0.31	n.d.	–	4.28	0.28	5.92	0.13	21.35	1.83	27.03	2.84	1.71	1.34
23.00	***	7.92	0.18	9.11	0.06	7.21	1.11	15.02	0.70	n.d.	–	10.87	0.11	19.96	0.26	n.d.	–	9.70	0.20	15.99	0.52	13.09	2.49	21.16	5.23	13.25	2.29
23.90	***	36.69	1.67	33.57	0.38	40.09	0.53	37.63	2.61	52.14	1.09	29.42	0.47	27.53	0.46	30.02	1.13	17.13	0.22	35.24	2.05	4.19	0.35	n.d.	–	11.96	2.11
24.10	***	n.d.	–	n.d.	–	n.d.	–	n.d.	0.00	n.d.	–	n.d.	–	n.d.	–	16.71	1.73	10.45	0.77	n.d.	–	24.07	2.74	8.69	2.37	4.13	0.53
25.00	***	4.60	0.14	4.17	–	7.01	0.18	5.84	0.69	4.27	0.95	1.13	0.06	6.80	0.44	n.d.	–	5.67	0.37	4.14	0.88	n.d.	–	3.80	0.79	4.51	0.75
26.00	***	13.97	0.33	12.64	0.36	8.65	0.20	5.63	0.90	12.76	1.15	18.10	0.79	6.85	0.84	21.60	0.65	11.92	0.90	7.12	0.59	n.d.	–	12.15	1.64	14.23	7.00
26.60	***	2.26	0.46	2.78	–	2.35	0.03	n.d.	–	n.d.	–	n.d.	–	n.d.	–	n.d.	–	3.24	0.42	2.18	1.08	n.d.	–	5.83	2.35	1.39	0.68
27.20	***	3.11	0.61	2.57	0.10	2.50	0.20	n.d.	–	2.62	0.40	2.71	0.13	n.d.	–	1.84	0.14	2.71	1.44	1.62	0.82	n.d.	–	n.d.	–	1.89	0.99

$\bar{X}$ = mean value; $s_{\bar{X}}$ = standard deviation ($n = 3$); Anova: ***$p \leq 0.001$; sample 1 Terrincho cheese; sample 2 Old Terrincho cheese; sample 3 Serra da Estrela Cheese; sample 4 Azeitão cheese; sample 5 Nisa cheese; sample 6 Évora cheese; sample 7 Serpa cheese; sample 8 Picante da Beira Baixa cheese; sample 9 Amarelo da Beira Baixa cheese; sample 10 Castelo Branco cheese; sample 11 Pico cheese; sample 12 S. Jorge cheese; sample 13 Cabra Transmontano cheese. n.d., not detected.

To the raw data of same matrix (13 × 19) principal component analysis (PCA) was applied (Fig. 3a). PCA is a way of identifying patterns in data, and expressing data in such a way as to emphasize theirs similarities and differences (Shin, Craft, Pegg, Phillips, & Eitenmiller, 2010). The factors are "extracted" from the data by statistical transformations involving the diagonalisation of the variables' correlation or variance–covariance matrix (Pereira & Sousa, 2000). Each factor describes a certain amount of the statistical variance of the analysed data and is interpreted according to the intercorrelated variables. After a principal component analysis the information is mainly concentrated on the first components. As a consequence, a plot of the scores of the objects on the first components allows the direct visualisation of the global information in a very efficient way; it is easy to detect similarity between objects (similar objects have a very similar position in the space) or the presence of outliers (they are very far from all other objects) or the existence of clusters (Leardi, 2003).

For the first factor, the loadings of the variables and the samples were the same, but for the third factor there were two samples with higher loadings: sample 11 with a high positive loading and sample 4 with a high negative loading. All the other samples have no significant loadings in either of the factors. Despite that only about 70% of the variance was effectively described by the first three factors, the analysis of Fig. 3a allowed to recognise a clear pattern of distribution since all samples, except bovine milk cheeses (samples 11 and 12), were grouped due to negative loadings in the first factor.

3.2. HPLC analysis of caseins

Reversed-phase high performance liquid chromatography (RP-HPLC) is a widely-used method for protein and peptide separation and it has been applied to the study of caseins in different types of milk (Bobe, Beitz, Freeman, & Lindberg, 1998; Ferreira et al., 2006; Mora-Gutierrez, Kumosinski, & Farrell, 1991; Morgan et al., 2003; Trujillo, Casals, & Guamis, 2000a, 2000b; Veloso, Teixeira, Peres, Mendonca, & Ferreira, 2004).

From 13 PDO cheeses and three milk samples of each species, three independent casein extractions were performed. All the extracts prepared were analysed three times (198 injections) to ensure repeatability and reproducibility of the equipment. In the present study, the HPLC conditions were optimised for mobile phase composition, gradient, operating procedure and flow-rate based on the procedure developed by Veloso et al. (2002) and the order of elution of bovine standards of κ-, α- and β-caseins resembled those obtained by these authors.

Each RP-HPLC profile of 13 PDO cheeses was compared with the casein profile of the different milk sources but it was not possible to extrapolate the milk cheese source from the milk chromatographic data. A total of 12 peaks were recognised and matched in the chromatographic profiles of the PDO cheeses taking into account the retention time of the fractions. Each peak represents a variable and each chromatographic profile represents a sample in the multivariate data set. The data are listed in Table 2 with the fractions where similar retention times have been identified. All peaks showed highly significant differences between 13 PDO cheeses after ANOVA analysis. The relative proportions of κ-casein, α-casein, β-caseins are subject to variations between milk from different species and within the breeds of the same species. In cheese however, the differences between samples are more likely to occur due to different degrees of proteolysis than the milk source (Pillonel et al., 2003).

Using the data from Table 2, cluster analysis was performed and the best dendrogram was obtained with the Euclidean distance using raw data (matrix 13 × 12). The hierarchical tree observed in Fig. 2b reveals the presence of one cluster with corresponding

similarity between the protein or peptide fractions, considering a linkage distance to be less than 25 units comprising samples 1–10, manufactured with only ovine milk or milk admixtures. Sample 13 (Cabra Transmontano cheese) clusters with this first cluster considering a linkage distance higher than 25 units. The same applies to samples 11 and 12, the only two cheese samples manufactured only with bovine, which also formed a different cluster.

Principal component analysis was applied in two different ways to evaluate if direct relationships can be found between samples based on milk origin (mentioned on the label or detected by DNA analysis (data not shown)) and proteolytic profiles. Firstly, PCA was performed based on 12 peaks selected from 13 PDO cheeses. The second approach was to apply PCA to all the data from the chromatograms of 13 PDO cheeses, without recognition and matching of equivalent peaks from proteolytic profiles.

On first approach, Fig. 3b, shows the factorial plane performed with the data of Table 2, on a matrix 13×12. The first three factors explained almost 80% of the sample variance. The first factor which explained 43.2% of the variance was highly correlated with samples 5, 11, 12 and moderate with samples 8, 6, 3 and 1. The third factorial plane which explained 16.2% of the data variance was positively highly correlated with sample 13 and negatively correlated with sample 12. From this score plot it was possible to identify that the first factor separates bovine cheese samples (samples 11 and 12) and caprine cheese (sample 13) from the other samples. All the samples that had ovine milk (1–10) were grouped in the left side of the plot.

The second approach relied on the application of principal component analysis to the chromatogram profile data without peak selection (Fig. 3c). On this circumstances only about 65% of the variance was effectively described by the first three factors.

The first two factors explained 49.2% of the data variance, a little lower compared to the first approach. Interestingly, the samples were organised in a very similar way regarding to bovine cheese samples (samples 11 and 12) and caprine cheese (sample 13). These three samples continued to behave like outliers and were grouped on the same area of the plot. The first factor was highly positively correlated with samples 4 and 11 and negatively correlated with samples 1, 2, 8 and 9. The second factor which explained 19.8% of the data variance was highly negatively correlated with samples 3 and 7 and positively correlated with sample 12. Samples 3, 4, 5, 7 and 10 were PDO cheeses manufactured with only ovine milk (confirmed by DNA analysis) and they were all grouped in the same area of the plot.

4. Conclusion

The evaluation of cheese authenticity by Urea-PAGE is a quite complex process involving many factors that are independent of the milk source, making it difficult to recognise a unique pattern to the proteolytic profile of a milk source. The results of this study were presented as observations and supportive data, but they can be statistically analysed.

The work presented here with Urea-PAGE did not allow the detection of admixtures but raised the possibility of using proteolytic profiles as a marker of authenticity, since the observance of cheese ripening conditions is one of the most important steps in PDO cheeses manufacture.

High performance liquid chromatography has been used to (i) understand the principles of proteolysis of milk caseins, (ii) to analyse cheese proteolysis and (iii) evaluate cheese authenticity. The evaluation of casein proteolysis by HPLC is normally followed by peptide identification using several approaches such as the sequence of N terminus (Singh, Fox, & Healy, 1997; Sousa & Malcata, 1998) or fast atom bombardment-mass spectrometry (FAB-MS) (Ferranti et al., 1997). Even though there is a great heterogeneity of milk caseins from the different species and proteolysis, this method combined with multivariate statistical tools allowed us to separate PDO cheeses made with bovine milk (Pico and S. Jorge) and PDO cheese made with caprine milk (Cabra Transmontano) from the other cheeses manufactured with ovine milk or admixtures. Interestingly, the caprine cheese sample (sample 13) using RP-HPLC approach and after PCA analysis was also placed in an outlier position relatively to samples 1–10, that are cheeses manufactured with goat and/or ewe milk.

First factor 1 through 3 together were found to be explanatory of almost 80% of the total variability in the data set. This results relies on recognition and matching of equivalent peaks from proteolytic profiles which brings us to the process of obtaining data as a subjective element, which is time consuming. The second approach, using all the data generated by the chromatogram, the first three factors only explain 65% the sample variance but allowed us to get results from chemometrical analyses immediately after the RP-HPLC. The basic pattern of samples organizion was very similar on both approaches.

Furthermore, this study clearly indicates that the combination of experimental Urea-Page and RP-HPLC data along with a chemometric approach (cluster analysis and PCA, in this case) revealed good potential concerning the assignment of proteolytic profiles as a marker of authenticity of technological process applied to PDO cheeses.

The proteolytic profiles could eventually be used as a marker of authenticity, a "fingerprint" of the process involved in cheese production, but large scale studies should be carried out to ensure good reliability of this approach.

References

Andrews, A. T. (1983). Proteinases in normal bovine milk and their action on caseins. *The Journal of Dairy Research, 50*(1), 45–55.

Arvanitotoyannis, I., Chalhoub, C., Gotsiou, P., Lydakis-Simantiris, N., & Kefalas, P. (2005). Novel quality control methods in conjunction with chemometrics (multivariate analysis) for detecting honey authenticity. *Critical Reviews in Food Science and Nutrition, 45*, 193–203.

Arvanitoyannis, I., Tsitsika, E. V., & Panagiotaki, P. (2005). Implementation of quality control methods (physicochemical, microbiological and sensory) in conjunction with multivariate analysis towards fish authenticity. *International Journal of Food Science and Technology, 40*, 237–263.

Arvanitoyannis, I., & Tzouros, N. E. (2005). Implementation of quality control methods in conjunction with chemometrics toward authentication of dairy products. *Critical Reviews in Food Science and Nutrition, 45*, 231–249.

Bobe, G., Beitz, D. C., Freeman, A. E., & Lindberg, G. L. (1998). Separation and quantification of bovine milk proteins by reversed-phase high-performance liquid chromatography. *Journal of Agricultural and Food Chemistry, 46*(2), 458–463.

Carmona, M. A., Sanjuan, E., Gómez, R., & Fernández-Salguero, J. (1999). Effect of starter cultures on the physico-chemical and biochemical features in ewe cheese made with extract from flowers of *Cyanara cardunculus* L.. *Journal of the Science of Food and Agriculture, 79*, 737–744.

Eigel, W. N., Butler, J. E., Ernstrom, C. A., Farrell, H. M., Harwalkar, V. R., Jenness, R., et al. (1984). Nomenclature of proteins of cows milk – 5th revision. *Journal of Dairy Science, 67*(8), 1599–1631.

Farkye, N. Y., Kiely, L. J., Allshouse, R. D., & Kindstedt, P. S. (1991). Proteolysis in Mozzarella cheese during refrigerated storage. *Journal of Dairy Science, 74*(5), 1433–1438.

Ferranti, P., Barone, F., Chianese, L., Addeo, F., Scaloni, A., Pellegrino, L., et al. (1997). Phosphopeptides from Grana Padano cheese: nature, origin and changes during ripening. *Journal of Dairy Research, 64*(4), 601–615.

Ferreira, I., Veiros, C., Pinho, O., Veloso, A. C. A., Peres, A. M., & Mendonca, A. (2006). Casein breakdown in Terrincho ovine cheese: Comparison with bovine cheese and with bovine/ovine cheeses. *Journal of Dairy Science, 89*(7), 2397–2407.

Freitas, A. C., Fresno, J. M., Prieto, B., Franco, I., Malcata, F. X., & Carballo, J. (1999). How milk type, coagulant, salting procedure and ripening time affect the profile of free amino acids in Picante da Beira Baixa cheese. *Journal of the Science of Food and Agriculture, 79*(4), 611–618.

Freitas, A. C., Fresno, J. M., Prieto, B., Malcata, F. X., & Carballo, J. (1997). Effects of ripening time and combination of ovine and caprine milks on proteolysis of Picante cheese. *Food Chemistry, 60*(2), 219–229.

Freitas, A. C., & Malcata, F. X. (1999). Technological optimisation of Picante cheese using microbiological, chemical and physical criteria. *Journal of Food Engineering, 41*(3–4), 163–175.

Furtado, M. M. (1983). Detection of cow milk in goat milk by polyacrylamide-gel electrophoresis. *Journal of Dairy Science, 66*(9), 1822–1824.

Please cite this article in press as: Guerreiro, J. S., et al. Principal component analysis of proteolytic profiles as markers of authenticity of PDO cheeses. *Food Chemistry* (2012), doi:10.1016/j.foodchem.2012.02.066

Ginger, M. R., & Grigor, M. R. (1999). Comparative aspects of milk caseins. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 124*(2), 133–145.

Lane, C. N., & Fox, P. F. (1996). Contribution of starter and adjunct Lactobacilli to proteolysis in cheddar cheese during ripening. *International Dairy Journal, 6*(7), 715–728.

Leardi, R. (2003). Chemometrics in data analysis. In M. Lees (Ed.), *Food authenticity and traceability* (pp. 299–320). Cambridge: Woodhead Publishing Limited.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951). Protein measurement with the folin phenol reagent. *Journal of Biological Chemistry, 193*(1), 265–275.

Mafra, I., Ferreira, I. M., & Oliveira, B. P. (2008). Food authentication by PCR-based methods. *European Food Research and Technology, 227*(3), 649–665.

Marcos, A., Esteban, M. A., Leon, F., & Fernandezsalguero, J. (1979). Electrophoretic patterns of European cheeses – comparison and quantitation. *Journal of Dairy Science, 62*(6), 892–900.

Mayer, H. K. (2005). Milk species identification in cheese varieties using electrophoretic, chromatographic and PCR techniques. *International Dairy Journal, 15*, 595–604.

McSweeney, P. L. H. (2004). Biochemistry of cheese ripening. *International Journal Dairy Technology, 57*(2–3), 127–144.

McSweeney, P. L. H., Fox, P. F., & Olson, N. F. (1995). Proteolysis of bovine caseins by cathepsin-D – Preliminary-observations and comparison with chymosin. *International Dairy Journal, 5*(4), 321–336.

Molina, E., Ramos, M., & Amigo, L. (2002). Characterisation of the casein fraction of Iberico cheese by electrophoretic techniques. *Journal of the Science of Food and Agriculture, 82*(10), 1240–1245.

Mora-Gutierrez, A., Kumosinski, T. F., & Farrell, H. M. (1991). Quantification of alpha s1-casein in goat milk from French-Alpine and Anglo-Nubian breeds using reversed-phase high performance liquid chromatography. *Journal Dairy Science, 74*(10), 3303–3307.

Morgan, F., Massouras, T., Barbosa, M., Roseiro, L., Ravasco, F., Kandarakis, I., et al. (2003). Characteristics of goat milk collected from small and medium enterprises in Greece, Portugal and France. *Small Ruminant Research, 47*(1), 39–49.

Pereira, H. J., & Sousa, A. J. (2000). Data analysis for the treatment of multidimensional tables: handbook of the intensive course of the data analysis 1988–2000. In CVRM (Ed.). Lisbon: Instituto Superior Técnico.

Pereira, C. I., Gomes, E. O., Gomes, A. M. P., & Malcata, F. X. (2008). Proteolysis in model Portuguese cheeses: Effects of rennet and starter culture. *Food Chemistry, 108*(3), 862–868.

Pillonel, L., Albrecht, B., Badertscher, R., Butikofer, U., Chamba, J. F., Tabacchi, R., et al. (2003). Analytical methods for the determination of the geographic origin of emmental cheese. Parameters of proteolysis and rheology. *Italian Journal of Food Science, 15*(1), 49–62.

Pripp, A. H., Stepaniak, L., & Sorhaug, T. (2000). Chemometrical analysis of proteolytic profiles during cheese ripening. *International Dairy Journal, 10*(4), 249–253.

Silva, S. V., Barros, R. M., & Malcata, F. X. (2002). Hydrolysis of caseins by extracts of *Cynara cardunculus* precipitated by ammonium sulfate. *Journal of Food Science, 67*(5), 1746–1751.

Singh, T. K., Fox, P. F., & Healy, A. (1997). Isolation and identification of further peptides in the diafiltration retentate of the water-soluble fraction of Cheddar cheese. *Journal of Dairy Research, 64*(3), 433–443.

Shin, E.-C., Craft, B. D., Pegg, R. B., Phillips, R. D., & Eitenmiller, R. R. (2010). Chemometric approach to fatty acid profiles in runner-type peanut cultivars by principal component analysis (PCA). *Food Chemistry, 119*(3), 1262–1270.

Sousa, M. J., & Malcata, F. X. (1998). Identification of peptides from ovine milk cheese manufactured with animal rennet or extracts of *Cynara cardunculus* as coagulant. *Journal of Agricultural and Food Chemistry, 46*(10), 4034–4041.

Sousa, P., & Sousa, J. (2000). Geoestatística Multivariada. Programa ANDAD. V. 7.10. In CVRM/IST (Ed.).

Tavaria, F. K., Sousa, M. J., Domingos, A., Malcata, F. X., Brodelius, P., Clemente, A., et al. (1997). Degradation of caseins from milk of different species by extracts of *Centaurea calcitrapa*. *Journal of Agricultural and Food Chemistry, 45*(10), 3760–3765.

Trujillo, A.-J., Casals, I., & Guamis, B. (2000a). Analysis of major ovine milk proteins by reversed-phase high-performance liquid chromatography and flow injection analysis with electrospray ionization mass spectrometry. *Journal of Chromatography A, 870*(1–2), 371–380.

Trujillo, A. J., Casals, I., & Guamis, B. (2000b). Analysis of major caprine milk proteins by reverse-phase high-performance liquid chromatography and electrospray ionization-mass spectrometry. *Journal of Dairy Science, 83*(1), 11–19.

Trujillo, A. J., Guamis, B., & Carretero, C. (1997). Proteolysis of goat casein by calf rennet. *International Dairy Journal, 7*(8–9), 579–588.

Tzouros, N. E., & Arvanitoyannis, I. S. (2001). Agricultural produces: Synopsis of employed quality control methods for the authentication of foods and for the classification of foods according to their variety of geographical origin. *Critical Reviews in Food Science and Nutrition, 41*(4), 287–319.

Van Hekken, D. L., & Thompson, M. P. (1992). Application of PhastSystem to the resolution of bovine milk proteins on urea-polyacrylamide gel electrophoresis. *Journal Dairy Science, 75*(5), 1204–1210.

Veloso, A. C., Teixeira, N., & Ferreira, I. M. (2002). Separation and quantification of the major casein fractions by reverse-phase high-performance liquid chromatography and urea-polyacrylamide gel electrophoresis. Detection of milk adulterations. *Journal Chromatography A, 967*(2), 209–218.

Veloso, A. C. A., Teixeira, N., Peres, A. M., Mendonca, A., & Ferreira, I. (2004). Evaluation of cheese authenticity and proteolysis by HPLC and urea-polyacrylamide gel electrophoresis. *Food Chemistry, 87*(2), 289–295.

Please cite this article in press as: Guerreiro, J. S., et al. Principal component analysis of proteolytic profiles as markers of authenticity of PDO cheeses. *Food Chemistry* (2012), doi:10.1016/j.foodchem.2012.02.066