

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Casein composition and differential translational efficiency of casein transcripts in donkey's milk

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1721772> since 2020-01-08T09:09:15Z

Published version:

DOI:10.1017/S0022029919000256

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

1 **Casein composition and differential translational efficiency of casein transcripts in**
2 **donkey's milk**

3

4 Gianfranco Cosenza^{1†*}, Rosalba Mauriello^{1†}, Giuseppina Garro¹, Barbara Auzino², Marco
5 Iannaccone¹, Angela Costanzo¹, Lina Chianese¹ and Alfredo Pauciullo³

6

7 ¹ Department of Agricultural Sciences, University of Naples Federico II, Portici (Naples), Italy

8 ² Department of Veterinary Sciences, University of Pisa, Pisa, Italy

9 ³ Department of Agricultural, Forest and Food Science, University of Torino, Grugliasco (TO), Italy

10

11 †These authors contributed equally to this work

12 *For correspondence; e-mail: giacosen@unina.it

13

14

15

16

17

18

19

20

21

22

23

24

25 **Abstract**

26 The amount of the four caseins (α_{s1} , α_{s2} , β and κ -CN) in donkey milk was evaluated by Urea-
27 PAGE analysis at pH 8.6, followed by immuno-detection with polyclonal antibodies, coupled to
28 densitometric analysis. The results showed the percentage of each casein in decreasing order: β (54.28) >
29 α_{s1} (35.59) > α_{s2} (7.19) > κ -CN (2.79). The mRNA quantification of donkey casein transcripts, carried out
30 by RT-qPCR, showed that the average percentage of corresponding gene transcripts (*CSN2*, *CSNIS1*,
31 *CSNIS2* I and *CSN3*) was 70.85, 6.28, 14.23 and 8.65, respectively. The observed translation efficiency,
32 assessed as percentage of single milk casein fraction out of single percentage of transcript, was 0.76, 5.66,
33 0.50 and 0.32, respectively. The analysis of the sequences flanking the start codon, the codon usage
34 frequencies and the coding sequence length might explain, at least in part, the differential transcriptional
35 and translational rate observed among the casein transcripts.

36

37 **Keywords:** Donkey, casein, mRNA, quantification

38

39 In recent years donkey's milk (DM) has attracted an increasing interest in human nutrition, since it may
40 represent the best natural substitute of cow's milk for children affected by milk protein allergy, a
41 condition of increasing incidence (Businco et al. 2000; Monti et al. 2012; Cunsolo et al. 2017). Allergic
42 manifestations to DM are rare and, to date, only one case of work-related DM allergy has been
43 documented (Giorgis et al. 2018). DM may be considered a valid alternative to powdered milks, soybean
44 milk replacement or other formulas employed in the diet therapy of these patients. The reason lies in the
45 low casein content and in the ratio casein to whey protein that is closer to human milk than what observed
46 in ruminant milk (Guo et al. 2007). Recently, the presence of all four casein fractions α_{s1} , β , α_{s2} and κ -
47 CN was demonstrated in donkey's milk (Chianese et al. 2010), as well as in the horse (Ochirkhuyag et al.
48 2000) and pony (Miranda et al 2004). The proteomic approach has also allowed characterization of the
49 casein compositional heterogeneity due to post-translational modifications, like phosphorylation (α_{s1} , α_{s2}

50 and β -CN), glycosylation (κ -CN) and non-allelic forms generated by RNA incorrect splicing (α_{s1} and β -
51 CN) (Cunsolo et al. 2009a; Cunsolo et al. 2009b; Chianese et al. 2010). In particular, the complete
52 primary structure of α_{s1} -casein (202 amino acids, Cunsolo et al. 2009a), β (226 amino acids, Cunsolo et
53 al. 2009b) and α_{s2} (221 amino acids, Chianese et al. 2010) have been determined. Moreover, the complete
54 sequences of the genes encoding for the β - (*CSN2*, EMBL No. FN598778), α_{s1} - (*CSNIS1*, EMBL No.
55 FN386610) and κ -casein (*CSN3*, Hobor et al. 2008; FR822990) and the related promoter regions have
56 been determined.

57 Similarly, two different donkey α_{s2} encoding genes (*CSNIS2 I* and *CSNIS2 II*) have been
58 identified (Cosenza et al. 2010). The first, spanning over a fragment of 1016 nt, is constituted by 19 exons
59 and it encodes for the protein of 221 amino acids (called α_{s2} -I) also characterized by Chianese et al.
60 (2010); the second, constituted by 16 exons, probably originated by gene duplication, encodes for a
61 predicted peptide (named α_{s2} -II) of 168 amino acids (Cosenza et al. 2010), not yet detected at proteomic
62 level. Studies on the genetic polymorphism of DM are limited when compared to those carried out in the
63 major dairy species, and it is only recently that researchers have paid particular attention to the proteomic
64 and genomic characterization of proteins in DM. In particular, Criscione et al. (2009) have identified an
65 individual DM sample lacking α_{s1} -casein, like in goats, known as the species expressing the highest
66 genetic variability for this casein fraction (Cosenza et al. 2008). In addition, Chianese et al. (2010) have
67 characterized a genetic variant of β -casein having a molecular weight value 28 mass units higher than the
68 common β -CN phenotype. Finally, regarding the *CSN3* and *CSNIS2 I* genes, the analysis of nucleotide
69 sequences has allowed the identification of several silent and missense polymorphisms (Hobor et al.
70 2008; Cosenza et al. 2010). On the contrary, no studies have been carried out on the expression of casein
71 genes in the donkeys, as well as on their translational efficiency, whereas cattle, sheep, goat (Bevilacqua
72 et al. 2006), buffalo (Cosenza et al. 2011) and yak (Bai et al. 2013) data have been reported.

73 The hypothesis of our study was that in donkey, similarly to what is observed in ruminants, a
74 significant difference in the translation efficiency characterises the genes encoding the four caseins. In

75 order to verify such hypothesis, we evaluated the expression of the four casein fractions in DM taking
76 into account the phenotypic and genotypic aspects. The protein quantification of α_{s1} , α_{s2} , β and κ -CN was
77 carried out by means of electrophoresis at alkaline pH and immunoblotting with polyclonal antibodies
78 coupled to densitometry analysis. The quantitative determination of the four casein mRNAs was assessed
79 by RT-qPCR and their translation efficiency was estimated through the percentage ratio of single milk
80 casein fractions/single percentage of transcripts.

81

82 **Materials and methods**

83 *Donkey milk sampling and casein extraction*

84 Individual milk samples from 8 donkeys of Martina Franca breed were collected in the same farm
85 (Aquila, Italy). Martina Franca are large-sized donkeys that originated in the Apulia region in the South-
86 East of Italy. In the past, the Martina Franca donkey breed has been considered useful for the production
87 of hybrids. Currently in Italy, the breed is used mainly in an amateur context, although different potential
88 uses (recreational, pet therapy, meat and milk production) are developing. The maximum milk yield per
89 milking corresponds to 700 grams (approximately 1.4 L) and regarding milk composition (g/100 g), the
90 maximum values are 0.97 for fat, 1.67 for protein, 6.87 for lactose and 9.05 for SCC (x 1000 cells/mL)
91 (D'Alessandro et al. 2009). All donkeys were free of clinical mastitis and were comparable for age (about
92 6 years old), lactation and parity order. Each casein sample was prepared by acid precipitation from
93 skimmed milk, as described by Aschaffenburg & Drewry (1959).

94

95 *Quantitative determination of the nitrogen fractions (TN, SN, CN, NPN) in donkey milk*

96 The total nitrogen in DM was determined by Kjeldahl method according to the IDF Method (1993). A
97 nitrogen protein conversion factor of 6.38 was used in all cases. All samples were analyzed in triplicate
98 and results presented as means \pm standard deviations.

99 *Urea polyacrylamide gel electrophoresis (Urea-PAGE) at pH 8.6 and immunoblotting analysis*

100 Urea- PAGE at pH 8.6 and the immunoblotting analysis were carried out according to the procedure
101 described by Chianese et al. (2009), using polyclonal antibodies against bovine peptides α_{s1} -CN (187-
102 199) and β -CN (195-199) and porcine κ and α_{s2} -CN. Each casein fraction were analyzed from the
103 Coomassie blue stained gel pattern by scanning with an Ultrosan XL enhanced laser densitometer
104 equipped with the software supplied by the manufacturer (Amersham Biosciences AB, Uppsala, Sweden).
105 Chemicals, the distribution of nitrogenous components, sample preparation and conditions of the
106 immunoelectrophoresis analysis were reported in supplementary materials.

107

108 *RNA analysis*

109 Total RNA was isolated from somatic cells present in the eight representative fresh milk samples using
110 Nucleospin Blood and NucleoSpin® Extract Kits (Macherey-Nagel). The quantity, quality, purity and
111 integrity of RNA, after DNase treatment, were estimated by means of Thermo Scientific NanoDrop 2000c
112 and by electrophoresis on a denaturing agarose gel. Reverse-Transcription reaction mix, quantitative PCR
113 amplification mix, thermal condition and primers sequences are reported in online Supplementary
114 Methods and supplementary table S1.

115

116 **Results and discussion**

117 *Quantitative analysis of the nitrogen fractions (TN, SN, CN, NPN) in donkeys' milk*

118 In the individual donkey milks analysed, the average protein content was $1.48\% \pm 0.2$, ranging between
119 1.10% and 1.81% (Supplementary Table S2) consistent with data reported by Salimei et al. (2004) and
120 Guo et al. (2007). In particular, the average content of caseins (34.61%) and whey proteins (49.80%),
121 with a casein to whey proteins ratio of 0.69, showed remarkable differences in comparison with bovine
122 and other ruminant milks but were within the range of donkey's milk variability, reported in literature
123 (Salimei et al. 2004; Guo et al. 2007). The one exception was CN content being lower than that reported
124 by Guo et al. (2007) for Chinese donkey milk. The high NPN content (15.55%) was very close to that of

125 human and mare's milk (Malacarne et al. 2002). The nutritional and biological significance of this milk
126 fraction is still far from being completely understood, but it seems to be related to the development of the
127 infant (Lonnerdal, 1994). It has been suggested that the high amount of whey protein (49.81%) in
128 donkey's milk, similar to mare's milk, may make it more favourable for human nutrition than cow's milk,
129 because of the relatively higher acute postprandial availability of essential amino acids.

130

131 *Qualitative and quantitative characterization of donkey's caseins by Urea-PAGE at pH 8.6,*
132 *immunoblotting and densitometry analysis*

133 The individual casein samples analysed by Urea-PAGE at pH 8.6 and shown in Fig. 1, were stained with
134 either Coomassie Brilliant Blue (CBB) or specific polyclonal antibodies against α_{s2} , α_{s1} , β and κ -CN to
135 identify each casein fraction in the electrophoretic pattern. In the Urea-PAGE profiles, at least three
136 components exhibiting the highest mobility toward the anode and migrating head α_{s1} -CN were detected
137 as α_{s2} -CN after immunoblotting; each component accounted for 10, 11 and 12 P/mole as previously
138 reported (Chianese et al. 2010). The α_{s1} -CN fraction showed a complex heterogeneity, after
139 immunostaining with specific antibodies, since five main components were identified as α_{s1} -CN,
140 exhibiting an intermediate anodic mobility between donkey β - and α_{s2} -CN. The compositional
141 heterogeneity of donkey α_{s1} -CN could be due to different phosphorylation degree of its components as
142 well as the presence of deleted forms (Cunsolo et al. 2009a), as in mare counterparts (Miranda et al. 2004;
143 Mateos et al. 2009) as well as in ruminants (Martin et al., 2003). After immunodetection the β -CN was
144 constituted of two/three main components, differing for the phosphorylation degree (5, 6 and 7 P/mole)
145 (Chianese et al. 2010), as found in mare's milk also (Girardet et al. 2006).

146 The electrophoretic profiles stained with CBB were quantitatively evaluated by densitometric
147 analysis. Taking into account the high intensity of electrophoretic bands, the donkey β -CN may be the
148 most abundant casein fraction. Finally, the CBB stained bands, characterised by a lower negative charge
149 than β -CN, were identified after immunoblotting as κ -CN, without overlapping with the other casein

150 fractions. It is known that κ -CN components exhibited a weak intensity to CBB, both owing to the poor
151 susceptibility of this fraction to staining and low content in the casein micelle.

152 After densitometric analysis, β -CN was by far the most abundant casein fraction ($54.28\% \pm 5.68$),
153 followed by α_{s1} -CN ($35.59\% \pm 5.06$), a composition certainly closer to that of human than cow's milk.
154 This latter, in fact, is rich in α_{s1} and α_{s2} -caseins, that are lacking or present in traces in breastmilk. The
155 allergenic advantage of non-bovine milks, such as goat's and now donkey's milk, might be attributed to
156 this difference (Bevilacqua et al. 2001). The amounts of α_{s2} -CN ($7.19\% \pm 2.55$) and κ -CN ($2.79\% \pm 0.85$)
157 were the lowest among casein fractions. However, it is well known that these latter casein fractions
158 represent the minor components also in the horse (Miranda et al. 2004). In Table 1, the percentage and
159 relative amounts of each casein fraction in donkey were reported in comparison with pony horse, goat,
160 yak, cattle, buffalo and camel milk.

161 Compared with ruminants' milk, the relatively low level of caseins observed in DM coupled with
162 the low protein content may be responsible for the soft curd produced in the stomach. For example, a
163 similar condition was observed also in goat carriers of defective alleles. Goat milk lacking the α_{s1} -CN has
164 poor coagulation properties in comparison with milk containing α_{s1} -CN, and it also decreases intestinal
165 and systemic sensitization to β -lactoglobulin in guinea pigs (Bevilacqua et al. 2001).

166 Although with different values, the trend of the casein fraction content in donkey ($\beta > \alpha_{s1} > \alpha_{s2} >$
167 κ) is similar to that observed for camel (Kappeler et al. 1998), but different from those observed for horse,
168 yak and goat ($\beta > \alpha_{s1} > \kappa > \alpha_{s2}$) (Miranda et al. 2004; Bevilacqua et al. 2006; Bai et al. 2013), cattle ($\beta =$
169 $\alpha_{s1} > \alpha_{s2} > \kappa$) (Miranda et al. 2004) and buffalo ($\beta > \alpha_{s2} > \alpha_{s1} > \kappa$) (Cosenza et al. 2011).

170 These data confirm that the casein-type composition (as well as the protein/fat ratio) is different in
171 most dairy animals, and the physicochemical properties of the milk depend on it, both contributing to the
172 functionality of milk and playing an important role in cheese making (Roncada et al. 2012). It is well-
173 known that the different proportion of casein fractions, besides genetic variants and post-translational

174 modifications of caseins family, directly affect the conformation and the sizes of the micelles in the milk
175 from different dairy animals and, consequently the technological properties.

176

177 *Transcripts quantification and translation efficiency*

178 In order to quantify the mRNA transcribed from the casein genes of eight lactating donkeys, we used a
179 RT-qPCR approach using the 18S rRNA as housekeeping gene and a standard curve for a complete
180 quantification of transcripts. The obtained results show that the average percentage of donkey casein
181 transcripts were 6.28, 70.85, 14.23 and 8.65 for *CSN1S1*, *CSN2*, *CSN1S2 I* and *CSN3*, respectively (Table
182 2). These values are somewhat different from that observed for the transcripts of homologous genes in
183 buffalo species from Cosenza et al. (2011), in yak (Bai et al. 2013) and in cattle, goat and sheep
184 (Bevilacqua et al. 2006). In particular, for the latter four species each casein transcript represents nearly
185 20-30% of the whole casein transcript population, while the incidence rate of buffalo *CSN1S1*, *CSN1S2*
186 transcripts are higher than those observed in the donkey (Table 2).

187 In order to evaluate the translation efficiency of the donkey gene casein transcripts, the ratio
188 between the percentage of single milk casein fractions and the single percentage of transcripts produced
189 in the milk somatic cells has been estimated.

190 The values obtained show a low translation efficiency for the *CSN1S2 I* (0.50), *CSN3* (0.32) and
191 *CSN2* (0.76) transcripts, whereas much higher efficiency (5.66) was found for the *CSN1S1*. The trend of
192 donkey casein translation efficiency is almost similar to that observed by Bai et al. (2013) for the yak
193 (0.30, 0.6, 1.5 and 1.8 for *CSN1S2*, *CSN3*, *CSN2* and *CSN1S1*, respectively) and for cattle, goat and sheep
194 by Bevilacqua et al. (2006). In particular, for the latter species β - and α 1- casein mRNA showed the
195 highest translational efficiency, with ratio values 2.5- to 4-fold over the values recorded for α 2- and κ -
196 casein transcripts (Bevilacqua et al. 2006). These results differ from those obtained in river buffalo, where
197 *CSN3* (2.69), *CSN2* (2.39) and *CSN1S1* (1.31) are characterized by a higher translation efficiency, while
198 *CSN1S2* showed the lowest value (0.25) (Cosenza et al. 2011).

199 The molecular mechanisms responsible for the observed differences in the individual transcript
200 efficiency can be different. Each mRNA is represented by various sequence-derived and functional
201 features related to translation. In order to investigate whether the mRNA sequences might be responsible
202 for the observed differences, a comparison of nucleotide sequences with the Kozak consensus sequence
203 (GCCA/GCCAUGG) was accomplished. Kozak consensus sequence is an element highly conserved in
204 the eukaryotic genomes, which represents the most efficient context for the correct translation initiation
205 (Kozak, 1994). In particular, more the sequence around the initiation codon is homologous to the Kozak
206 sequence (i.e., “strong” consensus), higher should be the efficiency of mRNA translation (Kozak, 1984).
207 The sequence comparison of the four casein transcripts in donkey (Table 3) showed for the *CSN2*,
208 *CSNIS2 I* and *CSN3* mRNAs the highest homology with the Kozak sequence. In particular, *CSN2* is
209 characterized by four conservative nucleotides (-5, -3, -2 and -1) directly upstream of the initiation
210 (nucleotide ‘A’ in AUG is numbered +1 and the number increases further downstream). Three of them (-
211 3, -2 and -1) are consecutive residues, similar to *CSNIS2 I*, while *CSN3* is characterized by a tandem
212 conservative nucleotides (-2, -3 and -5, -6). On the contrary, *CSNIS1* showed the worst combination.
213 Despite three nucleotides match with the consensus sequence, these are not consecutive (-5, -3 and -1)
214 and, therefore, it can be considered as a “weak” context (Table 3).

215 These observations are, apparently, in contradiction with the values obtained for the efficiency of
216 translation. However, it is worth noting that donkey *CSN2*, *CSNIS2 I* and *CSN3* are each characterized by
217 a single nucleotide substitution with respect to the canonic Kozak sequence, such as the G→T in position
218 -6 for *CSN2*, G→A in position -6 and C→T in position -5 for *CSNIS2 I* and C→G in position -1 for
219 *CSN3* (Table 3). Different studies demonstrated that mutations in these positions of the Kozak consensus
220 site decreased the efficiency of translation, thus confirming the hypothesized key role of the nucleotides -
221 6, -5 and -1 in the optimization of the translation process (Afshar-Kharghan et al. 1999; Usuki &
222 Maruyama, 2000; De Angioletti et al. 2004). For example, the G localized in position -6 with respect to
223 the AUG, is present in 44% of the 699 vertebrate mRNA sequences analyzed (Kozak, 1987). This high
224 conservation suggests that the G at position -6 is also important in the initiation of translation (De

225 Angioletti et al. 2004). An outstanding example exists in rabbit, where the substitution of the G at -6 with
226 a T in the β -globin 5'UTR reduced the efficiency of the translation initiation process *in vitro* (Kozak,
227 1994). In addition, in human, *in vitro* transcription/translation experiments demonstrated that the
228 substitution of -6G with a C decreased the efficiency of translation of the β -globin chain by about 30%
229 translation (De Angioletti et al. 2004).

230 Similarly, a polymorphism 5 bp upstream of the initiation codon in the Kozak sequence directly
231 influenced the *CSN1S2* translation in Norwegian Red cattle (Sodeland et al. 2011). Furthermore, in mouse
232 and human, a SNP at position -1 is associated with a significant reduction of CD40 gene product and with
233 a reduction in the translation efficiency (Jacobson et al. 2005; Pineda et a. 2008), analogous to what we
234 observed for donkey *CSN3*. Mechanistically, SNPs occurring at position -1 of the Kozak consensus
235 sequence would interfere with the ability of the ribosome to initiate translation, although not affecting the
236 ability of RNA polymerase to transcribe mRNA (Jacobson et al. 2005).

237 The ORF length is another element potentially affecting the translation efficiency. Valleriani et al.
238 (2011) demonstrated that the translational ratio decreases with increasing mRNA length. In this respect,
239 the calcium-sensitive casein genes in donkeys showed a higher translation efficiency of the *CSN1S1* vs
240 *CSN2* and *CSN1S2* I genes, which is consistent with the length of their coding sequence: 212 codons
241 (GeneBank FN386610) vs 241 (GeneBank FN598778) and 236 (GeneBank FM946022), respectively.
242 Therefore, based on these data, it is reasonable to suppose that the reduced ORF length counteracts the
243 negative effect of the “weak consensus site” and the impact of the SNP in position -6 on the *CSN1S1*
244 translation efficiency.

245 The coding region length could also explain some of the differences in translation efficiency
246 observed among the species. Donkey *CSN1S2* I and *CSN2* transcripts, which show a lower translation
247 efficiency than the homologous genes in ruminants, are characterized by a higher coding sequence length.
248 In particular, 236 codons for the donkey *CSN1S2* I vs 223 of goat and sheep (GenBank NM_001285585,
249 NM_001009363, respectively) and vs 222 of cattle, buffalo and yak (GenBank NM_174528, FM865618
250 and XP_014335716, respectively). Similarly, 241 codons for the donkey *CSN2* vs 222 for goat

251 (AJ011018) and sheep (NM_001009373), vs 224 for cattle (KC993858), buffalo (FM946182) and yak
252 (ELR51814).

253 A common feature in all species examined is the relatively low efficiency of translation of
254 *CSNIS2* compared to *CSN2*. The analysis of the mammary tissue collected from yak, goats, sheep and
255 cows has revealed that *CSN2* and *CSNIS2* mRNA are expressed at similar levels, but the β -casein
256 accumulation in milk is 4-5 times that of the α_{s2} -casein (Bevilacqua et al. 2006; Bai et al. 2013). In the
257 mammary tissue of water buffalo, the *CSN2* and *CSNIS2* represent 23 and 56% of casein transcripts,
258 respectively, while their corresponding protein concentrations in milk are 54 and 5%, respectively, of
259 total caseins, indicating approximately 10-fold more efficient translation of *CSN2* (Cosenza et al. 2011).
260 Analogously, in donkey lactating mammary gland the *CSN2* and *CSNIS2* I transcripts represent
261 respectively 70.85 and 14.23% of the total casein mRNAs, while the corresponding protein concentration
262 is 54.28 and 7.19 % respectively, with a greater *CSN2* translation efficiency of about 1.5 times. In the
263 bovine species, Kim et al. (2015) show that the usage of the last 28 codons of *CSNIS2* is the main
264 regulatory element attenuating its expression, and it is responsible for the differential translational
265 expression of the *CSNIS2* and *CSN2*. In particular, the authors reported that the codon usage and order
266 influenced the accuracy and the speed of translation.

267 Although the analysis of the sequences flanking the start codon, codon usage frequencies and the
268 coding sequence length can help to formulate hypotheses concerning some of the observed differences in
269 translation efficiency, other elements need to be analysed to fully understand the regulation mechanisms
270 of their expression. Factors like gene ontology enrichment scores, biochemical and physicochemical
271 features, minimum free energy, 5'UTR and 3'UTR length, number of transcription factors known to bind
272 the promoter region, number of RNA binding proteins known to bind its mRNA product, protein
273 abundance, mRNA and protein half-life, might affect gene expression (Huang et al. 2011). By
274 simultaneously measuring translational efficiencies (thus indirectly levels of protein synthesis) and
275 mRNA abundance, global analyses have shown evidence of significant mRNA destabilization and
276 translational repression. Since only slightly more translational repression is observed than mRNA

277 destabilization, it is possible that most of the loss in protein synthesis could directly result from effects on
278 mRNA stability (Djuranovic et al. 2012).

279

280 **Conclusions**

281 DM was characterized by a lower protein content with respect to ruminants milk and the different
282 proportions of caseins were closer to the human casein-type composition. β -CN was predominant with
283 respect to the alpha (s1), which may reduce allergenicity. This compositional feature might be responsible
284 for the soft curd produced in the stomach, determining a better digestibility of DM than cow's milk.
285 Moreover, the casein composition of DM could also be decisive for using it as a substitute when breast-
286 feeding is not possible.

287 The results obtained showed also a significant difference in the expression of donkey casein genes,
288 which revealed dissimilar patterns in comparison to those of the main species of ruminants (cattle,
289 buffalo, sheep, goats and yak). These data represent an important first step in the understanding of the
290 mechanisms regulating the expression of these genes in donkeys aimed at improving the milk production,
291 which fulfill special consumer requirements..

292

293 Acknowledgements. Supported by the Italian Ministry of Agricultural and Forestry Policies (SELMOL
294 project).

295

296 **Conflict of interest**

297 None of the authors have any conflict of interest to declare.

298

299 **References**

300 **Aschaffenburg R & Drewry J** 1959 New procedure for the routine determination of the various non-
301 casein proteins of milk. In: XVth International Dairy Congress, London vol. 3, pp. 1631-1637

302 **Afshar-Kharghan V, Li CQ, Khoshnevis-Asl M. & Lopez JA** 1999 Kozak sequence polymorphism of
303 the glycoprotein (GP) Iba gene is a major determinant of the plasma membrane levels of the platelet
304 GP Ib-IX-V complex. *Blood* **94** 186-191

305 **Bai WL, Yin RH, Jiang WQ, Ajayi, OO, Zhao SJ, Lu GB, Zhao ZH & Imumorin IG** 2013
306 Molecular analysis of alpha(s1)-, beta-, alpha(s2)- and kappa-casein transcripts reveals differential
307 translational efficiency in yak lactating mammary gland. *Livestock Science* **152** 74-78

308 **Bevilacqua C, Martin P, Candalh C, Fauquant J, Piot M, Roucayrol AM, Pilla F & Heyman M**
309 2001 Goats' milk of defective alpha(s1)-casein genotype decreases intestinal and systemic sensitization
310 to beta-lactoglobulin in guinea pigs. *Journal of Dairy Research* **68** 217-227

311 **Bevilacqua C, Helbling JC, Miranda G & Martin P** 2006 Translational efficiency of casein transcripts
312 in the mammary tissue of lactating ruminants. *Reproduction Nutrition Development* **5** 567-578

313 **Businco L, Giampietro PG, Lucenti P, Lucaroni F, Pini C, Di Felice G, Lacovacci P, Curadi C &**
314 **Orlandi M** 2000 Allergenicity of mare's milk in children with cow's milk allergy. *Journal of Allergy*
315 *and Clinical Immunology* **105** 1031-1034

316 **Chianese L, Quarto M, Pizzolongo F, Calabrese MG, Caira S, Mauriello R, De Pascale S, & Addeo**
317 **F** 2009 Occurrence of genetic polymorphism at the α_{s1} -casein locus in Mediterranean water buffalo
318 milk. *International Dairy Journal* **19** 181-189

319 **Chianese L, Calabrese MG, Ferranti P, Mauriello R, Garro G, De Simone C, Quarto M, Addeo F,**
320 **Cosenza G & Ramunno L** 2010 Proteomic characterization of donkey milk "caseome" *Journal of*
321 *Chromatography A* **1217** 4834-4840

322 **Cosenza G, Pauciullo A, Annunziata AL, Rando A, Chianese L, Marletta D, Iannolino G, Nicodemo**
323 **D, Di Bernardino D & Ramunno L** 2010 Identification and characterization of the donkey *CSNIS2* I
324 and II cDNAs and polymorphisms detection. *Italian Journal of Animal Science* **9 (e40)** 206-211

325 **Cosenza G, Pauciullo A, Coletta A, Di Francia A, Feligini M, Gallo D, Di Bernardino D & Ramunno**
326 **L** 2011 Translational efficiency of casein transcripts in Mediterranean river buffalo. *Journal of Dairy*
327 *Science* **94 (11)** 5691-5694

- 328 **Cosenza G, Pauciullo A, Gallo D, Colimoro L, D'avino A, Mancusi A & Ramunno L** 2008
329 Genotyping at the CSN1S1 locus by PCR-RFLP and AS-PCR in a Neapolitan Goat Population. *Small*
330 *Ruminant Research*, **74** 84–90
- 331 **Criscione A, Cunsolo V, Bordonaro S, Guastella AM, Saletti R, Zuccaro A, D'Urso G & Marletta D**
332 2009 Donkeys' milk protein fraction investigated by electrophoretic methods and mass spectrometric
333 analysis. *International Dairy Journal* **19** 190-197
- 334 **Cunsolo V, Cairone E, Fontanini D, Criscione A, Muccilli V, Saletti R & Foti S** 2009a Sequence
335 determination of α s1-casein isoforms from donkey by mass spectrometric methods. *Journal of Mass*
336 *Spectrometry* **44** 1742-1753
- 337 **Cunsolo V, Cairone E, Saletti R, Muccilli V & Foti S** 2009b Sequence and phosphorylation level
338 determination of two donkey beta-caseins by mass spectrometry. *Rapid Communication in Mass*
339 *Spectrometry* **23** 1907-1916
- 340 **Cunsolo V, Saletti R, Muccilli V, Gallina S, Di Francesco A & Foti S** 2017 Proteins and bioactive
341 peptides from donkey milk: The molecular basis for its reduced allergenic properties. *Food Research*
342 *International* **99** 41-47
- 343 **D'Alessandro AG, Rosanna De Petro R, Claps S, Pizzillo M & Martemucci G** 2009 Yield and quality
344 of milk and udder health in Martina Franca ass: effects of daily interval and time of machine milking.
345 *Italian Journal of Animal Science* **8 (2)** 697-699
- 346 **De Angioletti M, Lacerra G, Sabato V & Carestia C** 2004 β +45 G \rightarrow C: A novel silent β -
347 thalassaemia mutation, the first in the Kozak sequence. *British Journal of Haematology* **124** 224-231
- 348 **Djuranovic S, Nahvi A & Green R** 2012 miRNA-mediated gene silencing by translational repression
349 followed by mRNA deadenylation and decay. *Science* **336 (6078)** 237-240
- 350 **Giorgis V, Rolla G, Raie A, Geuna M, Boita M, Lamberti C, Nebbia S, Giribaldi M, Giuffrida M &**
351 **Brussino L** 2018 A case of work-related donkey milk allergy. *Journal of Investigational Allergology*
352 *and Clinical Immunology* **28(3)** 197-199

353 **Girardet JM, Miclo L, Florent S, Mollè D & Gaillard JL** 2006 Determination of the phosphorylation
354 level and deamination susceptibility of equine β -casein. *Proteomics* **6** 3707-3717

355 **Guo HY, Pang K, Zhang XY, Zhao L, Chen SW, Dong ML & Ren FZ** 2007 Composition,
356 physiochemical properties, nitrogen fraction distribution, and amino acid profile of donkey milk.
357 *Journal of Dairy Science* **90** 1635-1643

358 **Jacobson EM, Concepcion E, Oashi T & Tomer Y** 2005 A Graves' disease-associated Kozak sequence
359 single-nucleotide polymorphism enhances the efficiency of CD40 gene translation: a case for
360 translational pathophysiology. *Endocrinology* **146(6)** 2684-2691

361 **Hobor S, Kunej T & Dovc P** 2008 Polymorphisms in the kappa casein (CSN3) gene in horse and
362 comparative analysis of its promoter and coding region. *Animal Genetics* **39** 520-530

363 **Huang T, Wan S, Xu Z, Zheng Y, Feng KY, Li HP, Kong X & Cai YD** 2011 Analysis and prediction
364 of translation rate based on sequence and functional features of the mRNA. *PLoS One* **6 (1)** e16036

365 **IDF** (1993). *Milk. Determination of nitrogen content. Standard method 20B*. Brussels, Belgium:
366 International Dairy Federation

367 **Kappeler S, Farah Z & Puhan Z** 1998 Sequence analysis of *Camelus dromedarius* milk caseins.
368 *Journal of Dairy Research* **65 (2)** 209-222

369 **Kim JJ, Yu J, Bag J, Bakovic M & Cant JP** 2015 Translation attenuation via 3' terminal codon usage in
370 bovine CSN1S2 is responsible for the difference in α 2- and β -casein profile in milk. *RNA Biology* **12**
371 **(3)** 354-367

372 **Kozak M** 1984 Point mutations close to the AUG initiator codon affect the efficiency of translation of rat
373 preproinsulin in vivo. *Nature* **308 (5956)** 241-246

374 **Kozak M** 1987 Analysis of 50-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic*
375 *Acids Research* **15** 8125-8148

376 **Kozak M** 1994 Features in the 5' non-coding sequences of rabbit α and β -globin mRNAs that affect
377 translational efficiency. *Journal of Molecular Biology* **235** 95-110

- 378 **Lonnerdal B** 1994 Nutritional Importance of Non-protein Nitrogen Protein. Metabolism During Infancy.
379 In Nestle Nutrition Workshop Series edited by Niels C. R. Raiha, **33** (Ed. Nestec Ltd., Vevey/ Raven
380 Press, Ltd., New York)
- 381 **Malacarne M, Martuzzi F, Summer A & Mariani P** 2002 Protein and fat composition of mare's milk:
382 some nutritional remarks with reference to human and cow's milk. *International Dairy Journal* **12**
383 869-868
- 384 **Martin P, Ferranti P Leroux C & Addeo F** 2003 Non-bovine caseins: quantitative variability and
385 molecular diversity. In Advanced Dairy Chemistry Volume 1: Proteins, 3rd edn. Edited by PF Fox and
386 PLH McSweeney, Kluwer Academic/Plenum Publishers
- 387 **Mateos A, Miclo L, Molle D, Dary A, Girardet JM & Gaillard JL** 2009 Equine α 1- casein:
388 characterization of alternative splicing isoforms and determination of phosphorylation levels. *Journal*
389 *Dairy Science* **92** 3604-3615
- 390 **Miranda G, Mahé MF, Leroux C & Martin P** 2004 Proteomic tools to characterize the protein fraction
391 of Equidae milk. *Proteomics*, **4** 2496-2509
- 392 **Monti G, Viola S, Baro C, Cresi F, Tovo PA, Moro G & Bertino E** 2012 Tolerability of donkey's milk
393 in 92 highly-problematic cow's milk allergic children. *Journal of Biological Regulators and*
394 *Homeostatic Agents* **26 (3 Suppl)** 75-82
- 395 **Ochirkhuyag B, Chobert JM, Dalgalarrrondo M & Haertlé T** 2000 Characterization of mare caseins.
396 Identification of α 1- and α 2-caseins. *Lait* **80** 223-235
- 397 **Pineda B, Laporta P, Hermenegildo C, Cano A & García-Pérez MA** 2008 A C >T polymorphism
398 located at position -1 of the Kozak sequence of CD40 gene is associated with low bone mass in
399 Spanish postmenopausal women. *Osteoporosis International* **19(8)** 1147-1152
- 400 **Roncada P, Piras C, Soggiu A, Turk R, Urbani A & Bonizzi L** 2012 Farm animal milk proteomics.
401 *Journal of Proteomics* **75 (14)** 4259-4274
- 402 **Salimei E, Fantuz F, Coppola R, Chiofalo B, Polidori P & Varisco G** 2004 Composition and
403 characteristics of asses'milk. *Animal Research* **53** 67-78

404 **Sodeland M, Grove H, Kent M, Taylor S, Svendsen M, Hayes BJ & Lien S** 2011 Molecular
405 characterization of a long range haplotype affecting protein yield and mastitis susceptibility in
406 Norwegian Red cattle. *BMC Genetics* **12** 70

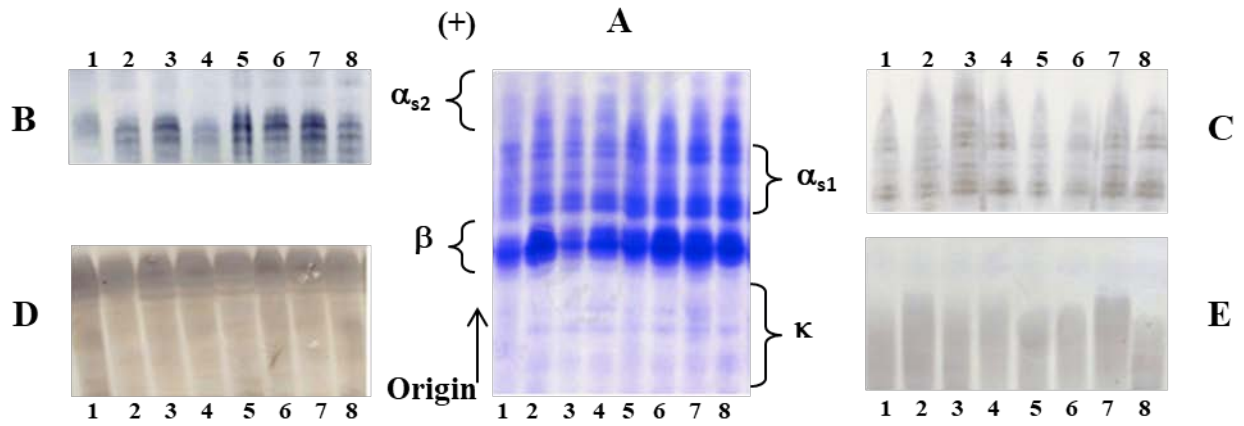
407 **Usuki F & Maruyama K** 2000 Ataxia caused by mutations in the alpha-tocopherol transfer protein gene.
408 *Journal of Neurology, Neurosurgery and Psychiatry* **69** 254-256

409 **Valleriani A, Zhang G, Nagar A, Ignatova Z & Lipowsky R** 2011 Length-dependent translation of
410 messenger RNA by ribosomes. *Physical review. E, Statistical, nonlinear, and soft matter physics* **83** (4
411 **Pt 1**) 042903

412

413 **Figure 1.**
414 PAGE analysis at pH 8.6 of the donkey's casein samples, after CBB staining (A) and identification
415 of the four casein fractions by immunoblotting with polyclonal antibodies against α_{s2} (B), α_{s1} (C),
416 β (D) and κ -CN (E).
417

418



419

Table 1.

Total casein and caseins' fraction content in DM in comparison with pony horse, cattle, buffalo, goat, yak and camel milk.

Species	α_{s1}		β		α_{s2}		κ		Total casein mg mL ⁻¹
	%	mg mL ⁻¹	%	mg mL ⁻¹	%	mg mL ⁻¹	%	mg mL ⁻¹	
Donkey ^a	35.59	1.82	54.28	2.77	7.19	3.68·10 ⁻¹	2.79	1.42·10 ⁻¹	5.12
Pony Horse ^b	17.92	2.50	78.85	11.00	1.43	0.20	1.80	0.25	13.95
Cattle ^b	36.77	10.00	36.77	10.00	13.69	3.70	12.86	3.50	27.20
Buffalo ^c	16.19	7.62	42.08	19.81	32.70	15.39	9.03	4.25	47.07
Goat ^b	26.12	7.00	41.05	11.00	15.67	4.20	17.16	4.60	26.80
Yak ^d	30.80	10.50	48.20	16.50	8.70	2.90	12.30	4.20	34.10
Camel ^e	22.00	5.20	65.00	15.60	9.60	2.30	3.30	0.80	24.00

^a Present work

^b Miranda et al. (2004)

^c Cosenza et al. (2011)

^d Bai et al. (2013)

^e Kappeler et al. (1998)

Table 2.

Comparison of average quantitative transcript levels for α_{s1} - (*CSN1S1*), β - (*CSN2*), α_{s2} - (*CSN1S2*) and κ -casein (*CSN3*) in donkey and in the main ruminant species.

Species	<i>CSN1S1</i> (%)	<i>CSN2</i> (%)	<i>CSN1S2</i> (%)	<i>CSN3</i> (%)
Donkey ^a	6.28 ± 1.93	70.85 ± 8.96	14.23 ± 6.82	8.65 ± 1.21
Cattle, sheep, goat ^b	~ 25	~ 25	~ 25	~ 25
Buffalo ^c	16.48 ± 4.99	23.18 ± 5.41	55.87 ± 8.22	4.47 ± 0.96
Yak ^d	17.5 ± 1.80	31.9 ± 1.90	29.6 ± 2.50	20.9 ± 2.10

^a Present work

^b Bevilacqua et al. (2006)

^c Cosenza et al. (2011)

^d Bai et al. (2013)

Table 3.

Comparison of start codon flanking sequences of the 4 casein transcripts in donkey.

Position ¹										Sequence ²
-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	
<u>G</u>	<u>C</u>	C	<u>R</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>G</u>	G	Kozak consensus sequence
U	<u>C</u>	A	<u>G</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>G</u>	A	<i>CSN2</i>
A	U	A	<u>A</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>G</u>	A	<i>CSNIS2 I</i>
A	<u>C</u>	A	<u>A</u>	<u>G</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>G</u>	A	<i>CSNIS1</i>
<u>G</u>	<u>C</u>	A	<u>A</u>	<u>C</u>	G	<u>A</u>	<u>U</u>	<u>G</u>	A	<i>CSN3</i>

¹The start codon (AUG) in the four casein transcripts is underlined; gray colour identifies a conserved nucleotide in comparison with the Kozak consensus sequence.

²Kozak consensus sequence = the optimal context for initiation of translation in mammals. *CSN2*, *CSNIS2 I*, *CSNIS1* and *CSN3* are the genes encoding β , α_{s2} , α_{s1} and κ -casein, respectively.