Mol Biol Rep
DOI 10.1007/s11033-011-0794-2

Study of polymorphisms in the promoter region of ovine *β-lactoglobulin* gene and phylogenetic analysis among the Valle del Belice breed and other sheep breeds considered as ancestors

S. Mastrangelo · M. T. Sardina · V. Riggio · B. Portolano

Received: 29 October 2010 / Accepted: 29 April 2011

Abstract The aim of this work was to sequence the promoter region of *β-lactoglobulin* (*BLG*) gene in four sheep breeds, in order to identify polymorphisms, infer and analyze haplotypes, and phylogenetic relationship among the Valle del Belice breed and the other three breeds considered as ancestors. Sequencing analysis and alignment of the obtained sequences showed the presence of 36 single nucleotide polymorphisms (SNPs) and one deletion. A total of 22 haplotypes found in "best" reconstruction were inferred considering the 37 polymorphic sites identified. Haplotypes were used for the reconstruction of a phylogenetic tree using the Neighbor-Joining algorithm. The number of polymorphisms identified showed high variability within breeds. Analysis of genetic diversity indexes showed that the Sarda breed presented the lowest nucleotide diversity, whereas the Comisana breed presented the highest one. Comparing the nucleotide diversity among breeds, the highest value was obtained between Valle del Belice and Pinzirita breeds, whereas the lowest one was between Valle del Belice and Sarda breeds. Considering that polymorphisms in the promoter region of *BLG* gene could have a functional role associated with milk composition, the lowest value of nucleotide diversity between Valle del Belice and Sarda breeds may be related to a higher similarity of milk composition of these two breeds compared to the others. Further analyses will be conducted in order to evaluate the possible correlation between the genetic diversity indexes and the BLG content in milk of our breeds.

Keywords *β*-Lactoglobulin · Polymorphisms · Sheep breeds · Phylogenetic analysis

S. Mastrangelo (✉) · M. T. Sardina · V. Riggio · B. Portolano
Dipartimento DEMETRA—Sezione Produzioni Animali, Università degli Studi di Palermo, Viale delle Scienze—Parco d'Orleans, 90128 Palermo, Italy
e-mail: salvatore.mastrangelo@senfimizo.unipa.it

S. Mastrangelo · M. T. Sardina
Consorzio Regionale di Ricerca Bioevoluzione Sicilia, c/o Dipartimento DEMETRA—Sezione di Produzioni Animali, Università degli Studi di Palermo, Viale delle Scienze—Parco d'Orleans, 90128 Palermo, Italy

Introduction

β-Lactoglobulin (BLG) is synthesized by secreting cells of mammary gland and it is the major whey protein in the milk of ruminants. It is also found in the milk of different mammalian species including cats [1], dogs and dolphins [2] but it is lacking in humans [3, 4], rodents, and lagomorphs [5]. It is a globular protein, belonging to the family of lipocalins, small proteins with some peculiarities, such as the ability to bind hydrophobic molecules [6]. Although no clear physiological functions have been defined for BLG, a role in the transport of retinol and fatty acids has been suggested [6, 7]. However, the general affinity of BLG with these hydrophobic molecules did not allow ascribing a specific role [7, 8].

The BLG encoding gene has been sequenced in sheep [9], cattle [10], and goats [11], and mapped on chromosome 3 in sheep and chromosome 11 in goats and cattle [12].

A large number of variants have been reported for bovine and ovine BLG protein. Three co-dominant alleles (A, B, and C) have been reported in sheep that differ by one or more amino acid changes. BLG variant A differs from BLG variant B in the amino acid sequence at position 20 ($Tyr_A \rightarrow His_B$) [13, 14], whereas it differs from BLG

Published online: 13 May 2011

variant C at position 148 ($Arg_A \rightarrow Gln_C$) [15]. Variants A and B are the most common and have been reported in several breeds [13, 14], whereas variant C has been reported only in milk from Merinoland, Hungarian Merino, Pleven [16], and Carranzana and Lacha [17] breeds. Many studies on the effect of ovine *BLG* polymorphisms on milk production traits have been carried out, but results are still conflicting. Some authors reported the positive effect of variant B on milk production and quality and whey protein content [18–22], whereas others reported the positive effect of variant A on fat and protein content and enzymatic properties [23–25]. However, other studies reported no direct association between genotypes at this locus and milk characteristics [26–29].

Several potential binding sites for specific mammary gland transcription factors (TFs) have been found by Watson et al. [30] within the ovine *BLG* promoter region. Since they have been identified in the 5′-flanking region of many expressed milk protein genes in different species [30–32], it has been suggested that these factors are important regulators of milk protein gene expression. Therefore, the presence of polymorphisms in the *BLG* promoter region could influence the binding affinity of TFs and affect both the expression level and the content of BLG in milk [33].

In Sicily, dairy sheep production represents an important resource for the economy of hill and mountain areas, in which other economic activities are difficult to develop [34]. The main breeds reared are Valle del Belice, Comisana, Pinzirita, and Sarda, which are genetically connected among them. Based on historical, geographical, and morphological information, it seems indeed that the Valle del Belice breed derives from the Pinzirita, to which is similar for the horned trait in males, crossed with the Comisana, to which is similar for the coat color (i.e. white with red head) and for the high milk production. Subsequently, the cross between these two breeds was likely crossed with the Sarda breed [35]. Nowadays, the Valle del Belice breed is the most appreciated dairy sheep breed reared on the island. The aim of this work was to sequence the full-length promoter region of *BLG* gene in four sheep breeds reared in Sicily, in order to: (i) identify polymorphisms; (ii) infer and analyze haplotypes; and (iii) analyze phylogenetic relationship among the Valle del Belice breed and the other three breeds.

Materials and methods

Amplification of sheep *BLG* promoter region

A total of 50 randomly chosen unrelated (i.e. without common parents) animals from several farms located in different areas of Sicily and belonging to the four breeds (Valle del Belice n = 20; Comisana n = 10; Pinzirita n = 10; and Sarda n = 10) were analyzed. Genomic DNA was extracted from blood buffy coats of nucleated cells using a salting out method [36]. Primers BLG-F1 and BLG-R1 (Table 1) were used to amplify a fragment of 2,255 bp, containing 2,138 bp of the promoter region and 117 bp of exon 1 of *BLG* gene, as reported by Sardina et al. [37] in goat gene (GenBank accession number Z33881). PCR reaction was performed in a final volume of 25 µl using 2× PCR Master Mix (Fermentas), 10 µM of each primer, and approximately 75 ng of genomic DNA. The thermal cycling conditions were: 95°C for 5 min, 30 cycles of 95°C for 30 s, 59°C for 1 min and 72°C for 1 min 30 s, and a final extension of 72°C for 5 min. The PCR products were checked by electrophoresis on 1% agarose gel stained with ethidium bromide.

Table 1 Primers used to amplify and sequence the promoter region of sheep *BLG* gene, as reported by Sardina et al. [37] in goat gene (GenBank accession number Z33881)

	Primer sequences
Forward	
BLG-F1	5′-AGG CCA GAG GTG CTT TAT TTC CGT-3′
BLG-F2	5′-TAG TCT CTG CCT CCG TGT TCA CAT-3′
BLG-F3	5′-AAC CTC CAA CCA AGA TGC TGA CCA-3′
BLG-F4	5′-AGG GTC AGG TCA CTT TCC CGT-3′
BLG-F5	5′-AGA AGG CCT CCT ATT GTC CTC GTA GA-3′
Reverse	
BLG-R1	5′-TCC ATG GTC TGG GTG ACG ATG ATG-3′
BLG-R2	5′-TTC CCG GAA TCC TAC TTG GCT CAT-3′
BLG-R3	5′-ACC AGC TCC TCC AAA CCA TGT GA-3′
BLG-R4	5′-AGT GAC TAA ACC ACT CAT CAC AGG G-3′
BLG-R5	5′-CAA CAA GGA ACT TCA GGT TGG AAT-3′

DNA sequencing reaction

PCR products were purified using 10 U of Exonuclease I and 1 U of Shrimp Alkaline Phosphatase (Fermentas). The resulting PCR products did not need additional purification before sequencing. Primers BLG-F1 and BLG-R1 and other eight internal primers (Table 1) were used for sequencing reaction with BigDye Terminator v3.1 Cycle Sequencing Kit in an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems).

Statistical analysis

SeqScape v3.1 software (Applied Biosystems) was used to analyze the nucleotide sequences, whereas Clustal W

Table 2 Polymorphic sites identified in the *BLG* promoter region (GenBank accession number X68105) of sheep breeds

Promoter	-1981	-1935	-1913	-1911	-1909	-1815	-1791	-1780	-1770	-1733	-1631	-1448	-1437	-1245	-1230	-983	-966	-941	-919	-903	-764	-722	-696	-654	-575	-545	-528	-496	-477	-447	-438	-246	-163	-134	-117	-46	-42
X68105	**A**	**T**	**C**	**C**	**T**	**G**	**T**	**T**	**C**	**A**	**G**	**G**	**G**	**C**	**A**	**G**	**C**	**A**	**A**	**T**	**G**	**G**	**T**	**G**	**G**	**C**	**A**	**C**	**T**	**A**	**C**	**A**	**T**	**G**	**G**	**T**	**T**
VdB-1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	Y	R	R	Y	-	-	-	-	Y	W	Y	K	R	K	Y
VdB-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	Y	R	R	Y	-	-	-	-	Y	W	Y	K	R	K	Y
VdB-3	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	-	-	R	R	-	R	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
VdB-4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-10	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	R	-	-	R	-	-	-	-	-	R	S	Y	M	-	W	Y	K	R	K	Y
VdB-11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-12	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	-	-	-	R	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
VdB-13	G	C	T	T	G	T	C	G	T	G	A	C	A	T	G	A	del	G	G	C	A	-	-	-	-	-	G	-	C	C	-	T	C	T	A	G	C
VdB-14	G	C	T	T	G	T	C	G	T	G	A	C	A	T	G	A	del	G	-	-	-	-	-	-	-	-	G	-	C	C	-	T	C	T	A	G	C
VdB-15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VdB-20	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	-	-	-	R	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
COM-1	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	-	-	-	-	-	R	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
COM-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	Y	R	R	Y	-	-	-	-	Y	W	Y	K	R	K	Y
COM-3	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	-	-	R	-	Y	R	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
COM-4	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	-	-	-	-	-	R	-	-	-	-	-	R	S	Y	M	-	W	Y	K	R	K	Y
COM-5	G	C	T	T	G	T	C	G	T	G	A	C	A	T	G	A	del	G	G	C	A	-	-	-	-	-	G	-	C	C	-	T	C	T	A	G	C
COM-6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
COM-7	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	-	-	-	R	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
COM-8	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	-	-	-	R	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
COM-9	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	-	-	R	-	Y	R	-	-	-	-	-	R	S	Y	M	-	W	Y	K	R	K	Y
COM-10	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	-	-	-	R	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
PIN-1	G	C	T	T	G	T	C	G	T	G	A	C	A	T	G	A	del	G	G	C	A	-	-	-	-	-	G	-	C	C	-	T	C	T	A	G	C
PIN-2	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	R	-	Y	R	-	-	-	-	-	R	S	Y	M	-	W	Y	K	R	K	Y
PIN-3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
PIN-4	R	Y	Y	Y	-	-	Y	K	Y	-	-	S	R	Y	-	R	-	-	-	-	-	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
PIN-5	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	-	-	-	-	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
PIN-6	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	-	-	-	-	-	-	R	R	Y	R	-	Y	M	Y	T	C	T	A	G	C
PIN-7	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	-	-	-	R	-	-	R	R	Y	R	-	Y	M	Y	T	C	T	A	G	C
PIN-8	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	-	-	-	R	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y
PIN-9	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	-	-	-	-	-	-	-	-	-	-	R	-	Y	M	Y	T	C	T	A	G	C
PIN-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAR-1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAR-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAR-3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAR-4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAR-5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAR-6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAR-7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAR-8	G	C	T	T	G	T	C	G	T	G	A	C	A	T	G	A	-	G	-	C	A	-	-	-	-	-	G	G	C	C	-	T	C	T	A	G	C
SAR-9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SAR-10	R	Y	Y	Y	K	K	Y	K	Y	R	R	S	R	Y	R	R	del	R	-	-	R	-	-	-	-	-	R	-	Y	M	-	W	Y	K	R	K	Y

VdB Valle del Belice, *COM* Comisana, *PIN* Pinzirita, *SAR* Sarda

software [38] was used to align the sequences (GenBank accession number FR821261–FR821310). TESS software [39] was used to predict TFs binding sites, using information collected in TRANSFAC database [40]. Genetic diversity indexes, such as number of polymorphic sites, nucleotide diversity (π), average number of nucleotide differences (k), number of haplotypes (h), and haplotype diversity (H_d) within and among breeds were estimated with the DnaSP v5.10.01 software [41]. PHASE v2.1.1 software [42, 43] (with -MR0 -d1 options), included in DNAsp v5.10.01 software package [41], was used to infer haplotypes within the whole analyzed sample. Finally, MEGA v4.0 software [44] was used for the reconstruction of a phylogenetic tree using the Neighbor-Joining (NJ) algorithm with nucleotide substitution model and 1,000 bootstrap replications.

Results and discussion

Identification of polymorphisms and genetic diversity indexes

Sequencing analysis and alignment of the obtained sequences showed the presence of 37 polymorphic sites in the *BLG* promoter region: 36 single nucleotide polymorphisms (SNPs) and one deletion (Table 2), which equates to about one polymorphism per approximately 60 bp. The number of polymorphisms identified in our breeds showed high variability of the *BLG* promoter region as reported by Sardina et al. [37] in goat and by Ganai et al. [45] in cattle. Valle del Belice and Comisana breeds had all point mutations (36 SNPs and the deletion) in common, whereas the Pinzirita breed presented 34 SNPs and the deletion, and

the Sarda breed 29 SNPs and the deletion. Using the TRANSFAC database [40], we found four binding sites for milk protein binding factor (MPBF) and five binding sites for nuclear factor-I (NF-I) within the promoter region of sheep *BLG* gene, as reported by Watson et al. [30]. Since at least five NF-I have been identified, these authors suggested that these factors could play a regulatory role in *BLG* transcription. The polymorphic site −246 A/T, we found in our breeds, lies within a region of sheep *BLG* promoter, in which a NF-I binding site is involved (−253/−240) (TESS—TRANSFAC Site Record R03872), causing the loss of the latter.

Analysis of genetic diversity indexes (Table 3) showed that the Sarda breed presented the lowest nucleotide diversity, which resulted in a reduced number of haplotypes with a consequent low haplotype diversity. The low nucleotide and haplotype diversity within the Sarda breed was due to the presence of low proportion of animals showing polymorphisms (Table 2). Although Comisana and Valle del Belice breeds presented the same number of SNPs, the former was characterized by the highest variability, presenting the highest nucleotide diversity, which resulted in the highest number of haplotypes with a consequent high haplotype diversity (Table 3), whereas the Valle del Belice breed showed lower nucleotide diversity, lower number of haplotypes, and consequently lower haplotype diversity. These results may be explained considering the higher proportion of polymorphic animals in the Comisana breed (90%) compared to the Valle del Belice breed (40%) (Table 2). Moreover, the lower number of haplotypes in the Valle del Belice breed compared to the Comisana breed can be explained by the fact that some positions that were in homozygous condition in the Valle del Belice breed were in heterozygous condition in the Comisana breed. The Pinzirita breed, which presented a lower number of SNPs compared to the Valle del Belice breed, presented higher nucleotide diversity, higher number of haplotypes, and higher haplotype diversity (Table 3), probably due to a higher proportion of polymorphic individuals in the Pinzirita breed (80%) compared to the Valle del Belice breed. Moreover, the heterozygous condition for some positions within the Pinzirita breed, compared to the Valle del Belice breed, led to a higher number of haplotypes.

It is interesting to highlight that among the four breeds, those characterized by lower proportion of polymorphic individuals were Valle del Belice and Sarda breeds. It is possible to hypothesize that this is influenced by the selection pressure. However, in the Sicilian farming system, natural mating is the common practice and the exchange of rams among flocks is quite unusual. This leads to an increase of inbreeding within the population and a consequent decrease of variability (heterozygous condition).

Table 4 shows the nucleotide diversity and the average number of nucleotide differences estimated between the Valle del Belice breed and the other three breeds. The highest value of nucleotide diversity was obtained between Valle del Belice and Pinzirita breeds, due to a higher presence of mutated sites in homozygous condition in the Pinzirita breed than in the Valle del Belice breed (Table 2). The lowest value of nucleotide diversity between breeds was obtained between Valle del Belice and Sarda breeds, due to a lower presence of mutated sites in homozygous condition in the Sarda breed (Table 2). These results were confirmed by the average number of nucleotide differences between breeds (Table 4). Valle del Belice and Pinzirita breeds presented the highest average number of nucleotide difference, whereas the lowest value was found in Valle del Belice and Sarda breeds. A previous study conducted on the genetic structure and relationship between the Valle del Belice breed and the other sheep breeds considered as ancestors, using the genetic polymorphisms of seven protein systems, has reported the lowest genetic distance between Valle del Belice and Pinzirita breeds and the highest one between Valle del Belice and Sarda breeds [46], which is not in agreement with our results. Considering that polymorphisms in the *BLG* promoter region gene could have a functional role associated with milk composition, the lowest value of nucleotide diversity between

Table 4 Nucleotide diversity (π) and average number of nucleotide differences (k) between Valle del Belice breed and the other three breeds

Breed	π	k
Valle del Belice–Comisana	0.00564	12.104
Valle del Belice–Pinzirita	0.00566	12.138
Valle del Belice–Sarda	0.00421	9.027

Table 3 Genetic diversity indexes in the four sheep breeds

Breed	Polymorphic sites	$\pi \pm$ SD	h	$H_d \pm$ SD
Valle del Belice	36	0.00459 ± 0.00097	8	0.438 ± 0.098
Comisana	36	0.00703 ± 0.00055	10	0.837 ± 0.076
Pinzirita	34	0.00695 ± 0.00047	9	0.826 ± 0.073
Sarda	29	0.00355 ± 0.00151	3	0.279 ± 0.123

π Nucleotide diversity, *h* number of haplotypes, H_d haplotype diversity, *SD* standard deviation

Table 5 Haplotypes identified in the four sheep breeds, frequencies (Freq.) and standard error (SE)

	Haplotypes	Freq.	SE
H1[a, b, c, d]	ATCCTGTTCAGGGCAGCAATGGTGGCACTACATGGTT	0.617	0.008
H2[c]	ATCCTGTTCAGGGCAGCAATGGTGGCACTATTCTAGC	0.007	0.004
H3[a, b]	ATCCTGTTCAGGGCAGCAATGACAATACTATTCTAGC	0.030	0.000
H4[c]	ATCCTGTTCAGGGCAGCAATAGTAATACTATTCTAGC	0.020	0.000
H5[b]	ATCCTGTTCAGGGCAGCGATGGTGGCACTACATGGTT	0.017	0.006
H6[d]	ATCCTGTTCAGGGCAG5AATGGTGGCACTACATGGTT	0.010	0.000
H7[c]	GCTTTGCGTAGCATAACAATAGTGGCGCCCCTCTAGC	0.010	0.000
H8[b]	GCTTGTCGTGACATGGCAATAGTGGCGCCCCTCTAGC	0.012	0.003
H9[b]	GCTTGTCGTGACATGGCAATAGTGGCGGCCCTCTAGC	0.014	0.007
H10[a, b]	GCTTGTCGTGACATGGCAGTAGTGGCGCCCCTCTAGC	0.020	0.000
H11[a]	GCTTGTCGTGACATGGCGACAGTGGCGGCCCTCTAGC	0.009	0.003
H12[c]	GCTTGTCGTGACATGACAATGGTGGCGCCCCTCTAGC	0.030	0.000
H13[c]	GCTTGTCGTGACATGACAATAGTGGCACTCCTTTAGC	0.007	0.004
H14[c]	GCTTGTCGTGACATGACAATAGTGGCGCTCCTCTAGC	0.009	0.002
H15[b]	GCTTGTCGTGACATGACAATAGTGGCGCCACTCTAGC	0.010	0.000
H16[a, b]	GCTTGTCGTGACATGACAATAGTGGCGCCCCTCTAGC	0.020	0.002
H17[a, b]	GCTTGTCGTGACATGACGATAGTGGCGCCCCTCTAGC	0.019	0.002
H18[c]	GCTTGTCGTGACATGACGATAGTGGCGGCCCTCTGGC	0.007	0.004
H19[d]	GCTTGTCGTGACATGACGACAGTGGCGGCCCTCTAGC	0.020	0.000
H20[a]	GCTTGTCGTGACATGA5GATGGTGGCACCCCTCTAGC	0.020	0.000
H21[d]	GCTTGTCGTGACATGA5GATAGTGGCGCCCCTCTAGC	0.010	0.000
H22[a, b, c]	GCTTGTCGTGACATGA5GGCAGTGGCGCCCCTCTAGC	0.060	0.000

[a] Haplotypes identified in Valle del Belice breed

[b] Haplotypes identified in Comisana breed

[c] Haplotypes identified in Pinzirita breed

[d] Haplotypes identified in Sarda breed

Valle del Belice and Sarda breeds may be related to a higher similarity of milk composition of these two breeds compared to the others.

Identification of haplotypes and phylogenetic analysis

On a total of 36 possible haplotypes, 22 haplotypes in "best" reconstruction were inferred considering the 37 polymorphic sites identified (Table 5). Of the 22 haplotypes, seven were specific for the Pinzirita breed, four for the Comisana breed, three for the Sarda breed, and two for the Valle del Belice breed (Table 5). Among the analyzed breeds only Valle del Belice and Comisana breeds shared four haplotypes. Haplotype H1 showed the highest frequency (0.617) and was found in all breeds, followed by haplotype H22 with a frequency of 0.060. In particular, haplotype H22 was the only one shared among Valle del Belice, Pinzirita, and Comisana breeds and it was specific of animals presenting the deletion at position −966.

Haplotypes were used for the reconstruction of a phylogenetic tree, using *BLG* promoter region of *Capra hircus* (GenBank accession number Z33881), *Bos taurus* (GenBank accession number Z48305), *Bos grunniens* (GenBank accession number AF194981), and *Bubalus bubalis* (GenBank accession number AM238696) as outliers. The NJ tree (Fig. 1) showed the presence of some haplotypes closely related to the consensus *BLG* promoter region of *Ovis aries* and in particular haplotypes H1 and H6, identical to the former except for the deletion at position −966. On the same branch were haplotypes H2, H3, H4, and H5 that showed polymorphisms in the proximal promoter region and in particular in the region between position −764 and position −42. The other haplotypes (H7–H22) were placed in a different branch and among them, haplotype H22 was the closest to the outlier sequences branch due to presence of all polymorphic sites in mutated homozygous condition compared to *O. aries* consensus.

Conclusion

Results showed high genetic variability in the *BLG* promoter region within our breeds. The presence of the

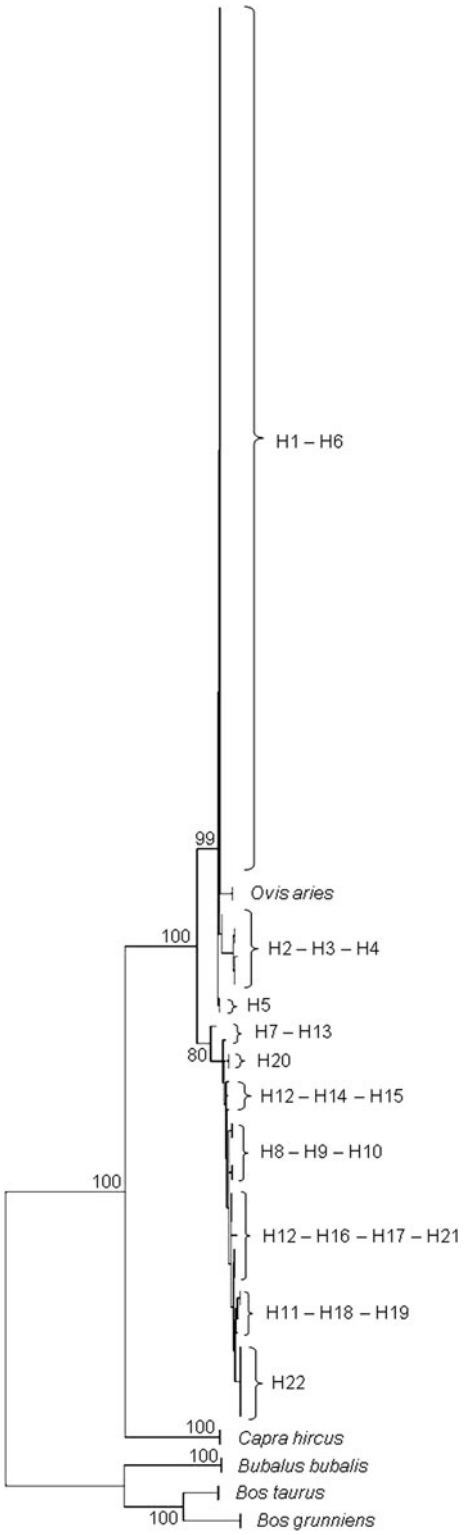

Fig. 1 Phylogenetic tree obtained using the Neighbor-Joining algorithm with nucleotide substitution model and 1,000 bootstrap replications

polymorphic site −246 A/T could influence the binding affinity of NF-I in the region −253/−240 of the *BLG* promoter. Analysis of genetic diversity of the *BLG* promoter region revealed the highest value of genetic diversity between Valle del Belice and Pinzirita breeds and the lowest one between Valle del Belice and Sarda breeds. The lowest value of genetic diversity between Valle del Belice and Sarda breeds may be related to a higher similarity of milk composition of these two breeds compared to the others. However, at present literature does not present any evidence about that. Further analyses will be conducted on a wider sample in order to estimate the possible effect that the loss of TF could have on *BLG* gene expression level and to evaluate the possible correlation between the genetic diversity indexes and the BLG content in milk of our breeds.

References

1. Halliday JA, Bell K, Shaw DC (1991) The complete amino acid sequence of feline β-lactoglobulin II and partial revision of the equine β-lactoglobulin II sequence. Biochim Biophys Acta Protein Struct Mol Enzymol 1077:25–30
2. Pervaiz S, Brew K (1986) Purification and characterization of the major whey proteins from the milks of the bottlenose dolphin (*Tursiops truncatus*), the Florida manatee (*Trichechus manatus latirostris*), and the beagle (*Canis familiaris*). Arch Biochem Biophys 246:846–854
3. Brignon G, Chtourou A, Ribadeau-Dumas S (1985) Does beta-lactoglobulin occur in human milk? J Dairy Res 52:249–254
4. Monti JC, Mermoud AF, Jollès P (1989) Anti-bovine beta-lactoglobulin antibodies react with a human lactoferrin fragment and bovine beta-lactoglobulin present in human milk. Experientia 45:178–180
5. Hambling SG, McAlpine A, Sawyer L (1992) β-lactoglobulin. In: Fox PF (ed) Advanced dairy chemistry—1. Proteins. Elsevier Applied Science, London, pp 141–190
6. Flower DR (1996) The lipocalin protein family: structure and function. Biochem J 318:1–14
7. Pérez MD, Calvo M (1995) Interaction of β-lactoglobulin with retinol and fatty acids and its role as a possible biological function for this protein: a review. J Dairy Sci 78:978–988
8. Puyol P, Pérez MD, Ena JM, Calvo M (1991) Interaction of bovine β-lactoglobulin and other bovine and human whey protein with retinol and fatty acids. Agric Biol Chem 55:2515–2520
9. Harris S, Ali S, Anderson S, Archibald AL, Clark AJ (1988) Complete nucleotide sequence of the genomic ovine beta-lactoglobulin gene. Nucleic Acids Res 16:10379–10380
10. Alexander LJ, Hayes G, Bawden W, Stewart AF, MacKinlay AG (1993) Complete nucleotide sequence of the bovine beta-lactoglobulin gene. Anim Biotechnol 4:1–10
11. Folch JM, Coll A, Sánchez A (1994) Complete sequence of the caprine β-lactoglobulin gene. J Dairy Sci 77:3493–3497
12. Hayes HC, Petit EJ (1993) Mapping of the β-lactoglobulin gene and of immunoglobulin M heavy chain-like sequence to homologous cattle, sheep and goat chromosomes. Mamm Genome 4:207–210
13. Bell K, McKenzie HA (1967) The whey proteins of ovine milk β-lactoglobulin A and B. Biochim Biophys Acta 147:123–134
14. King JWB (1969) The distribution of sheep β-lactoglobulins. Anim Prod 11:53–57
15. Erhard G, Godovac-Zimmermann J, Conti A (1989) Isolation and complete primary sequence of a new ovine wild-type beta-lactoglobulin C. Biol Chem Hoppe-Seyler 370:757–762

16. Erhardt G (1989) Evidence of a third allele at the beta-lactoglobulin (beta-Lg) locus of sheep milk and its occurrence in different breeds. Anim Genet 20:197–204
17. Calavia MC (1997). Componentes y fenotipos de las caseınas y proteınas del lactosuero de leche de oveja (razas Lacha y Carranzana) comportamiento de las mismas durante la coagulation por quimosina y estabilidad termica de la β-lactoglobulina (components and phenotypes of ewe milk whey, caseins and proteins (Lacha and Carranzana breeds), their behavior during coagulation by chimosin and thermic stability of β-lactoglobulin). PhD Thesis, Universidad de Zaragoza, Zaragoza
18. Caroli A, Bolla P, Spanu A, Piredda G, Fraghì A (1995) Effect of β-lactoglobulin genotype on milk yield in Sardinian sheep. In: Proceedings 11th congress ASPA, Udine, Italia, pp 181–182
19. Fraghì A, Carta A, Pilla F, Sanna SR, Piredda G (1996) β-lactoglobulin polymorphism in Sarda dairy sheep. In: 47th Annual meeting of the EAAP, 42, Den Haag, The Netherlands
20. Giaccone P, Di Stasio L, Macciotta NPP, Portolano B, Todaro M, Cappio-Borlino A (2000) Effect of β-lactoglobulin polymorphism on milk-related traits of dairy ewes analysed by a repeated measures design. J Dairy Res 67:443–448
21. Dario C, Carnicella D, Bufano G (2005) Effect of β-lactoglobulin genotypes on ovine milk composition in Altamurana breed. Arch Zootec 54:105–108
22. Dario C, Carnicella D, Dario M, Bufano G (2008) Genetic polymorphisms of β-lactoglobulin gene and effect on milk composition in Leccese sheep. Small Rumin Res 74:270–273
23. Garzon AI, Martinez J (1992) β-Lactoglobulin in Manchega sheep breed: relationship with milk technological indexes in handcraft manufacture of Manchego cheese. Anim Genet 23:106
24. Lopez-Galvez, G, Ramos M., Martin-Alvarez, PJ, Juarez M (1993) Influence of milk protein polymorphism on cheese producing ability in the milk of Manchega ewes breed. In: Proceedings of the International Dairy Federation seminar "Cheese Yield and Factors Affecting its Control", Cork, Ireland, pp 167–173
25. Gutiérrez-Gil B, Arranz JJ, Othmane MH, de la Fuente LF, San Primitivo F (2001) Influencia del genotipo de la β-lactoglobulina ovina sobre caracteres cualitativos y rendimiento quesero individual en la raza Churra. ITEA 22:15–17
26. Recio I, Fernandez-Fournier A, Ramos M (1995) Genetic polymorphism of the whey proteins for two Spanish ovine breeds. Influence of genetic polymorphism of β-lg on renneting properties. In: Proceedings of the International Dairy Federation seminar, 28–29 Mar 1995, Zürich, Switzerland
27. Recio I, Fernández-Fournier A, Martín-Álvarez PJ, Ramos M (1997) β-Lactoglobulin polymorphism in ovine breeds: influence on cheese making properties and milk composition. Lait 77:259–265
28. Pietrolà E, Carta A, Fraghì A, Piredda G, Pilla F (2000) Effect of β-lactoglobulin locus on milk yield in Sarda ewes. Zootec Nutr Anim 26:131–135
29. Staiger EA, Thonney ML, Buchanan JW, Rogers ER, Oltenacu PA, Mateescu RG (2010) Effect of prolactin, β-lactoglobulin, and κ-casein genotype on milk yield in East Friesian sheep. J Dairy Sci 93:1736–1742
30. Watson CJ, Gordon KE, Robertson M, Clark AJ (1991) Interaction of DNA-binding proteins with a milk protein gene promoter in vitro: identification of a mammary gland–gland-specific factor. Nucleic Acids Res 19:6603–6610
31. Mink S, Härtig E, Jennewein P, Doppler W, Cato ACB (1992) A mammary cell-specific enhancer in mouse mammary tumour virus DNA is composed of multiple regulatory elements including binding sites for CTF/NFI and a novel transcription factor, mammary cell-activating factor. Mol Cell Biol 12:4906–4918
32. Burdon TG, Demmer J, Clark AJ, Watson CJ (1994) The mammary factor MPBF is a prolactin-induced transcriptional regulator which binds to STAT factor recognition sites. FEBS Lett 350:177–182
33. Braunschweig MH, Leeb T (2006) Aberrant low expression level of bovine β-lactoglobulin is associated with a C to A transversion in the *BLG* promoter region. J Dairy Sci 89:4414–4419
34. Scintu MF, Piredda G (2007) Typicity and biodiversity of goat and sheep milk products. Small Rumin Res 68:221–231
35. Portolano N (1987) La pecora della Valle del Belice. In: Edagricole (ed) Pecore e capre italiane. Bologna, Italy, pp 117–124
36. Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215
37. Sardina MT, Rosa AJM, Braglia S, Scotti E, Portolano B (2009) Identification of SNPs in the promoter of β-lactoglobulin gene in three Sicilian goat breeds. In: Proceedings 18th congress ASPA, Palermo, Italia, pp 147–149
38. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680
39. Schug J (2008) Using TESS to predict transcription factor binding sites in DNA sequence. Curr Protoc Bioinform 21:2.6.1–2.6.15
40. Wingender E, Dietze P, Karas H, Knuppel R (1996) TRANSFAC: a database on transcription factors and their DNA binding sites. Nucleic Acids Res 24:238–241
41. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452
42. Stephens M, Smith NJ, Donnely P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989
43. Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet 76:449–462
44. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599
45. Ganai NA, Bovenhuis H, van Arendonk JAM, Visker MHPW (2008) Novel polymorphisms in the bovine *β-lactoglobulin* gene and their effects on β-lactoglobulin protein concentration in milk. Anim Genet 40:127–133
46. Di Stasio L, Rasero R, Giaccone P, Fiandra P (1992) Valle del Belice sheep: genetic structure and relationship with other sheep populations reared in Sicily. Agric Mediterr 122:66–69