VETERINARSKI ARHIV 93 (2), 169-190, 2023 DOI: 10.24099/vet.arhiv.1378

Molecular characterization of complete coding sequence of the *MBL1* gene in the Indian Buffalo (*Bubalus bubalis)* breed

Manali Baghel[1], Deepak Sharma[2], Satyendra P. Singh[2]*, Vijay Pandey[3] and Avneesh Kumar[4]

[1]College of Biotechnology, College of Veterinary Sciences and Animal Husbandry, DUVASU, Mathura, India
[2]Department of Animal Genetics and Breeding, College of Veterinary Sciences and Animal Husbandry, DUVASU, Mathura, India
[3]Department of Veterinary Biochemistry, College of Veterinary Sciences and Animal Husbandry, DUVASU, Mathura, India
[4]Department of Animal Genetics and Breeding, College of Veterinary Sciences and Animal Husbandry, DUVASU, Mathura, India

BAGHEL, M., D. SHARMA, S. P. SINGH, V. PANDEY, A. KUMAR: Molecular characterization of complete coding sequence of the *MBL1* gene in the Indian Buffalo (*Bubalus bubalis)* breed. Vet. arhiv 93, 169-190, 2023.

ABSTRACT

Mannose-binding lectin (MBL) is a member of the collectin protein family that binds a broad range of microorganisms and activates the lectin-complement pathway of innate immunity. A number of mutations have been found in both the coding as well as the non-coding regions of the *MBL1* gene in various species, of which several variations affected the assembly of *MBL1*, thus leading to a low level of plasmic MBL and innate immune dysfunctions. In the present study, we have reported molecular cloning and characterization of the complete coding sequence of the *MBL1* gene in the Indian buffalo breed Murrah. A 951 bp fragment of the *MBL1* gene was amplified, cloned and sequenced. Multiple sequence alignment with other buffalo and cattle breeds revealed that the Murrah buffalo *MBL1* CDS was 98.1-99.6% homologous to other buffalo breeds, and 98.3-98.5% similar to cattle breeds at nucleotide level. It was 96.8-98.8% homologous to buffalo breeds and 96.8-97.2% similar to cattle breeds at amino acid level. The amino acid sequence of the Murrah buffalo MBL1 contained two non-synonymous amino acid substitutions (L204P and S180P). Further, PCR-RFLP was performed to screen 50 Murrah buffalo for the presence of SNPs, g.855G>A in intron I and g.2686T>C, as well as g.2651G>A in exon 2 region of the *MBL1* gene. The *Apa*I/intron I PCR-RFLP assay revealed a polymorphic pattern with three genotypes *viz.*, AG (90%), GG (8%) and AA (2%), with allelic frequencies 0.94 for G and 0.06 for A. *Hae*III/exon 2 PCR-RFLP assay exhibited the presence of three genotypes, namely, TC (66%), TT (32%) and CC (2%) with allelic frequencies 0.15 for T and 0.85 for C. StyI/exon 2. PCR-RFLP assay showed a monomorphic pattern for g.2651G>A with GG genotype only. We further examined the association of these SNPs with milk production traits and somatic cell score (SCS), and found no significant difference for any of the traits. Since the present study has formulated the results on the basis of a relatively small sample size, further studies with a larger sample size are required to validate the effects of polymorphisms.

Key words: buffalo **(***Bubalus bubalis*); *MBL1* gene; cloning; characterization; PCR-RFLP; genetic polymorphism**.**

*Corresponding author:
Dr. Satyendra Pal Singh, assistant professor, Department of Animal Genetics and Breeding, College of Veterinary Sciences and Animal Husbandry, DUVASU, Mathura-281001, India. E-mail: satsinpal2l@gmail.com

ISSN 0372-5480
Printed in Croatia

Introduction

Mannan-binding lectins (MBL) are collagenous C-type lectins (collectins) involved in the innate immune response to various microbial pathogens (NETH et al., 2000; JACK and TURNER, 2003). MBL is a pathogen recognition molecule (PRM) that identifies a repeating pathogen-associated molecular pattern (PAMP) on the surface of microorganisms through a C-type carbohydrate recognition domain (CRD) in a Ca^{++}-dependent way. It recognizes and binds to glycoconjugates containing mannose, N-acetylmannoseamine, L-fucose, glucose, and N-acetylglucosamine present on microbial surfaces (KAWASAKI et al., 1983; SHERIFF et al., 1994; HOLMSKOV et al., 1994).

MBL, upon recognition of infectious agents, activates a number of cellular defense mechanisms, including phagocytosis, modulation of cytokines and immunoglobulin secretion. In addition, MBL can activate the complement system, independent of the antibody, and C1q in association with two serine proteases, MBL-associated serine protease-1 (MASP-1) and MBL-associated serine protease-2 (MASP-2) resulting in the cleavage of the complement components C4 and C2, thus generating the C3 convertase (MATSUSHITA and FUJITA, 1992; THIEL et al., 1997; WALLIS, 2007), which results in either opsonization or lysis of the microorganism. The basic subunit of MBL is a trimer composed of three identical monomers, each consisting of four distinct domains: an N-terminal cysteine-rich domain, a collagen-like domain, a neck region and a C-terminal carbohydrate recognition domain (LILLIE et al., 2005).

Most mammals have two MBL genes, *MBL1* and *MBL2*, which encode the MBL-A and MBL-C proteins, respectively. However, in humans and chimpanzees, *MBL1* is an expressed pseudogene (MBL1P1), so only the MBL-C protein is functional (GUO et al., 1998; SEYFARTH et al., 2005). An ancient *MBL* gene most likely duplicated prior to human–murine divergence, suggesting that the human homologue to *MBL1* was lost during evolution (SASTRY et al., 1995), and an expressed human *MBL1* pseudo-gene has been identified. Although both forms of the protein activate the complement system, they differ significantly in oligomerization and in serum concentrations (LIU et al., 2001). The major site of MBL synthesis is the liver. *MBL1* was also found to be expressed in the lung, testis and brain, whereas low expression of *MBL2* was detected in the testis and kidney (PHATSARA et al., 2007).

A number of variations have been found in both the coding and the non-coding regions of the MBL gene in humans, sheep and porcine, of which several affect the assembly of MBL, thus leading to a low level of plasmic MBL and innate immune dysfunctions (THIEL and GADJEVA, 2009; JUUL-MADSEN et al., 2011a and 2011b; LIU et al., 2011; ZHAO et al., 2011). Both *MBL1* and *MBL2* mutations have been shown to vary the susceptibility of animals to various infections (LILLIE et al., 2005). Impaired disease resistance has been associated with 3 single-nucleotide polymorphisms (SNPs) within the coding region of *MBL1* in various breeds of pigs (LILLIE et al., 2006b). Moreover, in chickens of low MBL type, weight gain rate was reduced after an infection with *Escherichia coli* (NORUP et al., 2009). In addition, it has been found that *MBL* in Hu sheep harbors a vast number of polymorphisms (SNPs and deletions), influencing MBL concentration in serum (ZHAO et al., 2011).

MBL1 has been isolated from cattle (LOVELESS et al., 1989) and its mutations have been reported to contribute to variations of animal susceptibility to different infections, including dairy mastitis (LILLIE et al., 2005; CAPPARELLI et al., 2008; WANG et al., 2011; LIU et al., 2011). The bovine *MBL1* has been mapped to BTA28, and contains four introns and five exons, encoding a 248 aa protein (GJERSTORFF et al., 2004). In this region, it contains a quantitative trait locus (QTLs) affecting the somatic cell score (SCS) (http:// www.animalgenome. org/cgi-bin/ QTLdb/BT/ qdetails?QTL_ID=16056) (WANG et al., 2011). Several studies have reported that *MBL1* gene polymorphisms were associated with milk SCS in cattle (LIU et al., 2010; WANG et al., 2011; LIU et al., 2011 and YUAN et al., 2013). Milk somatic cell count (SCC) can serve as a surrogate trait for dairy mastitis resistance (KOIVULA et

al., 2005; SHARMA et al., 2006; TAL-STEIN et al., 2010; YUAN et al., 2011; YUAN et al., 2012). The selection of lower SCC and SCS for dairy breeding strategy could reduce the occurrence of dairy mastitis (LUND et al., 1999; CARLEN et al., 2004). A case-control study demonstrated that polymorphism at the *MBL2* locus of water buffalo (*Bubalus bubalis*) is associated with susceptibility and resistance to *Brucella abortus* infection (CAPPARELLI et al., 2008). *MBL1* in porcine and bovine may be regarded as a functional and positional candidate gene for mastitis resistance and complement activity (PHATSARA et al., 2007; WANG et al., 2011; LIU et al., 2011). However, the *MBL1* gene has not been well characterized in buffalo. To date only a partial *MBL1* gene sequence has been reported in buffalo (KM087783). Thus, characterization of the buffalo *MBL1* gene will fill the information gap, and identification of polymorphisms in buffalo *MBL1* will help to define future breeding strategies through marker assisted selection (MAS) (LUND et al., 1999; CARLEN et al., 2004).

These results suggest that MBL may play a role in innate immunity against a wide range of microorganisms, which is why we hypothesized that the buffalo *MBL1* gene may also be a promising genetic marker for selection of improved disease resistance and high quality milk production, since the buffalo is an economically important resource for the dairy industry. Considering the significance of several functions of the *MBL1* gene and the impact of its polymorphism on various traits, the objective of the present study was to clone and characterize the *MBL1* gene in buffalo (*Bubalus bubalis*) and to investigate polymorphisms in the buffalo *MBL1* gene, and the effect of these polymorphisms on milk production traits and somatic cell score.

Materials and Methods

Sampling. For cloning and characterization of the *MBL1* coding sequence (CDS), a liver tissue sample was collected randomly from adult female buffalo (n=2) of the Indian buffalo breed Murrah in a RNAlater (Thermo Fisher Scientific, USA) from a slaughter house. For detection of genetic polymorphism in the buffalo *MBL1* gene, blood samples from adult females of the Murrah breed of buffalo (n=50) were collected randomly in vacutainer tubes. All of the animals were maintained at the Instructional Livestock Farm Complex (ILFC), DUVASU, Mathura (U.P.), India.

Primer design. To amplify the complete CDS region of the *MBL1* gene, a specific primer pair was designed using VNTI software, and commercially synthesized on the basis of the *MBL1* gene sequence of *Bos taurus* available in the GenBank (Acc. no. NM_001010994). The sequence of the primers are given in Table 1.

Table 1. Primer pair used to amplify complete coding region of buffalo *MBL1* gene

Region of *MBL1* gene	Primer Name	Primer Sequence (5′→3′)	Amplicon Size
MBL1 CDS	MBLcDNA F	GATGTTGCTGAGGCATCCGC	951 bp
	MBLcDNA R	TCCAGAGCAGCAGGGAGACA	

Molecular cloning of buffalo MBL1 CDS.

RNA extraction. The total RNA was isolated from freshly collected liver tissue using the GeneJET RNA purification kit (Thermo Fisher Scientific, USA), following the manufacturer's instructions. The purity of total RNA was checked using a Biospectrometer (Eppendorf, Germany). The quality of RNA was checked using 1% agarose gel electrophoresis.

cDNA synthesis. cDNA was synthesized from the RNA samples in a 12-µL reaction mixture using a Revert Aid First strand cDNA Synthesis Kit (Thermo Fisher Scientific, USA), according to the manufacturer's instructions.

Amplification of MBL1 cDNA. PCR was performed to amplify the complete coding region of the *MBL1* gene using specific primers as mentioned above, where PCR was carried out in a total volume of 25 µl containing 1 µl of first strand of cDNA, 10 X PCR buffer (New England Biolab, USA), 2 mM $MgCl_2$, 2 mM of dNTPs, 5 p moles of each primer, one unit of *Taq* DNA polymerase (New England Biolab, USA), and autoclaved double distilled water. The PCR profile consisted of a denaturation step at 94°C for 50 sec, an annealing step at 65°C for 50 sec and an elongation step at 72°C for 1 min, for a total of 34 cycles, followed by a final extension of 10 min at 72°C. The PCR product was checked by agarose (1%) gel electrophoresis containing ethidium bromide in 1X TAE buffer.

MBL1 CDS cloning and sequencing. Gene cloning was performed using a PUREGENE-Quickclone PCR cloning kit (Genetix Biotech Asia, India). Amplified PCR DNA products were ligated to the pTZ57R/T cloning plasmid vector, and the ligation reaction product was transformed into *Escherichia coli* DH5α host cells following the manufacturer's instructions. The recombinant clones (DH5α *E. coli* with PCR product) were identified from the transformed bacterial colonies using blue (non-recombinant) and white (recombinant) colony selection. DH5α bacterial culture derived from cloning was harvested by centrifugation, and recombinant plasmid (plasmid vector with PCR product) purification was carried out using a GeneJET plasmid miniprep kit (Thermo Fisher Scientific, USA), following the manufacturer's instructions. Further, the presence of the insert was confirmed by restriction digestion with *EcoR*I and PCR amplification of the insert, using recombinant plasmid as a template. The *MBL1* gene insert was amplified using gene specific primers, as described above. The positive clones were sequenced commercially by an automated sequencer using standard cycle conditions by Sanger′s dideoxy chain termination method with standard sequencing primers.

Sequencing analysis. The sequences obtained were subjected to BLAST analysis (www.ncbi.nlm.nih.gov/BLAST) to ascertain whether the obtained sequences corresponded to *MBL1*. The nucleotides as well as the deduced amino acid sequences of Murrah buffalo *MBL1* CDS were aligned with other buffalo and cattle breeds, and other species available in the GenBank database using the Clustal method of the MegAlign program of Lasergene software (DNASTAR, USA) and BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html). A phylogenetic and molecular evolutionary analysis was conducted using MEGA version 4.0 (TAMURA et al*.,* 2007). The amino acid sequence was also deduced to analyze the buffalo MBL1 protein sequence using online translation website (http://reverse-complement.com/translate-protein/ROOT/ .)

The comparative analysis of the sequence included the complete CDS of sub family Bovinae, such as *Bison bison* or the American Bison (American buffalo, XM_010834154), *Bubalus bubalis* (Italian Mediterranean buffalo, XM_006064773), *Bos taurus* (Mixed breed, NM_001010994), *Bos taurus* (Crossbred x Angus, BC109674), *Bos taurus* (AB178774), *Bos indicus* (Nelore, XM_019954216), hybrid *Bos indicus x Bos taurus* (Angus x Brahman F1 hybrid, XM_027530780) and *Bos mutus* (Wild Yak, XM_005909163), the sub family Caprinae, such as *Capra hircus* (San Clemente, XM_005699265), *Ovis aries* (Rambouillet, XM_004021518) and *Ovis aries* (Domestic sheep, LT852560). Two breeds of *Sus scrofa* or pig belonging to the Suidae family were also included in the comparative analysis, that is, Yorkshire (NM_001007194) and Landrace (EU421730). Animals belongs to the Camelidae family were also included, such as *Camelus bactrianus*/Bactrian camel (Alxa, XM_010963110) and *Camelus dromedaries*/Arabian camel (Breed African, XM_010984099).

PCR-RFLP for screening of polymorphisms in the MBL1 gene

DNA extraction and PCR reaction. Three SNPs, viz. g.855G>A in intron I and g.2686T>C as well as g.2651G>A in the exon 2 region of the *MBL1* gene were targeted and genotyped by PCR-RFLP. Blood samples were collected in 0.5 M EDTA (pH

8.0), and genomic DNA was isolated by using the standard phenol–chloroform DNA isolation protocol (SAMBROOK and RUSSELL, 2001). The concentration and purity of the genomic DNA was determined spectrophotometrically at OD260 and OD280. The integrity of the DNA was examined by agarose gel (1%) electrophoresis, and the gel was visualized under a UV light after staining with ethidium bromide (EtBr). To detect polymorphism, three primer pairs were used to amplify the intron I and exon 2 regions of the *MBL1* gene, adapted from WANG et al. (2011) and shown in Table 2. PCR reactions were carried out in 25µl reaction mixture containing 10X PCR buffer (New England Biolab), 2 mM $MgCl_2$, 2.5 mM of dNTPs, 5 pmoles of each primer, and one unit of *Taq* DNA polymerase (New England Biolab, USA). After an initial denaturation at 94ºC for 5min, the PCR amplification was performed in 35 cycles using the following parameters: denaturation at 94°C for 30 s for g.855G>A and g.2686T>C, while 94°C for 20 sec for g.2651G>A; annealing at 59, 62 and 53°C (Table 2) for 30 s, and elongation at 72ºC for 30 s. The reaction was continued for a final extension at 72 °C for 10 min.

Table 2. PCR–RFLP assay for genotyping SNPs in the buffalo *MBL1* gene.

SNP/ location	Primer Sequence (5′→3′)	SAF (bp)	AT (ºC)	RE	Possible genotypes
g.855G> A Intron I	F: CCCTTCCAACTCATTGCTTC R: AGTCCCAACCACCCTCA	588	62	*Apa*I	GG: 311 and 277 bp GA:588, 311and 277 bp AA: 588 bp
g.2686T> CExon 2	F: GCAGAGGTGGTGGCAAATGT R: CATCTTTAGAGAGAATGCCCC	401	59	*Hae*III	TT: 274 and 127 bp TC: 274, 184, 127 and 90 bp CC: 184, 127 and 90 bp
g.2561G> A Exon 2	F: GGTGGCAAATGTTGGCTA R: GTCTTCTGAGCATCCTCCA	162	53	*Sty*I	GG: 162 bp GA: 162, 141 and 21 bp AA: 141 and 21 bp

SAF: size of amplification fragment; AT: annealing temperature; RE: restriction enzyme, bp: base pair.

The amplicon of 588 bp obtained from the intron I region was digested using restriction enzyme (RE) *Apa*I (New England Biolab, USA). The restriction digestion was carried out at 25ºC for 12–14 hr. The amplified product of 401 bp and 162 bp obtained from the different sets of exon 2 region specific primers was digested using RE *Hae*III (New England Biolab) and RE *Sty*I (New England Biolab), respectively, and the digestion was carried out at 37ºC for 12–14 hr. A reaction mix of 15 µl total volume, containing 5 µlPCR products, 0.5 µl *Apa*I/*Hae*III/*Sty*I (10U/µl) and 1.5 µl 10X RE buffer (specific for each RE) was prepared for each restriction digestion reaction. For genotyping, digested products were checked on 2.0% agarose gel in 1X TAE buffer. The fragments were visualized under a UV light after staining with ethidium bromide.

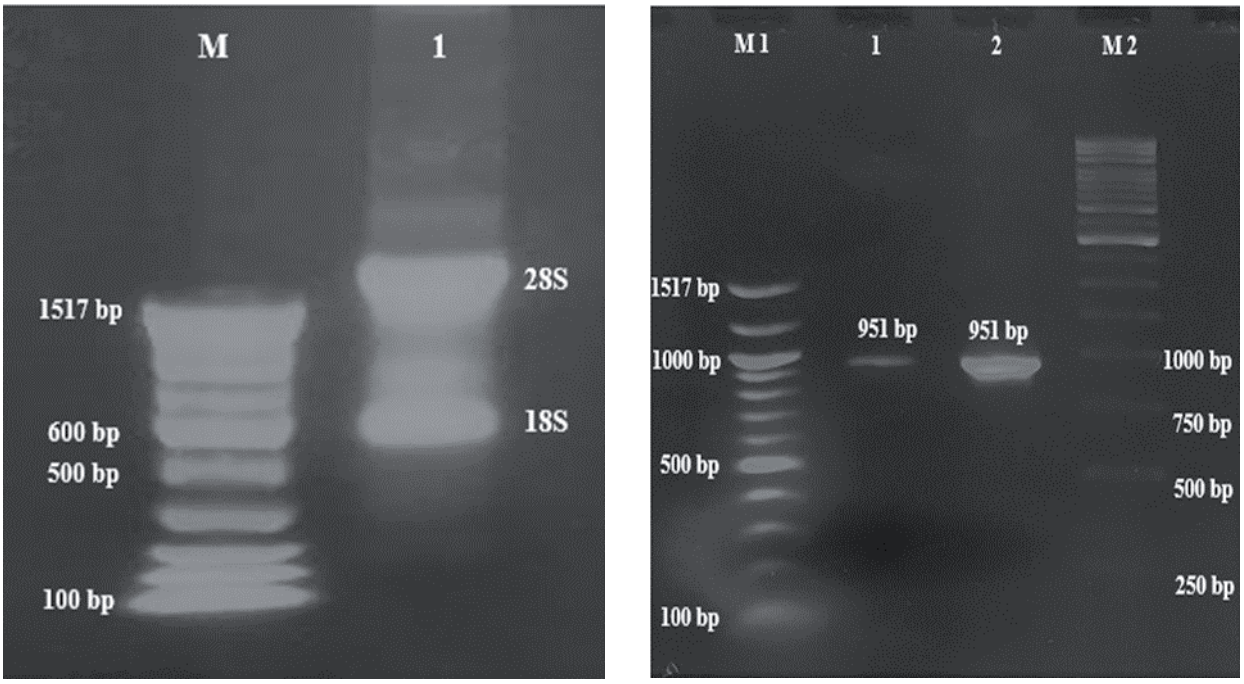
Statistical analysis. Milk samples were collected from the cattle and buffalo population within the polymorphic study and analyzed using a LACTOSCAN somatic cell counter using the florescent microscope technique of counting cells. Since the distribution frequency of the somatic cell count is usually skewed (ALI and SHOOK, 1980), we calculated the SCC (cells/µl) based parameter SCS using the following equation: SCS = log2 (SCC/100) + 3 (RUPP and BOICHARD, 1999).

The genotypic and allelic frequencies were estimated by the standard procedure (FALCONER

and MACKAY, 1996). An association study of genotyped animals was carried out with somatic cell count of milk and other milk production traits, including age at first calving (AFC), lactation period (LP), total milk yield (TMY) and milk yield at 300 days (MY300), to analyze the effect of polymorphisms.

Results

RNA isolation and cloning of MBL1 CDS. Total RNA was isolated and visualized under a gel documentation system and revealed two bands, *viz.,* 28S and 18S of rRNA (Fig. 1a). A 951 bp amplicon of *MBL1* cDNA encoding complete CDS was cloned. Two positive clones were picked up and the presence of a gene insert was confirmed by plasmid PCR using *MBL1* CDS specific primers (Fig. 1b).



a) b)

Fig. 1. (a). Agarose (1%) gel electrophoresis showing bands of RNA. Lane M: Marker (100 bp DNA ladder), Lane 1: 28S and 18S band of RNA. (b). Recombinant plasmid based PCR products of 951 bp. Lane M1: Marker (100 bp DNA ladder), Lane 1-2: PCR products of 951 bp obtained using plasmid of 54.5 ng/µl and 86.4 ng/µl, Lane M2: Marker (250 bp DNA Ladder).

Sequencing and sequence analysis. The size of the gene insert/amplicon was further confirmed by commercial sequencing of recombinant plasmid DNA. After sequencing, a sequence similarity search was carried out in the GenBank database using the NCBI BLAST tool. The BLAST result showed the most resemblance to *MBL1* sequences from other buffalo/cattle breeds, confirming that the nucleotide sequence obtained in the present study was exactly the expected CDS encoding MBL1. The *MBL1* CDS sequence of Indian Murrah buffalo breed was submitted to the GenBank database with accession no. MN990687.

Sequence analysis revealed that the 951 bp *MBL1* sequence contained a long 158 bp (part of exon 1 and exon 2) a 5′-untranslated region (UTR), followed by 57 bp analogous to a signal peptide of 19-aa residues, 690 bp analogous to the mature

protein, and a sort 46 bp 3′-UTR, as presented in Fig. 2. Further analysis revealed it consisted of five exons including: 149 bp long first non-coding exon 1 (1-149) that was transcribed but untranslated, 190 bp long exon 2 (150-339) which had 9 non-coding nucleotides (150-158) of 5′ UTR at start, 117 bp long exon 3 (340-456), 75 bp long exon 4 (457-531) and 420 bp long incomplete exon 5 having complete coding of part of it, as only 374 bp (from 532 to 905) of exon 5 were part of CDS while the remaining nucleotides (46 bp) were noncoding (part of 3′ UTR). Exon 2 encoded the signal peptide (1-19 aa), a cysteine rich N-terminus region (20-42 aa) and a part of a collagenous region (43-60 aa), exon 3 encoded the remainder of the collagenous region (61-98 aa), exon 4 encoded an α-helical coiled-coil structure known as the 'neck region' (99-124 aa), and exon 5 encoded the Carbohydrate Recognition Domain (CRD; 125-248 aa).

Fig. 2. Full-length sequence (951 bp) of *MBL1* cDNA encoding the Murrah buffalo MBL-A protein with predicted amino acid translation.
The numbers indicate amino acid (right side) and nucleotide (left side) position. The underlined sequences are non-coding regions of exon 1, 5′UTR and 3′UTR.

Sequence analysis revealed that at nucleotide level the Murrah buffalo *MBL1* sequence in the present study was more similar to that of exotic Mediterranean buffalo (99.6%) than another (partial) Indian Murrah buffalo (99.1%) and American buffalo (*Bison bison*) (98.1%) sequences, as shown in Fig. 4. It was more similar to exotic cattle breeds (98.3-98.5%) than American buffalo (98.1%). Among all the cattle breeds, the *MBL1* sequence of Murrah buffalo showed more similarity with Nelore cattle (*Bos indicus*, Brazilian origin) (98.5%). It showed 98.4% similarity with its wild relative *Bos mutus* (yak). The Murrah buffalo *MBL1* sequence showed 96.1% and 95.2% similarity

with sequences of goat and sheep, respectively. It showed 83.5-83.7% and 83.5% similarity with the camel breeds (Alxa Bactrian camel and Arabian camel) and the pig breeds (Landrace pig and Yorkshire pig), respectively.

Fig. 3. Location of putative structural domains based on exon structure, amino acid sequence, and homologous mammalian proteins: Signal peptide—aa 1–19; N-terminal domain—aa 20–42; collagen domain—aa 43–99; neck region—aa 100–124; CRD—aa 125–248.
Underlining shows conserved sequences of MBL-A protein

Percent Identity

Divergence

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1		98.8	98.8	96.8	97.2	96.8	97.6	97.2	96.8	95.2	94.4	97.6	79.9	79.9	81.9	81.9	1
2	1.2		99.6	97.6	98.0	97.6	98.4	98.0	97.6	95.2	94.4	98.4	79.8	79.8	82.3	82.3	2
3	0.8	0.4		97.6	98.0	97.6	98.4	98.0	97.6	95.2	94.4	98.4	79.5	79.5	81.9	81.9	3
4	2.9	2.5	2.0		99.2	98.8	98.8	99.2	98.8	94.8	94.0	98.8	79.1	79.1	80.3	80.3	4
5	2.5	2.1	1.6	0.4		99.2	99.2	99.6	99.2	95.2	94.4	99.2	79.5	79.5	80.7	80.7	5
6	2.9	2.5	2.0	0.8	0.4		98.8	99.2	99.6	95.6	94.4	98.8	79.5	79.5	80.7	80.7	6
7	2.0	1.6	1.2	0.8	0.4	0.8		99.2	98.8	94.8	94.0	98.8	79.1	79.1	80.7	80.7	7
8	2.5	2.1	1.6	0.4	0.0	0.4	0.4		99.2	95.2	94.4	99.2	79.5	79.5	80.7	80.7	8
9	2.9	2.5	2.0	0.8	0.4	0.0	0.8	0.4		95.6	94.4	98.8	79.5	79.5	80.7	80.7	9
10	4.6	5.0	4.6	5.0	4.6	4.1	5.0	4.6	4.1		98.4	95.6	80.7	80.7	81.1	81.1	10
11	5.4	5.9	5.4	5.9	5.4	5.4	5.9	5.4	5.4	1.2		94.8	80.7	80.7	80.7	80.7	11
12	2.0	1.6	1.2	0.8	0.4	0.8	0.8	0.4	0.8	4.1	5.0		79.9	79.9	81.1	81.1	12
13	21.4	22.0	21.9	22.4	21.9	21.9	22.4	21.9	21.9	20.3	20.3	21.4		99.6	81.1	81.1	13
14	21.4	22.0	21.9	22.4	21.9	21.9	22.4	21.9	21.9	20.3	20.3	21.4	0.0		81.1	81.1	14
15	20.3	20.4	20.3	22.4	21.9	21.9	21.9	21.9	21.9	21.4	21.9	21.4	19.8	19.8		99.6	15
16	20.3	20.4	20.3	22.4	21.9	21.9	21.9	21.9	21.9	21.4	21.9	21.4	19.8	19.8	0.0		16
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	

1 Murrah Buffalo_MN990687.pro
2 Murrah Buffalo_KM087783.pro
3 Mediterranean Buffalo_XM006064773.pro
4 American Buffalo_XM010834154.pro
5 Angus X Brahman Cattle_XM027530780.pro
6 Crossbred X Angus Cattle_BC109674.pro
7 Nelore Cattle_XM019954216.pro
8 Bos taurus Cattle_AB178774.pro
9 Mixed Breed Cattle_NM001010994.pro
10 San clemente Goat_XM005699265.pro
11 Rambouilet Sheep_XM004021518.pro
12 Yak_XM005909163.pro
13 Landrace Pig_EU421730.pro
14 Yorkshire Pig_NM001007294.pro
15 Alxa Bactrian Camel_XM010963110.pro
16 Arabian Camel_XM010984099.pro

Fig. 4. Percentage identity of MBL1 CDS of Murrah buffalo breed with other breeds of buffalo/cattle and other related species on the basis of amino acid sequences (above diagonal) and percentage divergence (below diagonal) estimated using DNASTAR.

At amino acid level, it was found that the MBL1 sequence in the present study was equally similar (98.8%) to another Murrah buffalo and exotic Mediterranean buffalo. It was least similar to exotic American buffalo (*Bison bison*) (96.8%) among buffalo breeds. It was more similar to exotic cattle breeds (96.8-97.2 %) than the American buffalo (96.8%). It showed 97.6% similarity with its wild relative, *Bos mutus* (yak), 95.2%, and 94.4% similarity with sequences of goats and sheep, respectively. It was least similar to the pig breeds (79.9%) followed by the camel breeds (81.9%).

Identification of polymorphisms in MBL1 CDS. The multiple sequence alignment of the deduced amino acid sequence of Murrah buffalo *MBL1* CDS with that of other buffalo and exotic cattle breeds, and other related species is presented in Fig. 5. We observed one unique synonymous nucleotide change G→A at position 87 which resulted in CAG (Glu) to CAA (Glu) at position 29 of the amino acid sequence, but did not change the amino acid composition. Additionally, it contained two unique non-synonymous nucleotide changes, including a T→C transition at nucleotide position 538, which led to amino acid substitution of "Ser" to "Pro" at the 180 position of the amino acid sequence, and a T→C transition at nucleotide position 611, which resulted in amino acid substitution of "Leu" to "Pro" at position 204 in the amino acid sequence (Table 3).

Table 3. Different polymorphic sites in the Indian buffalo breed and exotic buffalo/cattle breeds

Nucleotide	87	611	538	451	178	103
Amino acid	29	204	180	151	60	34
Region of *MBL1* CDS	Exon 1	Exon 4	Exon 4	Exon 4	Exon 1	Exon 1
Region of MBL-A protein	Mat_P (N-terminal domain)	Mat_P (CRD)	Mat_P (CRD)	Mat_P (CRD)	Mat_P (CLD)	Mat_P (N-terminal domain)
Codon Indian buffalo breed	CAA	CCT	CCT	GGC	GAG	GTG
Codon exotic buffalo/ cattle breeds	CAG	CTA	TCT	AGC	AAG	GTA
AA Indian buffalo breed	Q	P	P	G	E	V
AA exotic buffalo/cattle breeds	Q	L	S	S	K	V
Breed	Murrah buffalo	Murrah buffalo	Murrah buffalo	Murrah buffalo	American buffalo	All cattle breeds

Nucleotide: position of the polymorphism in the nucleotide sequence, amino acid: affected amino acid site, region of *MBL1* CDS: coding region in which the polymorphic site is located, region of MBL-A protein: coding region in which the polymorphic site is located, Mat_P: mature peptide domain, CRD: carbohydrate recognition domain, CLD: collagen like domain, AA: amino acid.

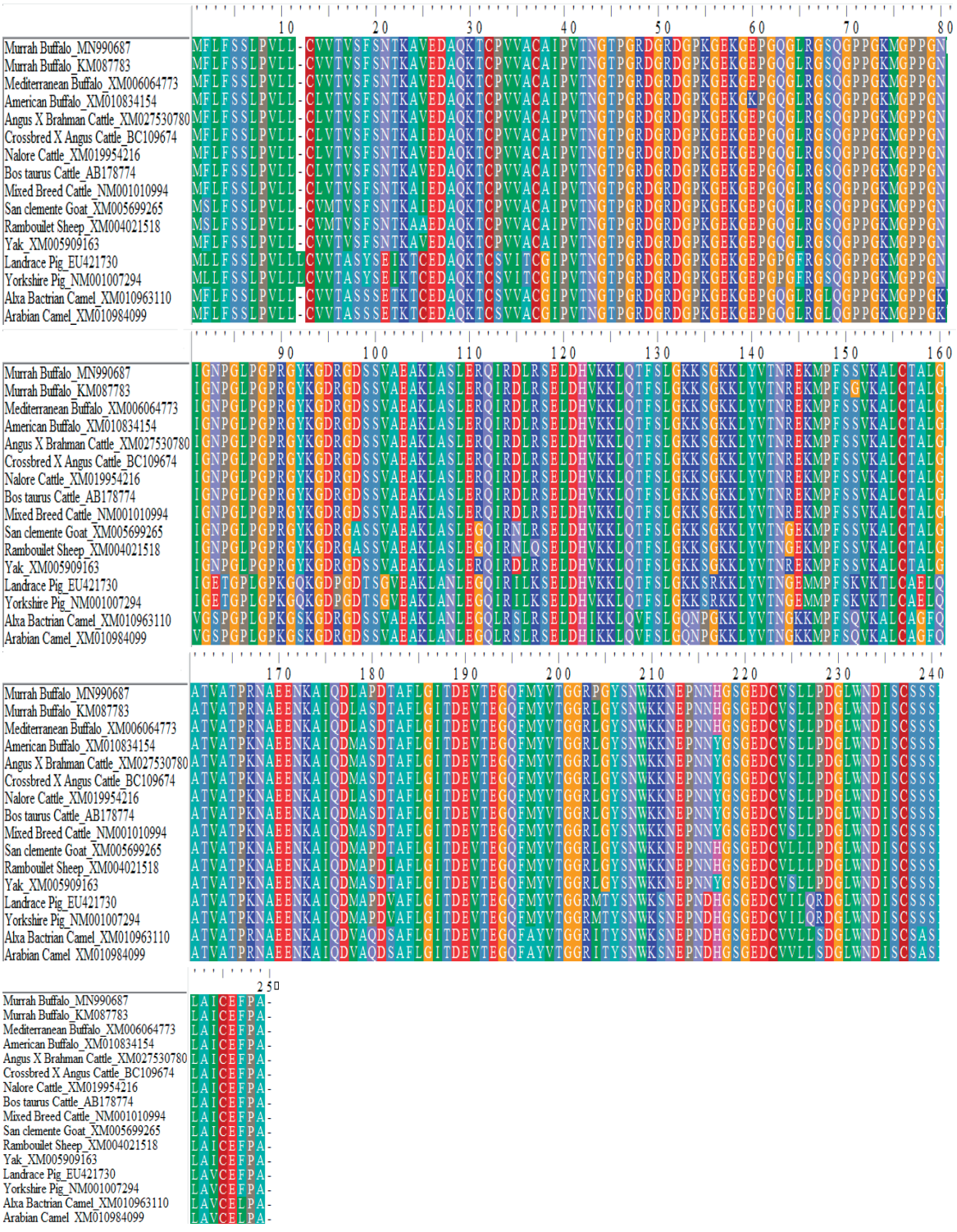


Fig. 5. Multiple sequence alignment of *MBL1* CDS of Murrah buffalo breed with other breeds of buffalo and other related species on the basis of deduced amino acid sequences estimated using Bioedit software.

Apart from these nucleotide changes found in the MBL1 CDS of Murrah buffalo, a unique non-synonymous G→A nucleotide change at nucleotide position 178 was observed in American buffalo, which resulted in amino acid substitution of "Glu" acid to "Lys" at position 60 in the amino acid sequence. A unique non-synonymous nucleotide change at position 451A→G was also observed in the previously reported partial Murrah buffalo sequence, which caused a non-synonymous mutation with amino acid substitution of "Ser" to "Gly" at position 151. Moreover, all the cattle breeds contained a synonymous nucleotide change, G→A transition at nucleotide position 103, which resulted in no change in amino acid composition (V34V).

Phylogenetic analysis. The phylogenetic tree (Fig. 6) clearly indicates the splitting of the Bovidae family into two clades: the bovinae and caprinae subfamilies, showing a greater degree of variation, leaving suidae and camelidae in two different clads. Within the bovinae group, the Murrah buffalo is seen to be most closely related to the Mediterranean buffalo. However, the American buffalo is seen to be most closely related to cattle breeds and Yak. The caprinae group is seen to be more closely related to cattle and buffaloes than the suidae and camelidae group. In the phylogenetic tree, all buffalo breeds exhibited a close phylogenetic evolutionary relationship with cattle and yak.

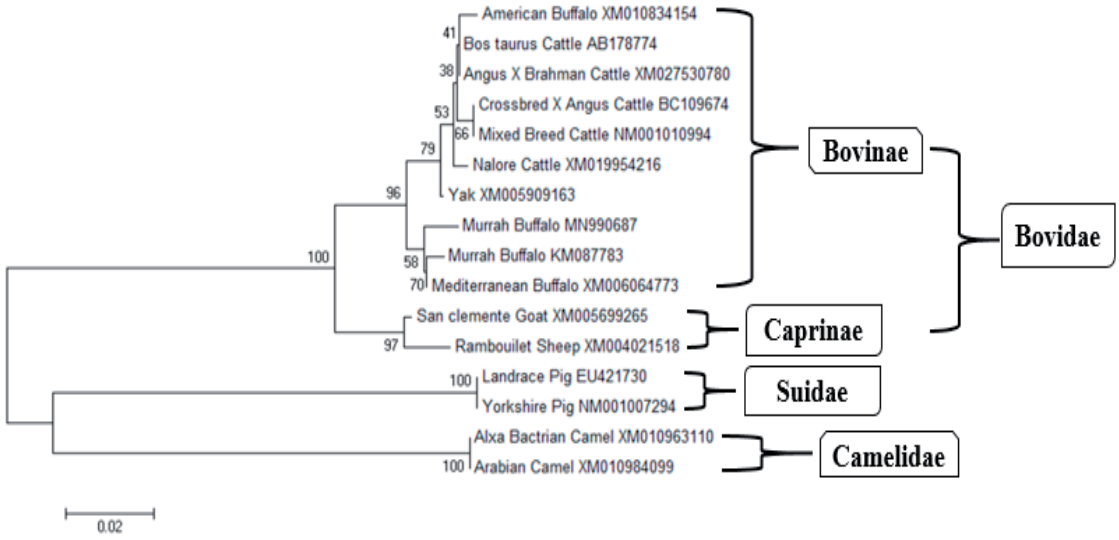


Fig. 6. Phylogenetic analysis based on deduced amino acids sequences of *MBL1* CDS for Murrah buffalo with other buffalo breeds and other related species using the Neighbor-Joining method (Bootstrap test of phylogeny). The percentage of replicate trees (bootstrap value) in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches.

PCR-RFLP assay to screen SNPs in MBL1 CDS of Murrah buffalo. Restriction digestion of 588 bp PCR product (containing g.855G>A locus) by *Apa*I produced fragments with lengths of 311 bp and 277 bp for genotype GG, 588 bp, 311 bp and 277 bp for genotype GA, and 588 bp for genotype AA, respectively (Fig. 7a). Further, digestion of the PCR fragment (containing g.2686T>C locus) by *Hae*III produced two bands (274 bp and 127 bp), four bands (274 bp, 184 bp, 127 bp and 90 bp) and three bands (184 bp, 127 bp and 90 bp) for TT, TA and AA genotypes, respectively. The 90 bp band was not clearly visible due to the long run (Fig. 7b). Restriction digestion of 162 PCR product (containing g.2651G>A locus) by *Sty*I produced a single banding pattern with undigested product of 162 bp for wild type genotype GG, as presented in Fig. 7c, which indicated the absence of a recognition site for RE *Sty*I.

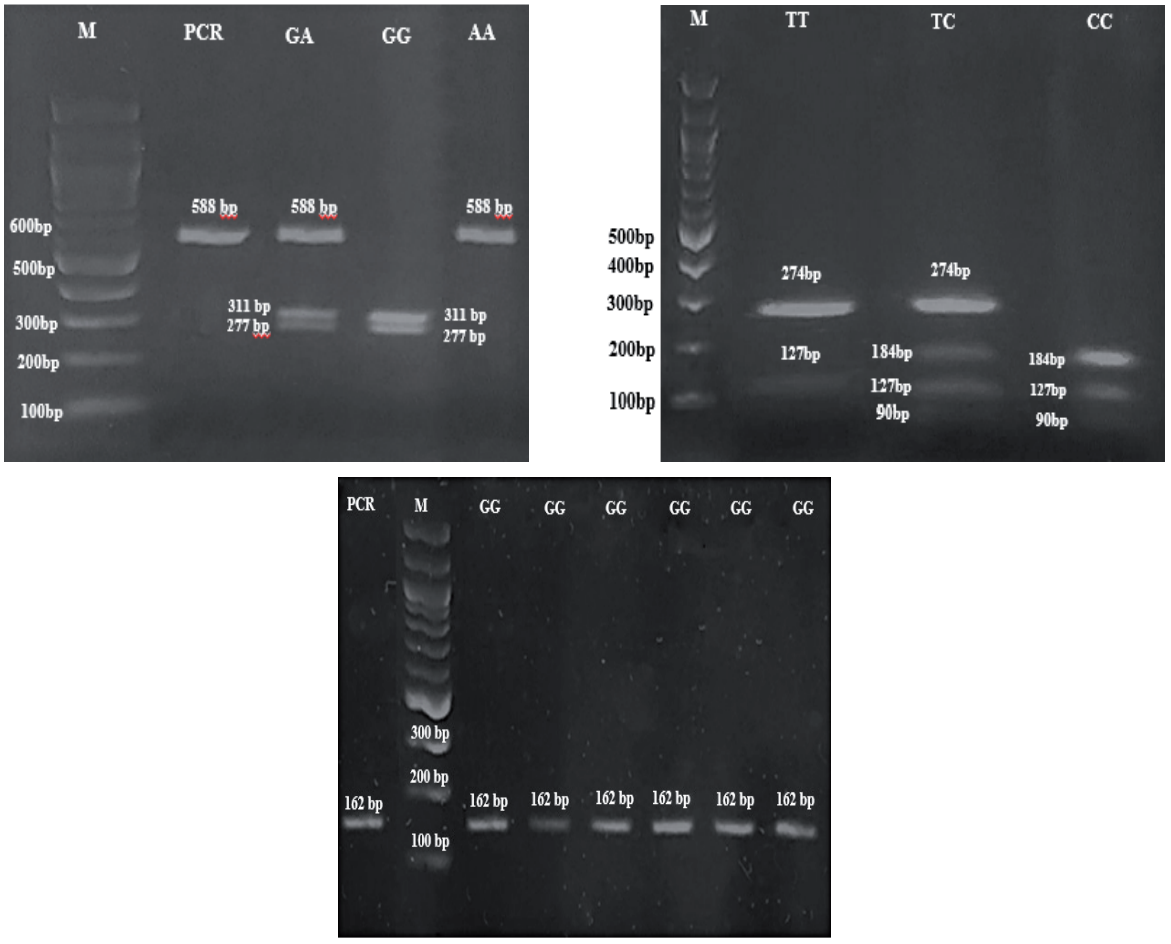


Fig. 7. Identification of SNPs g.855G>A, g.2686T>C and g.2561G>A in the MBL1 gene by PCR–RFLP genotyping. (a) Restriction digestion with *ApaI* at the MBL1 g.855G>A locus produced fragments with lengths of 311 and 277 bp for homozygous genotype GG, 588, 311 and 277 bp for heterozygous genotype GA, 588 bp for homozygous genotype AA, respectively; (b) Digestion of the PCR fragment by *Hae*III at the g.2686T>C locus produced two bands (274 and 127 bp), four bands (274, 184, 127 and 90 bp), and three bands (184, 127 and 90 bp) for TT, TA and AA genotypes, respectively; (c) Digestion of the PCR fragment by *Sty*I at the g.2651G>A locus produced a single band (162 bp) for genotype GG.

The allele and genotype frequencies of the three SNPs in the buffalo *MBL1* gene are shown in Table 4. The alleles G and C were the predominant alleles at positions g.855G>A and g.2686T>C, respectively, in the screened buffalo population. The frequency of G was 0.94 for SNP g.855G>A, and the frequency of C was 0.85 for SNP g.2686T>C in the genotyped animals. The Chi square ($\chi 2$) test indicated that SNP g.855G>A met the Hardy-Weinberg equilibrium ($p<0.05$), having a chi square value of 4.2, while position g.2686T>C did not meet the equilibrium ($p>0.05$) having a chi square value of 3.89 in all screened buffaloes. For SNP g.2651G>A, only GG genotypes having 162 bp band size (100%) with G allele (1.0) were present in all the screened buffalo.

Association of SNPs with milk production traits and SCS. The results of the association study of two SNPs, g.855G>A and g.2686T>C, of the buffalo *MBL1* gene with milk production traits are shown in Table 5 and Table 6, respectively. We observed

that both SNPs (g.855G>A and g.2686T>C) showed no signification difference (P>0.05) for any of the milk production traits between the genotypes in the first and second lactations. Association studies of g.855G>A and g.2686T>C genotypes with milk SCS revealed no significant (P>0.05) difference (Table 7). An association analysis of this g.2651G>A could not be performed with milk production traits and SCS due to the fact that the monomorphic pattern had only one genotype.

Table 4. Genotype distribution, allelic frequencies and chi square test of two SNPs found in the buffalo *MBL1* gene.

No. of samples	g.855G>A						g.2686T>C					
	Genotype frequency/ sample size			Allele frequency		Chi square (χ2)	Genotype frequency/ sample size			Allele frequency		Chi square (χ2)
50	GG	GA	AA	G	A	4.2 (p<0.05)	TT	TC	CC	T	C	3.89 (p>0.05)
50	90.00 (n=45)	8.00 (n=4)	2.00 (n=1)	0.94	0.06		6.00 (n=3)	18.00 (n=9)	76.00 (n=38)	0.15	0.85	

n: size of sample

Table 5. Association studies of g.855G>A genotypes with milk production traits in Murrah buffalo

Lactation	Genotype	AFC (days)	TMY (liters)	LP (days)	MY300 (liters)
I (N=49)	GG (44)	1960.0 ±105.6 (42)	1850.0 ±195.0 (42)	363.0 ±51.8 (42)	1600.0 ±91.8 (42)
	GA (4)	1620.0 ±216.0 (3)	1990.0 ±183.0 (3)	375.0 ±28.8 (3)	1590.0 ± 65.7 (3)
	AA (1)*	1735 (1)	1840 (1)	260 (1)	1840 (1)
II (N=43)	GG (39)	-----	2320.0 ±127.0 (39)	370.0 ±48.4 (39)	1860.0 ±72.7 (39)
	GA (3)	-----	1770.0 ±156.0 (3)	296.0 ±26.3 (3)	1770.0 ±61.2 (3)
	AA (1)*	-----	2162 (1)	358 (1)	1765 (1)

AFC: Age at first calving; TMY: Total milk yield; LP: Lactation period and MY300: Milk yield upon 300 days. *Only one AA genotype was present in the screened population so it was not included in association analysis.

Table 6. Association studies of g.2686T>C genotypes with milk production traits in Murrah buffalo

Lactation	Genotype	AFC (days)	TMY (liters)	LP (days)	MY300 (liters)
I (N=47)	TT (2)	1760.0±226.0 (2)	1400.0±194.0 (2)	271.0±10.5 (2)	1570.0±276.0 (2)
	TC (9)	1860.0±169.0 (9)	2010.0±204.0 (9)	392.0±39.9 (9)	1550.0±69.7 (9)
	CC (36)	1950.0±113.0 (36)	1940.0±98.6 (36)	363.0±17.1 (36)	1620.0±55.9 (36)
II (N=41)	TT (2)	------	2870.0±251.0 (2)	466.0±51.2 (2)	1850.0±87.3 (2)
	TC (7)	------	2580.0±339.0 (7)	371.0±36.8 (7)	2060.0±103.0 (7)
	CC (32)	------	2190.0±130.0 (32)	363.0±21.3 (32)	1800.0±60.3 (32)

AFC: age at first calving, TMY: total milk yield, LP: lactation period and 300 days MY: 300 days milk yield.

Table 7. Association studies of g.855G>A and g.2686T>C genotypes with somatic cell count in buffaloes

Genotype/sample size	Somatic cell score
g.855G>A	
GG	51.0±2.9 (41)
GA	56.2±7.8 (4)
AA*	57.2 (1)
g.2686T>C	
TT*	63.2 (1)
TC	50.7±6.8 (8)
CC	51.9±3.2 (35)

* Only one AA and TT genotype was present in the screened population so they were not included in association analysis.

Discussion

In the present study, we cloned and characterized the complete coding sequence of the *MBL1* gene in the Indian buffalo breed Murrah. It provides an opportunity to explore novel mutations and their exploitation to enhance milk production. We obtained 951 bp cDNA having a long 5′ UTR followed by a 747 bp coding region encoding 248 aa protein and a short 3′ UTR. Similarly, a 747 bp CDS encoding a protein of 248 aa residues in bovine (*Bos taurus*) was reported by KAWAI et al. (1997) and LILLIE et al. (2006a).

The 951 bp *MBL1* cDNA contained a 149 bp segment as exon 1 that was transcribed but

untranslated. This type of non-coding sequence was referred to as exon 0 or exon 1 in the previously characterized *MBL1* gene of several species. WANG et al. (2011) and LIU et al. (2011) also mentioned the non-coding exon in their study on China Holstein cattle and represented it as exon 1. Such an exon has also been found in mice (LAURSEN and NIELSEN, 2000), pigs (LILLIE et al., 2007) and murine (SASTRY et al., 1995) *MBL1* genes as well as in rats and mice *MBL2* genes (LAURSEN and NIELSEN, 2000). NAITO et al. (1999) also reported an alternative exon (exon 0) 1.0 kb upstream of the *MBL2* exon 1 in humans.

On comparison with previously reported partial *MBL1* CDS sequence (Accession no. KM087783) of Murrah buffalo, the present study sequence exhibited 99.1% homology. *MBL1* CDS of the present study was characterized by cloning and sequencing of *MBL* mRNA isolated from buffalo livers. In contrast, the previous sequence was characterized by genomic DNA sequencing of the *MBL1* gene. Moreover, the last three nucleotides, *viz.* TGA codes for stop codon lacked nucleotide 'A' in the previous sequencing of Murrah buffalo. The presence of a stop codon is generally a signal to terminate protein synthesis, and this process constitutes the last essential stage of translation, as it ensures the formation of full-sized proteins. Additionally, this partial sequence was lacking 5′-UTR and 3′-UTR, however, the present study sequence of Murrah buffalo *MBL1* CDS had 158 bp of 5′-UTR and 46 bp of 3′-UTR. In this way, the sequence in the present study filled this gap and characterized the complete CDS of the *MBL1* gene in Murrah buffalo. The *MBL1* CDS of Murrah buffalo was more similar to that of exotic Mediterranean buffalo (99.6%) than other (partial CDS) Murrah buffalo (99.1%) and American buffalo (98.1%) sequences. It showed more similarity with exotic cattle breeds (98.3-98.5%) and yak (98.4%) than exotic American buffalo (98.1%). No scientific report is available on the sequence similarity between cattle breeds, or between cattle and buffalo breeds. However, on comparison with the *MBL1* mRNA sequence of *Bos taurus* (NM_001010994) available in NCBI, it showed 98.3% homology at nucleotide level. It was 96.1% similar to goats, and 95.2% similar to sheep, indicating a high degree of sequence conservation. Further, it was 83.5% similar to pig breeds (Landrace and Yorkshire pig), that is, similar to the LILLIE et al. (2006a) report, who found 83.4% homology between porcine and bovine *MBL1 CDS*. The deduced amino acid sequence of the CDS of Murrah buffalo *MBL1* was found to have conserved sequences required for specific MBL functions, as shown in the results section. This finding was consistent with the kink observed in the collagen like domain in all other MBL proteins (WALLIS et al., 2004). Similar conserved sequences were reported in the deduced amino acid sequences of the previously characterized bovine *MBL1* gene (KAWAI et al., 1997; LILLIE et al., 2006a) and the porcine *MBL1* gene (AGAH et al., 2001; LILLIE et al., 2006a). According to the literature, at least two cysteine residues are needed in the N-terminal region for stabilization of trimers in higher order multimers (OHASHI and ERICKSON, 2004; JENSEN et al., 2005). A GEKGEP sequence in the collagen domain has been associated as a site for interaction with the C1q receptor C1qRp, and the PGKXGP sequence at the C-terminal end of the collagen-like domain is part of the putative MASP-binding motif (WALLIS et al., 2004) in the MBL of other species (ARORA et al., 2001). All the above regions are also present in currently characterized sequence, which suggests the Murrah buffalo MBLA protein has the potential to activate the lectin-complement pathway. It was stated that collagenous lectins with mannose-sugar-type binding specificity have a Glu–Pro–Asn (EPN) motif in their CRD, whereas those with galactose-sugar-type specificity, have Gln–Pro–Asp (WEIS et al., 1992; OGASAWARA et al., 1995). On this basis Murrah buffalo MBL-A was predicted to have mannose-type sugar specificity, similar to other mammalian MBLs. These previous findings about MBLs indicates that the Murrah buffalo MBL1 in the present study is a true mannose-binding lectin.

Murrah buffalo *MBL1* CDS had a several synonymous and non-synonymous substitutions. Likewise, YUAN et al. (2013) identified a synonymous mutation 2569T→C. WANG et al., (2011) and LIU et al. (2011) also identified a

synonymous mutation T→C in the N-terminal region of *Bos taurus MBL1*. Additionally, COSENZA et al. (2012) detected two synonymous mutations (g.993G→A) and g.1092C→T in the first coding exon of sheep *MBL*. However, several non-synonymous mutations in the first coding exon of the human *MBL* gene were frequently described as being associated with MBL plasma concentration, reduced ligand-binding capacity, and the failure to activate the complement system (LARSEN et al., 2004). Mutations in the N-terminal domain might interfere with oligomeric assembly and clustering, thus leading to alteration of the protein functions. Moreover, a non-synonymous mutation was detected G→A resulting in Val24Ile substitution in the N-terminal region of the bovine *MBL1* gene, and found to be significantly correlated with milk SCS and MBL-A serum levels, suggesting it might be useful as a marker for selection of dairy mastitis resistance in cattle (WANG et al., 2011; LIU et al., 2011; YUAN et al., 2013).

Further analysis revealed, non-synonymous SNPs in buffalo *MBL1* CDS including 611T→C and 538T→C located in the CRD region of the protein. A comparable mutation to leucine residue was identified by JUUL-MADSEN et al., 2006 in the porcine *MBL1* gene in CRD, which resulted in substitution of Leu233Ser, and they suggested it may affect the binding properties of CRD. Whether the CRD mutations found in the present study have similar effects on the structure of buffalo MBL-A is unknown, but this is of particular interest as evidence suggests that CRD specific mutations are associated with its binding properties and complement activation. LILLIE et al. (2006b) and JUUL-MADSEN et al. (2011a) identified one non-synonymous SNP 271G→T in porcine *MBL1* which was predicted to be functionally important because it substituted Gly91Cys in CLD. This triplet has been suggested to be involved in the initiation of the helical formation of the CLD, and the mutation may inhibit the assembly of MBL subunits into the basic trimer structure, thus reducing the amount of functional MBL in circulation. The phylogenetic tree revealed that all buffalo breeds exhibited a close phylogenetic evolutionary relationship with cattle and yak. The bovidae family divided into bovinae and caprinae subfamilies showing a greater degree of variation, leaving suidae and camelidae in two different subgroups. In contrast, LILLIE et al. (2006a) and PHATSARA et al. (2007) reported a phylogenetic analysis, and observed porcine and bovine MBL-A came off together as a separate branch, which might be because they were aligned with MBL1 of non-related species except bovine.

We screened three SNPs including g.855G>A, g.2686T>C and g.2561G>A in the buffalo *MBL1* gene. Our findings of genotypic and allelic frequencies for g.855G>A were similar to previous reports (LIU et al., 2010; WANG et al., 2011) in Chinese cattle breeds, but in contrast to the findings of ASAF et al., (2014) who reported only GG and GA genotypes in Vrindavani crossbred cattle. YUAN et al. (2013) and AKSEL et al. (2019) also reported an allelic variant with G→A mutation at location c.1252G>A in the intron I region of the bovine *MBL1* gene in Chinese, Turkish and European cattle breeds. However, intron sequences do not code for amino acids/proteins, and they have been proven to play vital regulating roles in gene expression and regulation (NOT et al., 2003), mRNA transcription and splicing (ZAN et al., 2007). There is no report available for this SNP in buffalo. Whether g.855G>A mutation affects gene expression and its regulation is unknown.

Among the two exonic SNPs, one is synonymous (g.2686T>C), with no amino acid change of Ala35Ala in the N-terminus region, and the other is non-synonymous (g.2651G>A), which caused substitution of Val24Ile in MBL-A in the N-terminus region. Both these SNPs were reported to present in the exon 2 region, which codes the signal peptide, N-terminus region and part of the CLD in bovine MBL-A protein (WANG et al., 2011). The screened Murrah buffalo population was found to be polymorphic for synonymous SNP g.2686T>C, and allele C was dominant. Similarly, previous studies (LIU et al., 2010; WANG et al., 2011; LIU et al., 2011; KAMALDEEP et al., 2017a and 2017b) also observed allele C as dominant in their study with a frequency of 0.49-0.74, while allele T was mutant with a frequency of 0.26-50.0 in different breeds of cattle. However, the pattern of genotypic frequencies in Murrah buffalo was

completely in contrast with the pattern in cattle breeds reported by other researchers, with 0.00-39.20 frequency for TT, 29.17-75.48 for TC, and 11.73-50.0 for CC, which might be due to the species differences. Further, the screened Murrah buffalo population was found monomorphic for g.2651G>A having a single genotype GG in all animals, whereas, several authors (WANG et al., 2011; LIU et al., 2011; ASAF et al., 2014a) found a polymorphic pattern for this SNP using different methods, such as CRS-PCR and allele specific PCR, respectively.

In the current study, the χ^2 indicated that screened Murrah buffalo animals were in Hardy-Weinberg equilibrium ($P<0.01$) for the SNPs (g.855G>A and for g.2686T>C). The association study of g.855G>A and g.2686T>C genotypes with milk production traits and SCS revealed no significant difference between genotypes, which is in agreement with the findings of other researchers (LIU et al., 2010; WANG et al., 2011; LIU et al., 2011; ASAF et al., 2014; KAMALDEEP et al., 2017a). On the other hand, interestingly, WANG et al. (2011) and LIU et al. (2011) found a significant difference between the combined genotypes of the three SNPs (g.855G>A, g.2651G>A and g.2686T>C) and milk production traits, suggesting that the genotype of one SNP may be influenced by the other SNPs. Moreover, YUAN et al., 2013 observed no significant correlation between c.1252G>A and c.2569T>C and milk SCS in the analyzed cattle populations. However, a significant correlation with milk SCS was detected in c.2534G>A. In the present study, the value of SCS for different genotypes showed very little difference, hence it cannot be predicted which genotype is favorable or susceptible to mastitis.

Conclusions

In the present study, we reported the cloning and characterization of the coding region of the *MBL* gene in the Indian buffalo breed Murrah, along with identification of two synonymous and four non-synonymous mutations in the *MBL* gene of buffalo breeds. The role of these SNPs found in *MBL1* CDS is yet to be analyzed as evidence suggests that *MBL*1 mutations have been associated with structural dysfunction, serum level, complement activation, milk production traits, SCS and microbial infections. The Murrah buffalo *MBL1* in the present study is a true MBL as it contains all the required functional conserved sequences of protein. Additionally, among the three screened *MBL1* SNPs, two (g.855G>A and g.2686T>C) polymorphic and one (g.2561G>A) monomorphic were found in Murrah buffalo *MBL1*. No significant different was observed between any of the SNPs and milk production traits or SCS. Since the present study was conducted with a relatively small sample size, the targeted locus needs to be explored further with a larger sample size to check the association with milk production traits and SCS.

References

AGAH, A., M. C. MONTALTO, K. YOUNG, G. L. STAHL (2001): Isolation, cloning and functional characterization of porcine mannose-binding lectin. Immunology 102, 338-343.
DOI: 10.1046/j.1365-2567.2001.01191.x

AKSEL, E. G., K. ARSLAN, F. ÖZDEMIR, B. AKYÜZ (2019): Investigation of *MBL-1* gene polymorphism in some cattle breeds raised in Turkey. Mediterr. Agri. Sci. 32, 25-30.
DOI: 10.29136/Mediterranean.457231

ALI, A. K. A., G. E. SHOOK (1980): An optimum transformation for somatic cell concentration in milk. J. Dairy Sci. 63, 487-490.
DOI: 10.3168/jds.S0022-0302(80)82959-6

ARORA, M., E. MUNOZ, A. J. TENNER (2001): Identification of a site on mannan-binding lectin critical for enhancement of phagocytosis. J. Biol. Chem. 276, 43087-43094.
DOI: 10.1074/jbc.M105455200

ASAF, V. N. M., K. AMOD, P. MNJIT, D. PRASHANT, G. K. GAUR, B. BHARAT (2014): Association of single nucleotide polymorphism in the MBL1 gene with mastitis in Vrindavani crossbred cattle. Indian J. Ani. Sci. 84, 703-705.

CAPPARELLI, R., M. PARLATO, M. G. AMOROSO, S. ROPERTO, R. MARABELLI, F. ROPERTO, D. IANNELLI (2008): Mannose-binding lectin haplotypes influence Brucella abortus infection in the water buffalo (Bubalus bubalis). Immunogenetics 60, 157-165.
DOI: 10.1007/s00251-008-0284-4

CARLEN, E., E. STRANDBERG, A. ROTH (2004): Genetic parameters for clinical mastitis, somatic cell score, and production in the first three lactations of Swedish Holstein cows. J. Dairy. Sci. 87, 3062-3070.
DOI: 10.3168/jds.S0022-0302(04)73439-6

COSENZA, G., A. PAUCIULLO, A. MANCUSI, A. D'AVINO, L. COLIMORO, D. DI- BERARDINO, L. RAMUNNO (2012): Identification of single nucleotide polymorphisms in the Nicastrese goat and Sardinia sheep mannose-binding lectin. Biochem. Genet. 50, 73-83.

FALCONER, D. S., T. F. C. MACKAY (1996): An Introduction to quantitative genetics. 4th ed., Songman Ltd. Esser, England.

GJERSTORFF, M., S. HANSEN, B. JENSEN, B. DUEHOLM, P. HORN, C. BENDIXEN, U. HOLMSKOV (2004): The gene encoding bovine SP-A, SP-D, MBL-A, conglutinin, CL–43 and CL–46 form a distinct collectin locus on *Bos taurus* chromosome 28 (BTA28) at position q.1.8–1.9. Anim. Genet.. 35, 333-337.
DOI: 10.1111/j.1365-2052.2004.01167.x

GUO, N., T. MOGUES, S. WEREMOWICZ, C. C. MORTON, K. N. SASTRY (1998): The human ortholog of rhesus mannose-binding protein-A gene is an expressed pseudogene that localizes to chromosome 10. Mamm. Genome. 9, 246-249.
DOI: 10.1007/s003359900735

HOLMSKOV, U., R. MALHOTRA, R. B. SIM, J. C. JENSENIUS (1994): Collectins: collagenous Ctype lectins of the innate immune defense system. Immunol. Today 15, 67-74.
DOI: 10.1016/0167-5699(94)90136-8

JACK, D. L., M. W. TURNER (2003): Anti-microbial activities of mannose-binding lectin. Biochem. Soc. Trans. 31, 753-757.
DOI: 10.1042/bst0310753

JENSEN, P. H., D. WEILGUNY, F. MATTHIESEN, K. A. MCGUIRE, L. SHI, P. HOJRUP. (2005): Characterization of the oligomer structure of recombinant human mannan-binding lectin. J. Biol. Chem. 280, 11043-11051.
DOI: 10.1074/jbc.M412472200

JUUL-MADSEN, H. R., R. M. KJAERUP, C. TOFT, M. HENRYON, P. M. H. HEEGAARD, P. BERG, T. S. DALGAARD (2011a): Structural gene variants in the porcine mannose-binding lectin 1 (MBL1) gene are associated with low serum MBL-A concentration. Immunogenetics 63, 309-317.
DOI: 10.1007/s00251-011-0512-1

JUUL-MADSEN, H. R., L. R. NORUP, P. H. JØRGENSEN, K. J. HANDBERG, E. WATTRANG, T. S. DALGAARD (2011b): Crosstalk between innate and adaptive immune responses to infectious bronchitis virus after vaccination and challenge of chickens varying in serum mannose-binding lectin concentrations. Vaccine 29, 9499-9507.
DOI: 10.1016/j.vaccine.2011.10.016

KAMALDEEP, D., A. MGOTRA, B. L. PANDER, A. MALIK, A. S. GARG, D. S. DALAL, B. S. MALIK (2017a): Association of G.2686T>C mutation of MBL1 gene with reproduction traits in Sahiwal cattle. Int. J. of Agri. Sci. and Res. 7, 159-164.

KAMALDEEP, D., A. MAGOTRA, B. L. PANDER, M. MANJEET, B. S. MALIK, D. S. DALAL (2017b): Identification of point mutation in Exon 2 of MBL1 gene and its relationship with Somatic Cell Score (SCS) in Sahiwal cattle. Indian J. Anim. Res. 53, 200-203.
DOI:10.18805/ijar.B-3483

KAWAI, T., Y. SUZUKI, S. EDA, K. OHTANI, T. KASE, Y. FUJINAGA, T. SAKAMOTO, T. KURIMURA, N. WAKAMIYA (1997): Cloning and characterization of a cDNA encoding bovine mannan-binding protein. Gene 186, 161-165.
DOI: 10.1016/s0378-1119(96)00664-6

KAWASAKI, N., T. KAWASAKI, I. YAMASHINA (1983): Isolation and characterization of a mannan-binding protein from human serum. J. Biochem. 94, 937-947.
DOI: 10.1093/oxfordjournals.jbchem.a134437

KOIVULA, M., E. A. MANTYSAARI, E. NEGUSSIE, T. SERENIUS (2005): Genetic and phenotypic relationships among milk yield and somatic cell count before and after clinical mastitis. J. Dairy Sci. 88, 827-833.
DOI: 10.3168/jds.S0022-0302(05)72747-8

LARSEN, F., H. O. MADSEN, R. B. SIM, C. KOCH, P. GARRED (2004): Disease associated mutations in human mannose-binding lectin compromise oligomerization and activity of the final protein. J. Biol. Chem. 279, 21302-21311.
DOI: 10.1074/jbc.M400520200

LAURSEN, S. B., O. L. NIELSEN (2000): Mannan-binding lectin (MBL) in chickens: Molecular and functional aspects. Dev. Comp. Immunol. 24, 85-101.
DOI: 10.1016/s0145-305x(99)00066-x

LILLIE, B. N., N. D. KEIRSTEAD, E. J. SQUIRES, M. A. HAYES (2007): Gene polymorphisms associated with reduced hepatic expression of porcine mannan-binding lectin C. Dev Comp. Immunol. 31, 830-846.
DOI: 10.1016/j.dci.2006.11.002

LILLIE, B. N., J. D. HAMMERMUELLER, J. I. MACINNES, M. JACQUES, M. A. HAYES (2006a): Porcine mannan-binding lectin A binds to Actinobacillus suis and Haemophilus parasuis. Dev. Comp. Immunol. 30, 954-965.
DOI: 10.1016/j.dci.2005.12.008

LILLIE, B. N., N. D. KEIRSTEAD, E. J. SQUIRES, M. A. HAYES (2006b): Single nucleotide polymorphisms in porcine mannan-binding lectin A. Immunogenetics 58, 983-993.
DOI: 10.1007/s00251-006-0160-z

LILLIE, B. N., A. S. BROOKS, N. D. KEIRSTEAD, M. A. HAYES (2005): Comparative genetics and innate immune functions of collagenous lectins in animals. Vet. Immunol. Immunopathol. 108, 97-110.
DOI: 10.1016/j.vetimm.2005.07.001

LIU, J., Z. JU, Q. LI, J. HUANG, R. LI, J. LI, L. MA, J. ZHONG, C. WANG (2011): Mannose-binding lectin haplotypes influence serum MBL-A concentration, complement activity, and milk production traits in Chinese Holstein cattle. Immunogenetics 63, 727-742.
DOI: 10.1007/s00251-011-0548-2

LIU, M., Z. H. JU, Q. L. LI, H. M. WANG, J. M. HUANG, J. B. LI, J. F. ZHONG, C. F. WANG (2010): Polymorphisms of MBL1 Gene Intron 1 and Exon 2 and Its Association with Mastitis and Milk Quality in Chinese Holstein Cattle. Scientia Agricultura Sinica 43, 2363-2371.
DOI: 10.3864/j.issn.0578-1752.2010.11.021

LIU, H., L. JENSEN, S. HANSEN, S. V. PETERSEN, K. TAKAHASHI, A. B. EZEKOWITZ, F. D. HANSEN, J. C. JENSENIUS, S. THIEL (2001): Characterization and quantification of mouse mannan-binding lectins (MBL-A and MBL-C) and study of acute phase responses. Sc. J. Immunol. 53, 489-497.
DOI: 10.1046/j.1365-3083.2001.00908.x

LOVELESS, R. W., T. FEIZI, R. A. CHILDS, T. MIZUOCHI, M. S. STOLL, R. G. OLDROYD, P. J. LACHMANN (1989): Bovine serum conglutinin is a lectin which binds non-reducing terminal N-acetylglucosamine, mannose and fucose residues. Biochem. J. 258, 109-113.
DOI: 10.1042/bj2580109

LUND, M. S., J. JENSEN, P. H. PETERSEN (1999): Estimation of genetic and phenotypic parameters for clinical mastitis, somatic cell production deviance, and protein yield in dairy cattle using Gibbs sampling. J. Dairy. Sci. 82, 1045-1051.
DOI: 10.3168/jds.S0022-0302(99)75325-7

MATSUSHITA, M., T. FUJITA (1992): Activation of the classical complement pathway by mannose-binding protein in association with a novel C1s-like serine protease. J. Exp. Med. 176, 1497-502.
DOI: 10.1084/jem.176.6.1497

NAITO, H., A. IKEDA, K. HASEGAWA, S. OKA, K. UEMURA, N. KAWASAKI, T. KAWASAKI (1999): Characterization of human serum mannanbinding protein promoter. J. Biochem. 126, 1004-1012.
DOI: 10.1093/oxfordjournals.jbchem.a022543

NETH, O., D. L. JACK, A. W. DODDS, H. HOLZEL, N. J. KLEIN, M. W. TURNER (2000): Mannose-binding lectin binds to a range of clinically relevant microorganisms and promotes complement deposition. Infect. Immun. 68, 688-693.
DOI: 10.1128/IAI.68.2.688-693.2000

NORUP, L. R., T. S. DALGAARD, N. C. FRIGGENS, P, SØRENSEN, H. R. JUUL-MADSEN (2009): Influence of chicken serum mannose-binding lectin levels on the immune response towards Escherichia coli. Poult. Sci. 88, 543-553.
DOI: 10.3382/ps.2008-00431

NOTT, A., S. H. MEISLIN, M. J. MOORE (2003): A quantitative analysis of intron effects on mammalian gene expression. RNA 9, 607–617.
DOI: 10.1261/rna.5250403

OGASAWARA, Y., D. R. VOELKER (1995): Altered carbohydrate recognition specificity engineered into surfactant protein D reveals different binding mechanisms for phosphatidylinositol and glucosylceramide. J. Biol. Chem. 270, 14725-14732.
DOI: 10.1074/jbc.270.24.14725

OHASHI, T., H. P. ERICKSON (2004): The disulfide bonding pattern in ficolin multimers. J Biol. Chem. 279, 6534-6539.
DOI: 10.1074/jbc.M310555200

PHATSARA, C., D. G. JENNEN, S. PONSUKSILI, E. MURANI, D. TESFAYE, K. SCHELLANDER, K. WIMMERS (2007): Molecular genetic analysis of porcine mannose-binding lectin genes, MBL1 and MBL2, and their association with complement activity. Int. J. Immunogenet. 34, 55-63.
DOI: 10.1111/j.1744-313X.2007.00656.x

RUPP, R., D. BOICHARD (1999): Genetic parameters for clinical mastitis, somatic cell score, production, udder type trait and milking ease. J. Dairy Sci. 82, 2198-2204.
DOI: 10.3168/jds.S0022-0302(99)75465-2

SASTRY, R., J. S. WANG, D. C. BROWN, R. A. EZEKOWITZ, A. I. TAUBER, K. N. SASTRY (1995): Characterization of murine mannose-binding protein genes Mbl1 and Mbl2 reveals features common to other collectin genes. Mamm. Genome 6, 103-110.
DOI: 10.1007/BF00303252

SEYFARTH, J., P. GARRED, H. O. MADSEN (2005): The 'involution' of mannose binding lectin. Hum. Mol. Genet. 14, 2859-2869.
DOI: 10.1093/hmg/ddi318

SHARMA, B. S., I. LEYVA, F. SCHENKEL, N. A. KARROW (2006): Association of toll-like receptor 4 polymorphisms with somatic cell score and lactation persistency in Holstein bulls. J. Dairy Sci. 89, 3626-3635.
DOI: 10.3168/jds.S0022-0302(06)72402-X

SHERIFF, S., C. Y. CHANG, R. A. EZEKOWITZ (1994): Human mannose- binding protein carbohydrate recognition domain trimerizes through a triple alpha-helical coiled-coil. Nat. Struct. Biol. 1, 789-794.
DOI: 10.1038/nsb1194-789

TAL-STEIN, R., L. FONTANESI, M. DOLEZAL, E. SCOTTI, A. BAGNATO, V. RUSSO, F. CANAVESI, A. FRIEDMANN, M. SOLLER, E. LIPKIN (2010): A genome scan for quantitative trait loci affecting milk somatic cells core in Israeli and Italian Holstein cows by means of selective DNA pooling with single- and multiple-marker mapping. J. Dairy Sci. 93, 4913-4927.
DOI: 10.3168/jds.2010-3254

TAMURA, K., J. DUDLEY, M. NEI, S. KUMAR (2007): MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596-1599.
DOI: 10.1093/molbev/msm092

THIEL, S., M. GADJEVA (2009): Humoral pattern recognition molecules: mannan-binding lectin and ficolins. Adv. Exp. Med. Biol. 653, 58-73.
DOI: 10.1007/978-1-4419-0901-5_5

THIEL, S., T. VORUP-JENSEN, C. M. STOVER, W. SCHWAEBLE, S. B. LAURSEN, K. POULSEN, A. C. WILLIS, P. EGGLETON, S. HANSEN, U. HOLMSKOV, K. B. REID, J. C. JENSENIUS (1997): A second serine protease associated with mannan-binding lectin that activates complement. Nature 386, 506-510.
DOI: 10.1038/386506a0

WALLIS, R (2007): Interactions between mannose-binding lectin and MASPs during complement activation by the lectin pathway. Immunobiology 212, 289-299.
DOI: 10.1016/j.imbio.2006.11.004

WALLIS, R., J. M. SHAW, J. UITDEHAAG, C. B. CHEN, D. TORGERSEN, K. DRICKAMER (2004): Localization of the serine protease-binding sites in the collagen-like domain of mannose-binding protein: indirect effects of naturally occurring mutations on protease binding and activation. J. Biol. Chem. 279, 14065-14073.
DOI: 10.1074/jbc.M400171200

WANG, C., M. LIU, Q. LI, Z. JU, J. HUANG, J. LI, H. WANG, J. ZHONG (2011): Three novel single-nucleotide polymorphisms of MBL1 gene in Chinese native cattle and their associations with milk performance traits. Vet. Immunol. Immunopathol. 139, 229-236.
DOI: 10.1016/j.vetimm.2010.10.023

WEIS, W. I., K. DRICKAMER, W. A. HENDRICKSON (1992): Structure of a C-type mannose-binding protein complexed with an oligosaccharide. Nature 360, 127-134.
DOI: 10.1038/360127a0

YUAN, Z., J. LI, J. LI, X. GAO, S. XU (2013): SNPs identification and its correlation analysis with milk somatic cell score in bovine MBL1 gene. Mol. Biol. Rep. 40, 7-12.
DOI: 10.1007/s11033-012-1934-z

YUAN, Z., G. CHU, Y. DAN, J. LI, L. ZHANG, X. GAO, H. GAO, S. XU, Z. LIU (2012): BRCA1: a new candidate gene for bovine mastitis and its association analysis between single nucleotide polymorphisms and milk somatic cell score. Mol. Biol. Rep. 39, 6625-6631.
DOI: 10.1007/s11033-012-1467-5

YUAN, Z. R., J. LI, L. LIU, L. P. ZHANG, L. M. ZHANG, C. CHEN, X. J. CHEN, X. GAO, J. Y. LI, J. B. CHEN, H. J. GAO, S. Z. XU (2011): Single nucleotide polymorphism of CACNA2D1 gene and its association with milk somatic cell score in cattle. Mol. Biol. Rep. 38, 5179-5183.
DOI: 10.1007/s11033-010-0667-0

ZAN, L.S., J. L. ZHANG, X. W. LIU (2007): Association study on AGPAT6 intron 3 polymorphism and milk performance of dairy cattle. Scientia Agricultura Sinica. 40, 1498-1503.

ZHAO, F., Z. ZHAO, G. YAN, D. WANG, Q. BAN, P. YU, W. ZHANG, Y. LUO (2011): Polymorphisms in mannose-binding lectin (MBL) gene and their association with MBL protein levels in serum in the Hu sheep. Vet. Immunol. Immunopathol. 140, 297-302.
DOI: 10.1016/j.vetimm.2010.12.009

Received: 19 November 2020
Accepted: 25 February 2023

BAGHEL, M., D. SHARMA, S. P. SINGH, V. PANDEY, A. KUMAR: Molekularna karakterizacija cjelovite kodirajuće sekvencije gena *MBL1* u indijskog vodenog bivola (*Bubalus bubalis*). Vet. Arhiv 93, 169-190, 2023.

SAŽETAK

Lektin koji veže manozu (MBL), član je porodice proteina kolektina koja veže veliki broj mikroorganizama i aktivira lektinski put komplementa prirođene imunosti. U različitim vrstama pronađene su brojne mutacije u kodirajućim i nekodirajućim regijama gena *MBL1*. Za nekoliko je mutacije utvrđeno da utječu na strukturu gena *MBL1*, što dovodi do snižavanja razine MBL-a i prirođenog poremećaja imunosnog sustava. U ovom je istraživanju provedeno molekularno kloniranje i karakterizacija cjelovitog slijeda gena *MBL1* u indijskog vodenog bivola pasmine Murrah. Fragment s ukupno 951 bp gena *MBL1* umnožen je kloniran i sekvenciran. Višestruko poravnanje sekvencija s drugim pasminama bivola i goveda otkrilo je da je kodirajuća sekvencija gena *MBL1* bila 98,1 – 99,6 % podudarna s drugim pasminama bivola i 98,3 – 98,5 % podudarna s drugim pasminama goveda na nukleotidnoj razini. Na razini aminokiselina ustanovljena je podudarnost od 96,8 do 98,8 % s pasminama bivola i 96,8 – 97,2 % s pasminama goveda. Sekvencija aminokiselina *MBL1* u vodenih bivola pasmine Murrah sadržavala je dvije nepodudarne supstitucije aminokiselina (L204P i S180P). Osim toga proveden je PCR-RFLP kako bi se 50 bivola pasmine Murrah analiziralo na prisutnost jednonukleotidnih polimorfizama (SNP), g. 855G > A u intronu I i g. 2686T > C, kao i g. 2651G > A, u eksonu 2 regije gena *MBL1*. Test ApaI/intron I PCR-RFLP otkrio je polimorfni obrazac s tri genotipa: AG (90 %), GG (8 %) i AA (2 %) s učestalošću alela od 0,94 za G i 0,06 za A. Test HaeIII/exon 2 PCR-RFLP pokazao je prisutnost triju genotipova: TC (66 %), TT (32 %) i CC (2 %) s učestalošću alela od 0,15 za T i 0,85 za C. Test StyI/exon 2 PCR-RFLP pokazao je monomorfni obrazac za g. 2651G > A, samo s genotipom GG. Osim navedenog istražena je i povezanost SNP-ova s proizvodnim svojstvima mlijeka i omjerom somatskih stanica (SCS). Nije pronađena znakovita povezanost ni za jedno svojstvo. S obzirom na to da se ovo istraživanje temelji na malom broju uzoraka, potrebna su istraživanja koja će uključiti veći broj životinja kako bi se provjerili učinci polimorfizama.

Ključne riječi: vodeni bivol (*Bubalus bubalis*); gen *MBL1*; kloniranje; karakterizacija; PCR-RFLP; genetski polimorfizam