

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

**Remarkable genetic diversity detected at river buffalo prolactin receptor (PRLR) gene and association studies with milk fatty acid composition**

**This is the author's manuscript**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1664866> since 2021-11-15T09:27:25Z

*Published version:*

DOI:10.1111/age.12645

*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 **Remarkable genetic diversity detected at river buffalo prolactin receptor (*PRLR*) gene and**  
2 **association studies with milk fatty acid composition**

3 G. Cosenza,\* M. Iannaccone\*,<sup>1</sup> B. Auzino,<sup>†</sup> N.P.P Macciotta,<sup>‡</sup> A. Kovitvadhi,<sup>§</sup> I. Nicolae,<sup>#</sup> and A.  
4 Pauciullo <sup>||</sup>

5

6 \* Department of Agricultural Sciences, University of Naples “Federico II”, 80055, Portici (NA), Italy

7 <sup>†</sup> Department of Veterinary Sciences, University of Pisa, 56124, Pisa, Italy

8 <sup>‡</sup> Department of Agricultural Sciences, University of Sassari, 07100, Sassari, Italy

9 <sup>§</sup> Department of Physiology, Kasetsart University, 10900, Bangkok, Thailand

10 <sup>#</sup> Research and Development Institute for Bovine Breeding, 077015, Balotesti, Bucharest, Romania

11 <sup>||</sup> Department of Agricultural, Forest and Food Science, University of Torino, 10095, Grugliasco  
12 (TO), Italy

13

14 Buffalo *PRLR* variability and milk FA association

15

16 <sup>1</sup> Corresponding author: [m.iannaccone@unina.it](mailto:m.iannaccone@unina.it)

17

18

19 **Summary**

20 Prolactin is an anterior pituitary peptide hormone involved in many different endocrine activities and  
21 is essential for reproductive performance. This action is mediated by its receptor, the prolactin  
22 receptor, encoded by the *PRLR* gene. In this study, we sequenced and characterized the Mediterranean  
23 river buffalo *PRLR* gene (from the exon 3 to 10) and a remarkable genetic diversity was found. In  
24 particular, we found 24 intronic polymorphisms and 13 exonic SNPs, 7 of which were non-  
25 synonymous. Furthermore, the polymorphisms identified in the 3'UTR were investigated to establish  
26 their possible influence on microRNA binding sites. Considering all the amino acid changes and the  
27 observed allelic combinations, it is possible to deduce at least 6 different translations of the buffalo  
28 prolactin receptor and, consequently, the presence at the *PRLR* gene of at least 6 alleles. Furthermore,  
29 we identified a deletion of a CACTACC heptamer between the nucleotides 1102-1103 of exon 10  
30 (3'UTR) and we developed an allele specific PCR to identify the carriers of this genetic marker.

31 Finally, the SNP g.11188A>G, detected in exon 10 and responsible for the amino acid  
32 replacement p.His328Arg, was genotyped in 308 Italian Mediterranean river buffaloes and an  
33 association study with milk fat traits was carried out. The statistical analysis showed a trend that  
34 approached the significance for the AA genotype with higher contents of odd branched-chain fatty  
35 acids (OBCFA). Thus, our results suggest that *PRLR* gene is a good candidate for gene association  
36 studies with qualitative traits related to buffalo milk production.

37

38 **Keywords:** Mediterranean river buffalo, *PRLR*, alleles, association, milk, OBCFA

39

## 40 **Introduction**

41 Prolactin (encoded by *PRL* gene) is peptide hormone produced by anterior pituitary and it plays a role  
42 in many different endocrine activities and is essential for reproductive performance. These actions  
43 are mediated by the prolactin receptor, which has been detected in various tissues including brain,  
44 ovary, placenta and uterus of several mammalian species. The prolactin receptor, encoded by *PRLR*  
45 gene, belongs to the growth hormone/prolactin receptor gene family sharing high homology sequence  
46 regions.

47 In cattle, *PRLR* gene is mapped on chromosome 20 (20q17) (Hayes *et al.*, 1996). It is made  
48 by 10 exons and 3 of them can be alternative spliced with at least 2 different known isoforms, long  
49 and short (Lü *et al.*, 2011). Both the isoforms are differentially expressed and regulated during estrous  
50 cycle and pregnancy, indicating a role in different metabolic pathways (Buck *et al.*, 1992; Clarke *et*  
51 *al.*, 1993, Clarke & Linzer 1993). In particular, the long isoform of *PRLR* binds *PRL* and contributes  
52 to activation of the transcription factor JAK2 (janus kinase 2) and subsequent phosphorylation of  
53 STAT5 (signal transducer and activator of transcription 5), which recognize sequences located in  
54 promoters of milk protein genes (Bole-Feysot *et al.* 1998). In ruminants, the long and short isoforms  
55 encode for a protein of 581 and 296 amino acids, respectively. The short protein is originated by an  
56 alternative splicing of 39 bp (in the exon 10) that generate 2 close stop codons (Bignon *et al.* 1997;  
57 Lü *et al.*, 2011).

58 Several studies have highlighted the influence of *PRLR* on the reproductive performance in  
59 sheep and pig (Tomás *et al.* 2006; Chu *et al.*, 2007), the induction of maternal behavior in sheep  
60 (Wang *et al.*, 2015) and the litter size in goat (Hou *et al.*, 2015). Moreover, *PRLR* expression is  
61 important during mouse mammary gland development (Bole-Feysot *et al.* 1998).

62 In cattle, quantitative trait *locus* (QTL) on chromosome 20 and *PRLR* gene polymorphisms  
63 have been associated with milk production traits (Zhang *et al.* 2008; Iso-Touru *et al.* 2009, Lü *et al.*,  
64 2011), as well as in pig (Skrzypczak *et al.*, 2015) and sheep (Hou *et al.*, 2013, 2014). Based on these

65 studies, *PRLR* gene is considered as a strong candidate for gene association studies with quantitative  
66 and qualitative traits related to milk production.

67 Recently, different studies have shown that the *PRLR* gene is polymorphic also in *Bubalus*  
68 *bubalis* (mainly Murrah river buffalo) (Javed *et al.*, 2011a,b; Shi *et al.*, 2012), but to date none of the  
69 polymorphisms identified have been investigated for association with milk production traits.

70 Therefore, the aim of the present research was to identify and analyse the variability of the  
71 *PRLR* locus in Mediterranean river buffalo reared in Italy and test the possible associations between  
72 polymorphisms and milk fat traits in order to contribute to the improvement of the productive  
73 efficiency of the breed.

74

## 75 **Materials and Methods**

### 76 *Sample collection DNA isolation and PCR amplification conditions*

77 Individual blood samples were collected from 308 female Mediterranean river buffaloes belonging  
78 to 14 farms, located in Salerno and Caserta province (Southern Italy) plus 150 samples collected from  
79 Romanian Murrah (100) and Thai Mehsana (50). Sampling of Italian river buffaloes was carried out  
80 in collaboration with the Italian National Association of Buffalo Breeders (ANASB).

81 DNA was isolated from leukocytes, using the procedure described by Goossens & Kan (1981).  
82 DNA concentration and OD<sub>260/280</sub> ratio of the samples were measured by the Nanodrop ND-2000C  
83 Spectrophotometer (Thermo Scientific).

84 Using primers designed on bovine (GenBank acc. no. AF426315.1), bubaline genome  
85 sequence (NW\_005785465) and bubaline mRNA sequence (GenBank acc. no. GQ339914.1), we  
86 amplified the DNA region of the *PRLR* gene spanning from the exon 3 to the exon 10 in ten individual  
87 Mediterranean river buffaloes reared in 10 different farms. Additional primers, designed on newly  
88 determined sequences, were also used for sequencing. PCR products were sequenced on both strands  
89 at CEINGE - Biotecnologie Avanzate (Naples, Italy).

90 The estimation of SNPs frequency at exon 10 and the determination of possible haplotypes  
91 was accomplished by the re-sequencing of additional 48 and 10 samples of Mediterranean and  
92 Mehsana breeds, respectively. All sequenced samples were randomly selected.

93 PCR Reaction mix and thermal condition are reported in supplementary materials.

94 The entire panel of 308 Mediterranean river buffalo DNA samples was genotyped in  
95 outsourcing (KBiosciences, Herts, UK, <http://www.kbioscience.co.uk>) for the SNP g.11188A>G  
96 (Nuova Tabella)

97

98 *Allele Specific-Polymerase Chain Reaction (AS-PCR) for the heptamer detection*

99 An allele-specific PCR (AS-PCR) was developed to identify carriers of the CACTACC heptamer  
100 between nucleotides 1102 and 1103 of exon 10 (3'UTR) of *PRLR* gene in river buffalo. The two  
101 allele-specific reverse primers differ for absence (5'-AACTGGGAGTTGTCATTCTAA-3') or  
102 presence (5'-AACTGGGAGTTGGGTAGTG-3') of the sequence corresponding to the insertion  
103 (herein underlined). The allele-specific reverse primers and the forward common primer sequence  
104 (5'-CTCTGCTAAACCCTTGGAAT-3') are part of the exon 10. Primers were designed according  
105 to the newly determined sequences and by the comparison with the corresponding goat  
106 (XM\_018065467.1), sheep (KC734660.1) and bovine (AJ966356.4) sequences. Reaction mix and  
107 thermal condition were performed as reported above.

108 Genotyping was carried out on buffaloes belonging to the Italian Mediterranean (300), Murrah  
109 (100) and Mehsana (50) buffalo breeds, as well as cattle (50), sheep (50) and goats (50) belonging to  
110 different breeds/genetic types randomly chosen (Nuova Tabella).

111

112 *Bioinformatic and statistical analysis*

113 Allelic frequencies and Hardy-Weinberg equilibrium ( $\chi^2$  test) were calculated. Homology searches,  
114 comparison among nucleotide and amino acid sequences, multiple alignments for polymorphism  
115 discovery were accomplished using DNAsis-Pro (Hitachi Software Engineering Co., Japan).

116 Measures of linkage disequilibrium ( $D'$  and  $r^2$ ) were estimated using Haploview software ver.  
117 4.2 (<http://www.broadinstitute.org/haploview/haploview>). Haplotype structure was defined  
118 according to Gabriel *et al.* (2002).

119 Associations between *PRLR* genotype, fatty acid composition traits, fatty acid classes and fat  
120 percentage were tested using a mixed linear model (Supplementary materials).

121 MicroRNAs screening was performed using TargetScan (Lewis *et al.*, 2005) based on  
122 homology between our sequence and the bovine 3'UTR of *PRLR* gene.

123

## 124 **Results and Discussion**

### 125 *Structure and analysis of Mediterranean river buffalo PRLR gene*

126 Using genomic DNA from 10 Mediterranean river buffaloes reared in Campania, we sequenced the  
127 coding region of prolactin receptor gene (GeneBank accession number MF461277).

128 The sequenced region (over 11900 bp, including 1138 bp of exonic regions) spans from the  
129 exon 3 to the exon 10 (excluding mostly of intron 5, about 12000 bp). The sequenced *PRLR* exons  
130 and introns have different length: from 70-100 bp for the exons 9 and 8 to over 1180 bp for the exon  
131 10, while introns start from 693 bp (intron 8) to 2756 bp (intron 3). Splicing acceptor and donor sites  
132 followed the rule 5' GT/ 3'AG (Figure S1). The homology with the bovine counterpart is 95% with  
133 a G/C content of 45%.

134 Signal peptide is 24 amino acid long (MKENAASRVVFILLFLNFSLLNG) and it is  
135 encoded by the last 70 nt of exon 3 and the first 2 nt of exon 4. The stop codon (TGA) is localized  
136 between the nt 873-875 of the exon 10, which encodes for the majority of intracellular domain. This  
137 structure is highly conserved among mammals and, depending on the size of cytoplasmic domain,  
138 two mRNAs named short and long have been described (Bignon *et al.*, 1997; Zhang *et al.*, 2008). The  
139 only difference between short and long coding sequences was the presence or absence, respectively,  
140 of an insert 39 bp long at the beginning of the cytoplasmic domain. The insertion creates two  
141 contiguous *in frame* stop codons at its 3' end. As described by Bignon *et al.* (1997), this insertion,

142 present also in Mediterranean river buffalo (Figure S1), is an alternative splicing strategy generated  
143 by the use of a splice donor site at 40–41 bp upstream the exon 10. As consequence, the short and  
144 long form are 296 aa and 581 aa, respectively.

145

#### 146 *Intronic regions analysis and polymorphisms detection in Mediterranean river buffalo*

147 Alignments of the intron sequences among the 10 subjects showed 24 polymorphisms (6  
148 transversions, 15 transitions and 3 deletions/insertions) (Table 1) and several microsatellites that were  
149 not polymorphic in these samples. However, none of these polymorphisms would seem to affect  
150 expression regulation sequences (splicing sites, enhancers/silencers).

151 As it was already demonstrated for several ruminant (Ramunno *et al.*, 2004) and non-ruminant  
152 species (Pauciullo *et al.*, 2013; Pauciullo *et al.*, 2014), also the buffalo *PRLR* gene is characterized  
153 by the presence of several retroposons located in the introns and representing about 13% of the entire  
154 gene.

155 According to Lenstra *et al.* (1993), three different types of artiodactyla retroposons have been  
156 described. In particular, we have identified 6 retroposon sequences: A, B, C, D, E and F in buffalo  
157 *PRLR* gene (Figure S1). The elements A and B are localized in the intron 4. In particular, sequence  
158 A (between nt 1338 and 1600, taking as reference the first nucleotide of the corresponding intron) is  
159 a Bov-A2 element-like, while the sequence B (from nt 1648 to nt 1880) shows high homology with  
160 the retroposon Bov-tA2. The element C is located in the intron 6 (Bov-tA2 like, from 457 to 678),  
161 while the sequence D is in the intron 7 from nt 388 to nt 725 (Bov-B like). The last elements E (Bov-  
162 tA2, from 317 to 496) and F (Bov-A2, between nt 241 and 522) are located in the introns 8 and 9  
163 respectively. The elements A, B, C and F are flanked by imperfect direct repeats, underlined in Figure  
164 S1. All the aforementioned retroposons are well conserved among ruminants for which genomic  
165 sequences are available (homology higher than 94%).

166

#### 167 *Exon polymorphism analysis in Mediterranean and Mehsana*



168 The alignment of *PRLR* exons from the 10 investigated subjects of Mediterranean river buffalo  
169 identified 13 polymorphisms, 7 of which were non-synonymous SNP. In particular, the first two of  
170 the 13 polymorphisms identified, are located at the exon 3 and 7 respectively, while the remaining at  
171 10<sup>th</sup> exon (Table 1). Two of these mutations (g.114T>G and g.11188A>G, responsible for the aa  
172 changes p.Phe19Val and p.His328Arg, respectively) have already been reported in Indian water  
173 buffalo breeds by Javed *et al.* (2011a).

174 Moreover, the comparison between the sequences available in GenBank for the Mediterranean  
175 river buffalo (GenBank NW\_005785465, XR\_328736.1, XM\_006071901.1, XM\_006071900.1)  
176 pointed out other two polymorphic sites in exon 10: g.11465T>C and g.11563C>T, the latter  
177 responsible for the amino acid change p.Ser453Leu. All the buffaloes investigated were  
178 monomorphic for the last two *loci*,

179 Beside the polymorphisms observed for the Mediterranean buffalo, the sequencing of the exon  
180 10 in ten Mehsana samples showed a new non-synonymous SNP: g.11206C>T responsible for the  
181 amino acid change p.Thr334Met (Table SX).

182 Another non-synonymous mutation at the exon 10 of the *PRLR* gene has been described by  
183 Javed *et al.* (2011b) for Indian water buffalo breeds (Murrah and Mehsana): the transition  
184 g.11488A>G responsible for the amino acid replacement p.Ser428Gly. However, this polymorphism  
185 was not observed in any of the subjects examined in the present study.

186 Based on the comparison with the sequences available in GenBank for the main species  
187 belonging to the order Cetartiodactyla and Perissodactyla (Table Sx), at least 7 polymorphisms  
188 (g.114T>G, g.11188A>G, g.11206C>T, g.11563C>T, g.11687A>G, g.11882G>A, g.12100A>G)  
189 showed particular variability among the different mammalian species, so that haplotypes specific of  
190 specie/order/suborder can be identified .

191 Furthermore, comparing our sequences with those reported in Table Sx, two polymorphisms  
192 at the buffalo *PRLR* locus also characterize other species of ruminants and, therefore they may  
193 represent possible examples of trans-specific polymorphisms (TSPs).

194 The first of these mutations (g.11188A>G responsible for the change p.His328Arg) also  
195 characterizes the *PRLR locus* in goat: CGC<sup>Arg328</sup> (GenBank HQ667935, NM\_001285669); CAC<sup>His3288</sup>  
196 (GenBank KC906265, JX087440). The second polymorphism (g.11936G>A, p.Pro577) was also  
197 reported in cattle (FJ901289; Iso-Touru *et al.*, 2009).

198 Moreover, the comparison with the sequences available in GenBank shows a further possible  
199 TSP for the *PRLR locus* which, however, affects only the ovine and caprine species. This is a silent  
200 transition A>G in position 276 of 10<sup>th</sup> exon (Pro<sup>377</sup>): goat A<sup>276</sup> (GenBank EU662222, EU662223),  
201 goat G<sup>276</sup> (GenBank NM\_001285669, JX087440), sheep A<sup>276</sup> (GenBank FJ901301, AF041979),  
202 sheep G<sup>276</sup> (GenBank HM437209, HM437208).

203 TSPs are ancient genetic variants whose origin predates speciation events, resulting in shared  
204 alleles between evolutionarily related species (Klein *et al.*, 1987). TSPs can be explained by three  
205 main reasons: 1) SNP survival due to random chance, 2) coincidental mutations occurred after  
206 speciation and 3) balancing selection. TSPs as striking signature of persistent balancing selection are  
207 not very common (Hedrick, 2006). Examples of trans-specific polymorphisms have been reported in  
208 different species, including humans. To date, putative TSPs between humans and other primate  
209 species have been identified for the highly polymorphic major histocompatibility complex (MHC),  
210 the histo-blood ABO group, two antiviral genes (ZC3HAV1 and TRIM5), an autoimmunity-related  
211 gene LAD1 and several non-coding genomic segments with a putative regulatory role (Azevedo *et*  
212 *al.*, 2015). On contrary, there are very few documented examples of TSPs in ruminants. Only recently,  
213 Cosenza *et al.* (2017a) detected a first possible example of TSPs in the oxytocin-neurophysin I  
214 hormone encoding gene (*OXT*) in goat and sheep. Therefore, the identification of the detected  
215 polymorphisms could represent a new report of TSPs at DNA-level in ruminants and the first in  
216 Mediterranean river buffalo.

217 Moreover, the further comparison of the obtained sequences and the sequences available in database  
218 for the main ruminant species shows for the Mediterranean and Mehsana breeds the deletion of the  
219 CACTACC heptamer between the nucleotides 1102 and 1103 of 10th exon (3'UTR).

220 The presence of the heptamer might represent the ancestral condition of the gene, because it  
221 has also been found in the main non-ruminant species belonging to order Cetartiodactyla and in  
222 Perissodactyla (Figure S2). In particular, the members of suborder Tylopoda and Suiformes are  
223 characterized by a similar molecular event: CACTACCA and CACTACCC insertions, respectively,  
224 while members of suborder Hippomorpha and Ceratomorpha (order Perissodactyla) are characterized  
225 by CACTACTA (Figure S2).

226 In order to verify the specificity of this genetic marker, an allele specific PCR method was  
227 developed. By means of this method, the *PRLR* homozygote samples for the insertion showed  
228 successful PCR amplification (462 bp) only using the reverse primer for the “insertion”. Conversely,  
229 the homozygote samples for “non-insertion” allele (455 bp) are successfully amplified only by reverse  
230 primer with no insertion.

231 The genotyping of buffaloes from Mediterranean (300), Murrah (100) and Mehsana (50)  
232 breeds (all belonging to the river type) in comparison to cattle (50), sheep (50) and goats (50)  
233 belonging to different breeds/genetic types randomly chosen, confirmed that only our buffaloes  
234 samples (450) are characterized by the absence of the heptamer (Figure 1). However, our results from  
235 Murrah and Mehsana breeds conflict with a *PRLR* genomic sequence deposited in GenBank  
236 (GQ339914.1) by Javed *et al.* (2011) for Indian buffalo breeds (Murrah, Bhadawari, Tarai,  
237 Pandharpuri, Marathwada and Mehsana). Considering the conflicting information, we can not claim  
238 that the eptamer deletion is peculiar also of Murrah and Mehsana breeds, but we can surely state that  
239 it is specific for the Mediterranean breed.

240

#### 241 *Observed haplotypes at the river buffalo PRLR locus*

242 In order to estimate the frequency of the 11 polymorphic sites identified at the exon 10 of the  
243 Mediterranean river buffalo *PRLR* gene and to determine the possible haplotypes, the entire exon was  
244 sequenced a total of 68 samples (58 Mediterranean subjects reared in Southern Italy and 10 Mehsana  
245 buffaloes reared in Thailand) (Nuova tabella).

246 Three different allelic combinations were observed: haplotype 1, 2 and 3. Of these, the  
247 haplotype 1 is the most represented with a frequency of 0.62, followed by the haplotype 2 (0.25) and  
248 3 (0.13) (Figure 2).

249 According to Gabriel *et al.* (2002), the first 9 SNPs of the exon 10 (position 11188, 11434,  
250 11577, 11580, 11683, 11687, 11768, 11882, 11936) can be considered as belonging to one haplotype  
251 block, whereas the *loci* 12014 and 12100 constitute a second haplotype block (Figure 2).

252 To these haplotypes it is possible to add the three haplotypes inferred from sequences available  
253 in the GenBank for the same breed (haplotype 4, GenBank XR\_328736, XM\_006071901,  
254 NW\_005785465 and XM\_006071900) and for the buffalo reared in India (haplotypes 5 and 6,  
255 GenBank HQ236497, GQ339914) (Table S2).

256 Therefore, taking into consideration all the 7 amino acid changes (exons 3, 7 and 10) and the  
257 observed allelic combinations, it is possible to deduce at least 6 different possible translations of the  
258 buffalo prolactin receptor and, consequently, the presence at this *locus* of at least 6 alleles (*PRLR* A,  
259 B, C, D, E and F, Figure 3).

260 The SNPs in position 11465 and 11563 were excluded from this analysis because they were  
261 identified only by the comparison of the sequences available in GenBank for Mediterranean  
262 (GenBank XR\_328736, XM\_006071901, NW\_005785465 and XM\_006071900) and Mehsana  
263 breeds (present work) (Table S1).

264 The high level of variability found at the exon 10 of the *PRLR* buffalo gene, which represents  
265 most of the intracellular domain, is probably related primarily to the alternative splicing of the  
266 primary transcript relative to this exon (Trott *et al.*, 2003; Tomás *et al.*, 2006), suggesting that this  
267 particular DNA region (coding) can accumulate a significant amount of polymorphisms, similar to  
268 that observed in other species.

269 The analysis of polymorphisms at the *PRLR locus* in European cattle (*Bos taurus*), African  
270 (*Bos indicus*), yak (*Bos grunniens*), and bison (*Bison bison*) has, in fact, revealed a comparable degree

271 of variability and the existence of several (at least 20) haplotypes (Iso-Touru *et al.*, 2009). A similar  
272 level of genetic diversity was described also at *PRLR* locus in pig (*Sus scrofa*) (Tomás *et al.*, 2006).

273

274 *Fatty acids composition profile, genotyping and association of PRLR polymorphism with milk FA*  
275 *composition traits*

276 Fatty acid composition, fatty acid classes and fat percentage of milk samples have been assessed and  
277 previously reported by Cosenza *et al.* (2017b). The same phenotypes were also used in the present  
278 work to assess possible associations with the genetic diversity found at the *PRLR*. Since the haplotype  
279 analysis evidenced that most of SNP found at the exon 10 belong to one haplotype block, only one  
280 SNP (g.11188A>G) was genotyped and considered in running model.

281 The genotyping of 308 buffalo showed that major allele (adenine) had a relative frequency of  
282 0.9 and the  $\chi^2$  value showed that there was no evidence of departure from the Hardy-Weinberg  
283 equilibrium ( $P \leq 0.05$ ). Genotype distribution of buffalo cows for the SNP g.11188A>G is reported in  
284 Table 2.

285 Results of the model (Supplementary materials) showed a tendency that approached a  
286 significance effects ( $P < 0.06$ ) of the SNP g.11188A>G (AA genotype) at the *PRLR* locus on the milk  
287 FA composition traits with higher contents of odd branched-chain fatty acids (OBCFA), belonging  
288 to polyunsaturated FA class (Table S3). OBCFA acids are an emerging class of bioactive FA that  
289 reduce the incidence of necrotizing enterocolitis in newborns, have anti-tumor effects on lymphomas,  
290 and improve pancreatic  $\beta$ -cell function. However, the interest in OBCFA arises from their  
291 anticarcinogenic effects on cancer cells comparable to that of conjugated linoleic acid (Bainbridge *et*  
292 *al.*, 2016). OBCFA acids are typically of exogenous origin, deriving from forages or supplements  
293 given to the animal and from metabolism of bacteria living in the rumen, respectively. Because  
294 OBCFA are constituents of rumen bacterial membranes (Kaneda, 1991; Mackie *et al.*, 1991), dairy  
295 and meat products from ruminants are a unique source of these FA. This allows for OBCFA to be

296 used as biomarkers of dairy intake in humans and also as biomarkers of rumen function in cattle  
297 (Stefanov *et al.*, 2010).

298 It is possible to hypothesize that the SNP g.11188A>G may be in linkage disequilibrium with  
299 another mutation in *PRLR* gene affecting a greater extent yield traits rather than being causal. For  
300 example, interrogating the software targetscan (Lewis *et al.*, 2005) looking for microRNA targeting  
301 the bovine *PRLR* (ENST00000310101.5) and based on the high homology between *Bos taurus* and  
302 *Bubalus bubalis*, we found that g.12100A>G influences binding of miR-125b (Figure S3) and,  
303 potentially, affect *PRLR* expression.

304 The studies carried out to date in bovine, caprine and porcine species have already detected  
305 different polymorphisms at *PRLR* locus and many of these were significantly associated with  
306 phenotypic traits related to milk production, in particular, milk yield and fat content (Zhang *et al.*  
307 2008; Iso-Touru *et al.* 2009, Lü *et al.*, 2011; Hou *et al.*, 2013, 2014; Skrzypczak *et al.*, 2015).

308 Our results confirm the influence of *PRLR* gene on buffalo milk acidic composition and,  
309 consequently, milk quality, particularly the mutations that fall into the exon 10 of the gene. In fact, it  
310 was suggested that the *PRLR* long form plays a more important role in the regulation of the lactation  
311 (Zi *et al.*, 2012). In particular, the up-regulation or down-regulation of goat *PRLR* long form causes  
312 the change of expression of different genes involved in fatty acids synthesis, and finally controls the  
313 content of lipid droplets in culture of goat mammary epithelial cells. In contrast, none of these roles  
314 have been ascribed to the *PRLR* short form (Shi *et al.*, 2016).

315

316

## 317 **Conclusions**

318 In this study, we provided the first contribution to the characterization of the genomic sequence of  
319 the prolactin receptor encoding gene (*PRLR*) in the Mediterranean river buffalo. Based on the detected  
320 high degree of genetic variability, we defined at least 6 alleles at this locus each codifying for a  
321 different predicted amino acid sequences, four of which specific of the Mediterranean breed (Figure

322 3). All the detected mutations represent the first examples of markers found at the *PRLR* in this breed.  
323 The deletion of a CACTACC heptamer at exon 10 (3'UTR) can be considered as a good molecular  
324 markers to distinguish between the milk of different origins (e.g. from Mediterranean buffalo *vs* all  
325 others mammalian) This finding would be useful for the dairy industry. In fact, the buffalo dairy  
326 sector is one of the most affected by frauds. One of the main goal of the Consortium for the  
327 “Mozzarella di bufala Campana” PDO (Protected Designation of Origin) is to prevent the deception  
328 of consumers with similar cheese products made of less-expensive ingredients (milk of other species,  
329 curd imported from other countries, etc.). Conversely, the authentic buffalo mozzarella must be  
330 prepared only with fresh milk of Italian Mediterranean breed reared in specific geographical areas of  
331 Southern Italy, and following well definite guidelines. Finally, our data revealed that SNPs detected  
332 in *PRLR* gene could influence OBCFA content and, consequently, milk fat composition. Therefore,  
333 the use of these polymorphisms can be a useful tool for the improvement of buffalo milk quality by  
334 marker assisted selection.

335

### 336 **Conflict of interest**

337 The Authors have no conflict of interest

338

### 339 **Acknowledgments**

340 This work was financially supported by the Italian Ministry for Agriculture and Forestry Policy –  
341 MiPAAF (INNOVAGEN project)

342

343

344

### 345 **References**

346

347 Azevedo L., Serrano C., Amorim A. & Cooper D.N. (2015) Trans-species polymorphism in humans  
348 and the great apes is generally maintained by balancing selection that modulates the host immune  
349 response. *Human Genomics* **4**, 9–21.

350

351 Bignon C., Binart N., Ormandy C., Schuler L.A., Kelly P.A. & Djiane J. (1997) Long and short forms  
352 of the ovine prolactin receptor: cDNA cloning and genomic analysis reveal that the two forms arise  
353 by different alternative splicing mechanisms in ruminants and in rodents. *Journal of Molecular*  
354 *Endocrinology* **19**, 109–20.

355

356 Bole-Feysot C., Goffin V., Edery M., Binart N. & Kelly P.A. (1998) Prolactin (PRL) and its receptor:  
357 actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice.  
358 *Endocrine Reviews* **19**, 225–68

359

360

361 Buck K., Vanek M., Groner B. & Ball R.K. (1992) Multiple forms of prolactin receptor messenger  
362 ribonucleic acid are specifically expressed and regulated in murine tissues and the mammary cell line  
363 HC11. *Endocrinology* **130**, 1108–14.

364

365 Chu M.X., Mu Y.L., Fang L., Ye S.C. & Sun S.H. (2007). Prolactin receptor as a candidate gene for  
366 prolificacy of small tail Han sheep. *Animal Biotechnology* **18**, 65–73.

367

368 Clarke D.L., Arey B.J. & Linzer D.I.H. (1993) Prolactin receptor messenger ribonucleic acid  
369 expression in the ovary during rat estrous cycle. *Endocrinology* **133**, 2594–603.

370

371 Clarke D.L. & Linzer D.I.H. (1993) Changes in prolactin receptor expression during pregnancy in  
372 the mouse ovary. *Endocrinology* **133**, 224–32.



373

374

375 Cosenza G., Ianaccone M., Pico B.A., Gallo D., Capparelli R. & Pauciuillo A. (2017a) Molecular  
376 characterization, genetic variability and detection of a functional polymorphism influencing the  
377 promoter activity of OXT gene in goat and sheep. *Journal of Dairy Research* **84(2)**, 165-9.

378

379 Cosenza G., Macciotta N.P.P., Nudda A., Coletta A., Ramunno L. & Pauciuillo A. (2017b) A novel  
380 polymorphism in the Oxytocin receptor encoding gene (OXTR) affects milk fatty acid composition  
381 in Italian Mediterranean river buffalo. *Journal of Dairy Research* **84(2)**, 170-80.

382

383 Gabriel S.B., Schaffner S.F., Nguyen H., Moore J.M., Roy J., Blumenstiel B., Higgins, J., DeFelice  
384 M., Lochner A., Faggart M., Liu-Cordero S.N., Rotimi C., Adeyemo A., Cooper R., Ward R., Lander  
385 E.S., Daly M.J. & Altshuler D. (2002) The structure of haplotype blocks in the human genome.  
386 *Science* **296**, 2225–9.

387

388 Goossens M. & Kan Y.W. (1981) DNA analysis in the diagnosis of hemoglobin disorders. *Methods*  
389 *in Enzymology* **76**, 805–17.

390

391 Hayes H., Le C.C. Goubin G., Mercier D., Payen E.J., Bignon C. & Kohno K. (1996) Localization of  
392 ZNF164, ZNF146, GGTA1, SOX2, PRLR and EEF2 on homeologous cattle, sheep, and goat  
393 chromosomes by fluorescent in situ hybridization and comparison with human gene map.  
394 *Cytogenetics and Cell Genetics* **72**, 342–6.

395

396 Hedrick P.W. (2006) Genetic polymorphism in heterogeneous environments: the age of genomics.  
397 *Annual Review of Ecology, Evolution, and Systematics* **37**, 67–93.

398

399 Hou J.X., An X.P., Song Y.X., Wang J.G., Ma T., Han P., Fang F. & Cao B.Y. (2013) Combined  
400 effects of four SNPs within goat PRLR gene on milk production traits. *Gene* **529**, 276–81.  
401

402 Hou J.X., Fang F., An X.P. Yan Y., Ma T., Han P., Meng F.X., Song Y.X., Wang J.G., Cao B.Y.  
403 (2014) Polymorphisms of PRLR and FOLR1 genes and association with milk production traits in  
404 goats. *Genetics and Molecular Research* **13(2)**, 2555–62.  
405

406 Hou J.X., An X.P., Han P., Peng J.Y. & Cao B.Y. (2015) Two missense mutations in exon 9 of caprine  
407 PRLR gene were associated with litter size. *Animal Genetetics* **46(1)**, 87–90.  
408

409 Iso-Touru T., Kantanen J., Li M.H. Gizejewski, Z. & Vilkki J. (2009) Divergent evolution in the  
410 cytoplasmic domains of PRLR and GHR genes in Artiodactyla. *BMC Evolutionary Biology* **9**, 172.  
411

412 Javed R., Gautam S.K., Vijh R.K. & Tantia M.S. (2011a) Six novel PCR-RFLP loci in milk quality  
413 candidate genes in *Bubalus bubalis*. *International Journal of Livestock Production* **2(6)**, 79–83.  
414

415 Javed R., Gautam S.K., Vijh R.K. & Tantia M.S. (2011b) Characterization of PRLR and PPARGC1A  
416 genes in buffalo (*Bubalus bubalis*). *Genetics and Molecular Biology* **34**, 592–4.  
417

418 Kaneda T. (1991) Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic  
419 significance. *Microbiological Reviews* **55**, 288–302.  
420

421 Klein J. (1987) Origin of major histocompatibility complex polymorphism: the trans-species  
422 hypothesis. *Human Immunology* **19**, 155–62.  
423

424 Lenstra J.A., Van Boxtel J.A.F., Zwaagstra K.A. & Schwerin M. (1993) Short interspersed nuclear  
425 elements (SINE) sequences of the Bovidae. *Animal Genetics* **24**, 33–9.  
426

427 Lewis B.P., Burge C.B. & Bartel D.P. (2005) Conserved seed pairing, often flanked by adenosines,  
428 indicates that thousands of human genes are microRNA targets. *Cell* **120(1)**, 15–20.  
429

430 Lü A., Hu X., Chen H., Dong Y., Zhang Y. & Wang X. 2011. Novel SNPs of the Bovine PRLR Gene  
431 Associated with Milk Production Traits. *Biochemical Genetics* **49**, 177–89.  
432  
433

434 Mackie R.I., White B.A. & Bryant M.P. (1991) Lipid metabolism in anaerobic ecosystems. *Critical*  
435 *Reviews in Microbiology* **17**, 449–79.  
436

437 Pauciullo A., Shuiep E., Cosenza G., Ramunno L. & Erhardt G. (2013) Molecular characterization  
438 and genetic variability at  $\kappa$ -casein gene 1 (CSN3) in camels. *Gene* **513**, 22–30.  
439

440 Pauciullo A., Giambra I.J., Iannuzzi L. & Erhardt G. (2014) The  $\beta$ -casein in camels: molecular  
441 characterization of the CSN2 gene, promoter analysis and genetic variability. *Gene* **547**, 159–68.  
442

443 Ramunno L., Cosenza G., Rando A, Illario R., Gallo D., Di Bernardino D. & Masina P. (2004) The  
444 goat  $\alpha$ s1-casein gene: gene structure and promoter analysis. *Gene* **334**, 105–11.  
445

446 Shi D.S., Wang J., Yang Y., Lu F.H., Li X.P. & Liu Q.Y. (2012) DGAT1, GH, GHR, PRL and PRLR  
447 polymorphism in water buffalo (*Bubalus bubalis*). *Reproduction in Domestic Animals* **47**, 328–34.  
448

449 Shi H., Zhang T., Yi, Y., Wang H. & Luo J. (2016) Long form PRLR (lPRLR) regulates genes  
450 involved in the triacylglycerolsynthesis in goat mammary gland epithelial cells. *Small Ruminant*  
451 *Research* **139**, 7–14.

452

453 Skrzypczak E., Babicz M. & Pastwa M. (2015) Effect of Prolactin Receptor (PRLR) and Beta-Casein  
454 (CSN2). Gene Polymorphism on the Chemical Composition of Milk Sows. *Folia Biologica (Kraków)*  
455 **63(2)**, 135–44.

456

457 Stefanov I., Baeten V., Abbas O., Colman E., Vlaeminck B., De Baets B. & Fievez V. (2010) Analysis  
458 of milk odd- and branched-chain fatty acids using Fourier transform (FT)-Raman spectroscopy.  
459 *Journal of Agricultural and Food Chemistry* **58**, 10804–11.

460

461 Tomás A., Casellas J., Ramírez O., Muñoz G., Noguera J.L. & Sánchez A. (2006) High amino acid  
462 variation in the intracellular domain of the pig prolactin receptor (PRLR) and its relation to ovulation  
463 rate and piglet survival traits. *Journal of Animal Science* **84(8)**, 1991–8.

464

465 Trott J.F., Hovey R.C., Koduri S. & Vonderhaar B.K. (2003) Alternative splicing to exon 11 of human  
466 prolactin receptor gene results in multiple isoforms including a secreted prolactin-binding protein.  
467 *Journal of Molecular Endocrinology* **30**, 31–47.

468

469 Wang L.P., Geng R.Q., Zhang X.N. & Sun W. (2015) Identification of SNPs within the PRLR gene  
470 and effects on maternal behavior in sheep. *Genetics and Molecular Research* **14(4)**, 17536–43.

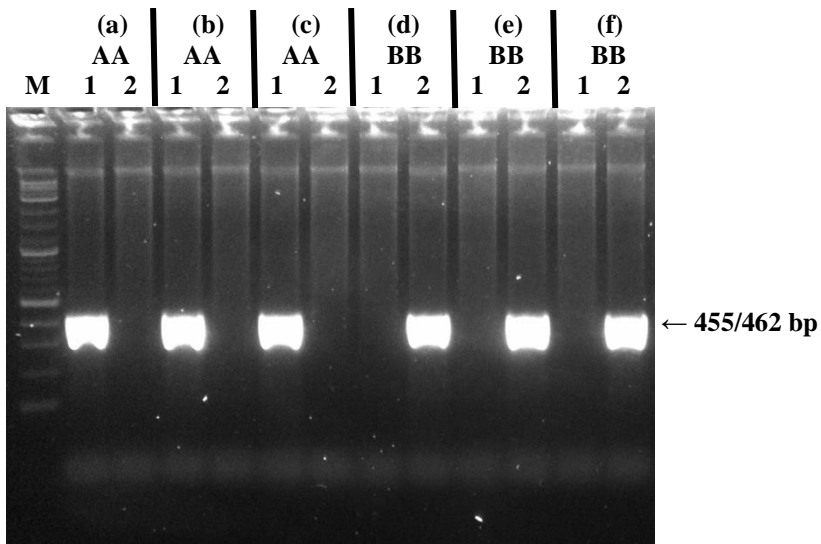
471

472 Zhang J.L., Zan L.S., Fang P., Zhang F., Shen G. & Tian W.Q. (2008) Genetic variation of *PRLR*  
473 gene and association with milk performance traits in dairy cattle. *Canadian Journal of Animal Science*  
474 **88**, 33–9.

475

476 Zi X.D., Chen D.W. & Wang H.M. (2012) Molecular characterization, mRNA expression of prolactin  
477 receptor (PRLR) gene during pregnancy, nonpregnancy in the yak (*Bos grunniens*). *General and*  
478 *Comparative Endocrinology* **175**, 384–8.

479



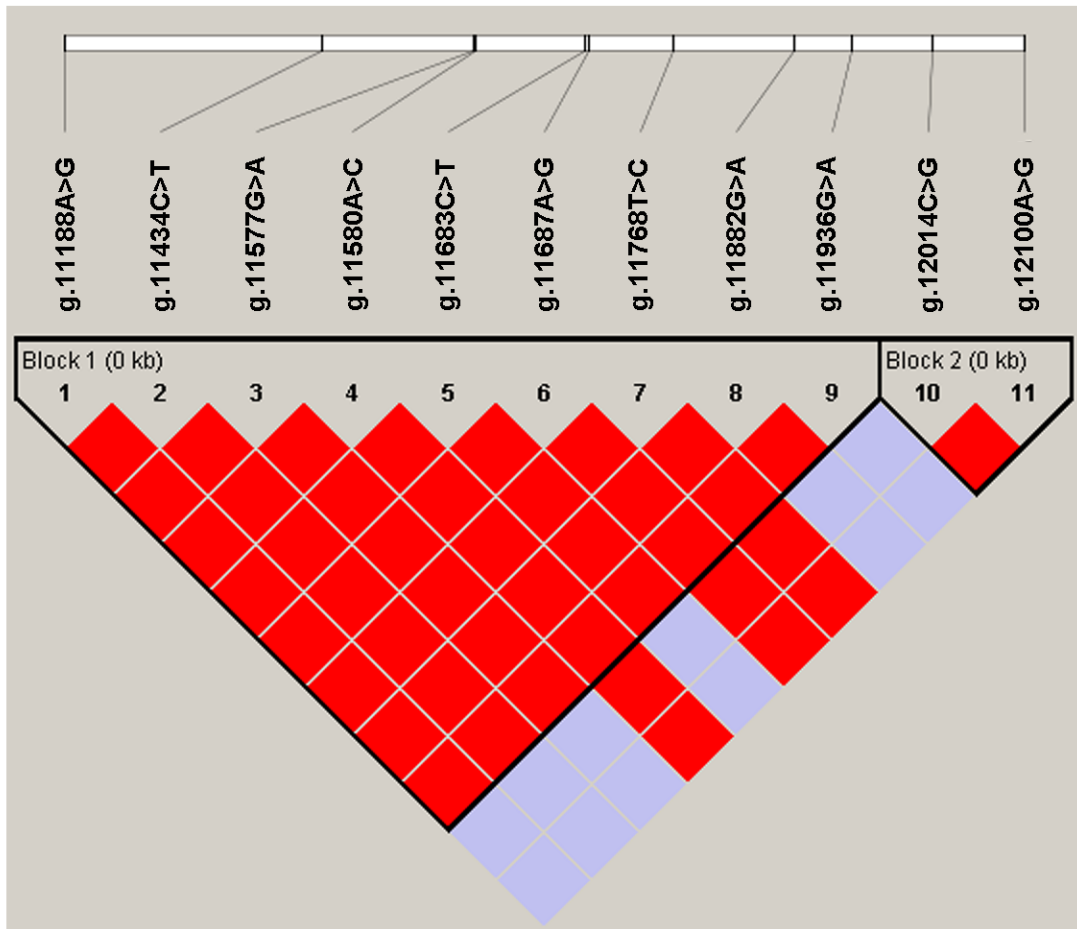
480  
481  
482  
483  
484  
485  
486  
487  
488  
489

**Fig. 1** Identification by AS-PCR of carriers of the heptamer CACTACC at the exon 10 of the *PRLR* gene in Mediterranean (a), Murrah (b) and Mehsana (c) river buffaloes, goat (d), sheep (e) and cattle (f). AA, homozygous for the “non CACTACC insertion”; BB, homozygous for the “CACTACC insertion”; 1, “non CACTACC insertion” primer; 2, “CACTACC insertion” primer, M, Marker (2-Long DNA Ladder (0.1-10 kb) (Biolabs).

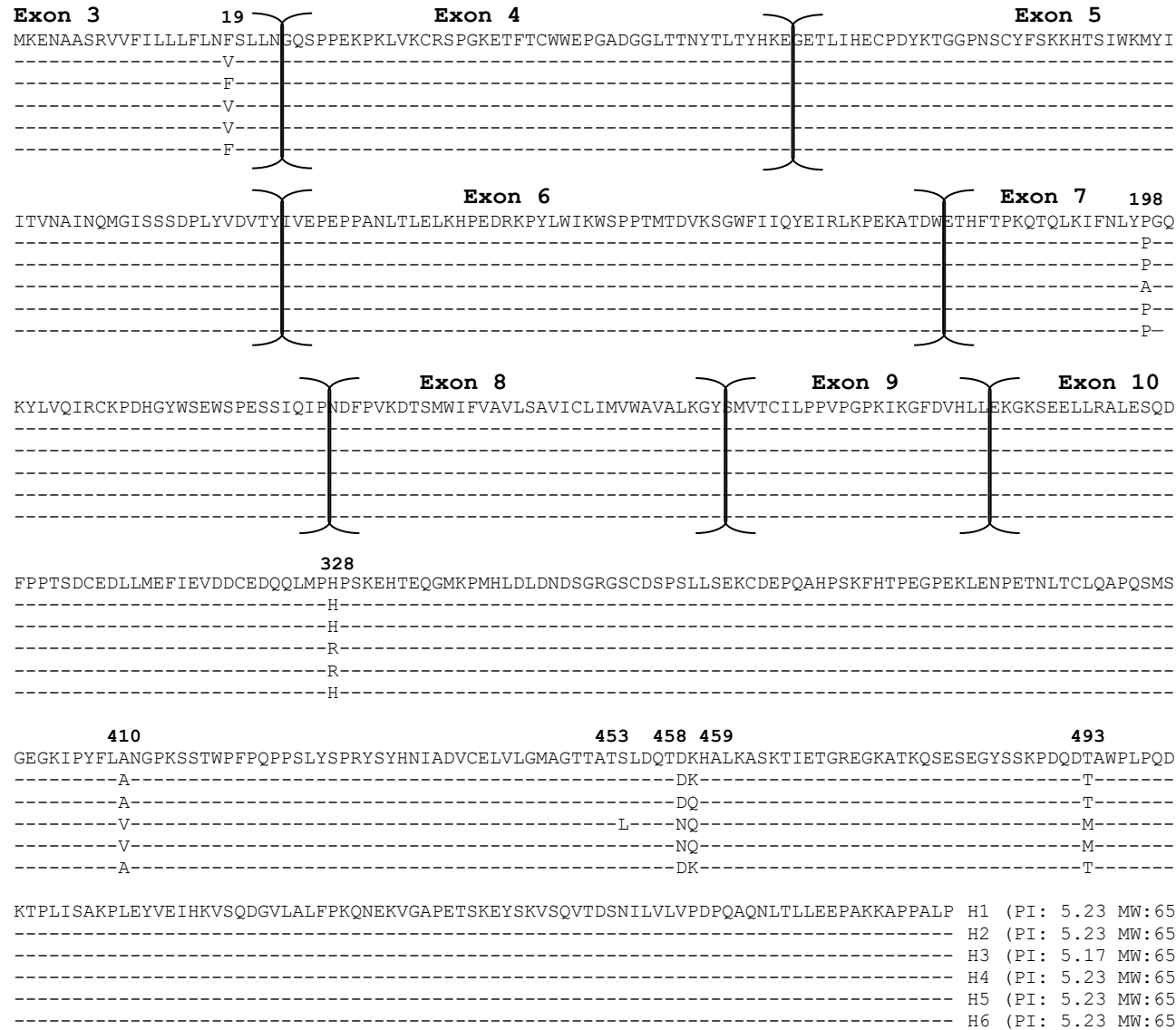
490 **Fig. 2.** Linkage disequilibrium analysis, observed haplotypes (H1, H2 and H3) and their relative  
 491 frequencies at the Mediterranean river buffalo *PRLR* locus (exon 10).

492

	Block 1									Block 2		
H1: 0.614	A	C	G	A	C	A	C	G	G	G	G	0.754
H2: 0.245	A	C	G	C	C	G	T	G	G	C	A	0.237
H3: 0.132	G	T	A	C	T	G	T	A	A			



493  
 494



**Fig. 3.** Amino acid comparison (deduced mRNA) of different haplotypes (H1 to H6) observed at the buffalo *PRLR* locus. Dashes represent nt identical to those in upper lines. PI: Isoelectric Point; MW: Molecular Weight.



<b>Description</b>	<b>Mutation</b>	<b>aa change</b>
Exon 3	g.114T>G	p.Phe19Val
Intron 3	g.216T>C g.689A>G g.947A>G g.1046C>A g.1119_1121insA g.1999G>A g.2100A>G g.2550T>G	
Intron 4	g.3266C>T g.3720G>A g.3777C>T g.3835G>A	
Intron 6	g.7011G>T g.7333C>T	
Exon 7	g.7401C>G	p.Pro198Ala
Intron 7	g.7953T>A g.8948_8951delG	
Intron 8	g.9504G>T	
Intron 9	g.10240G>A g.10254C>T g.10276C>T g.10352A>T g.10494C>T g.10558_10561delAGTC g.11007C>T	
Exon 10	g.11188A>G g.11434C>T g.11577G>A g.11580A>C g.11683C>T g.11687A>G g.11768T>C g.11882G>A g.11936G>A	p.His328Arg p.Ala410Val p.Asp458Asn p.Lys459Gln p.Thr493Met p.Ala494 p.Asp521 p.Pro559 p.Pro577
Exon 10 3'UT	g.12014C>G g.12100A>G	

**Tab. 1.** Polymorphisms detected at Mediterranean river buffalo *PRLR* locus

**Tab. 2.** Genotyping data, allele frequency, relative frequencies of buffalo cows of the SNP g.11188A>G at the exon 10 of the *PRLR* gene in Mediterranean river buffalo population.

Position	Genotype distribution				Allele freq		Relative frequencies used in the model			
		AA	AG	GG	TOT	A	G	Genotype	n. of cows	n. of farms
g.11188A>G	Obs.	253	49	6	308	0.9	0.1	AA	82.14	14
	Exp.	250.02	54.96	3.02				AG	15.91	
	1 degree of freedom P value <0.05; $\chi^2= 3.84$							GG	1.95	
								Total	100	