SciDoc Publishers
Infer, Interpret & Inspire Science

International Journal of Microbiology & Advanced Immunology (IJMAI)
ISSN 2329-9967

OPEN ACCESS

# Partial characterization of immunoglobulin Cμ gene of water buffalo (Bubalus bubalis) predicts distinct structural features of C1q-binding site in Cμ3 domain

**Short Communication**

Surinder S. Saini[1], N. K. Maiti[2], Azad K. Kaushik[1*]

[1] Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada.
[2] Principal Scientist, Microbiology Section, Division of Fish Health Management, Central Institute of Freshwater Aquaculture, Bhubaneswar, Orissa, India

**Abstract**

The Partial characterization of immunoglobulin Cμ gene of water buffalo (Bubalus bubalis) has revealed unique features of the IgM isotype. Although, buffalo Cμ domain shares high amino acid sequence similarity with Cμ of cattle (94.28%) and sheep (91.71%), four distinct amino acid replacements (Met-301, Val-310, Asn-331 and Thr-432) spread across Cμ2, Cμ3 and Cμ4 diverge from those in cattle and sheep. As compared to cattle, buffalo Cμ gene has a codon deletion at position 507 (GTG encoding valine present in cattle) and insertion of GGC encodingglycine at position 532 in the Cμ4 domain. Unlike cattle and sheep, buffalo IgM has three potential N-linked glycosylation (Asn-X-Thr/Ser) sites, one at position 325-327 in the Cμ2 domain and two at positions 372-374 and 394-396 in the Cμ3 domain. Similar to cattle, buffalo IgM has fewer proline residues in Cμ2 domain that acts as hinge resulting in restricted flexibility of Fab arms. The increased structural flexibility in the C1q-binding site in Cμ3 may compensate for the rigid Cμ2 domain of buffalo IgM. The secondary structure of C1q binding site reveals its distinct features in buffalo and cattle IgM where a long alpha-helical structure is predominant which seems to be relevant to complement fixation function. The conserved protein motif 'Thr- Cys-Thr-Val-Ala-His' provides protein signatures of C1q binding region across ruminant species. The distinct structural features of C1q binding site of buffalo IgM, closest to cattle, are likely to be of functional significance relevant to designing antibody based therapeutics and diagnostics.

**Keywords:** Buffalo; Cμ Gene; Immunoglobulin; IgM; C1q-Binding Site

***Corresponding Author**:
Azad K. Kaushik,
Department of Molecular and Cellular Biology,
University of Guelph, Guelph,
Ontario, Canada.
E-mail: akaushik@uoguelph.ca

**Received:** May 10, 2013
**Accepted:** June 22, 2013
**Published:** June 26, 2013

**Citation:** Surinder S. Saini, N. K. Maiti, Azad K. Kaushik (2013) Partial characterization of immunoglobulin Cμ gene of water buffalo (Bubal-us bubalis) predicts distinct structural features of C1q-binding site in Cμ3 domain. Int J Microbiol Adv Immunol. 1(3), 19-23. **doi: http://dx.doi.org/10.19070/2329-9967-130004**



## Introduction

The water buffalo (Bubalus bubalis), member of fami-ly Bovidae domesticated approximately 5000 years ago in Asia, is raised for milk, meat and draught purposes.

Approximately, 170 million water buffaloes are main-ly found in Asia (97%) but their number is growing across Africa, Australia, Europe and South America [1]. India possesses the best of dairy breeds (Murrah, Nili-Ravi and Surti) that produce 72 million tones of milk annually, 5 % of world's total milk ouput [1,2]. Buffalo milk is rich in fat, protein and minerals but low in cholesterol [3] and is, thus, perfect source of good quality dairy products, especially the traditional Italian mozzarella cheese [4]. The demand for buffalo meat is high as it is relatively lean with low fat and high min-eral content as compared to beef or pork. Buffaloes provide an excellent source of draught power in more than 50 countries [1,2] (Borghese and Mazzi, 2005; Singh and Barwal, (2010). Buffalo utilizes poorly di-gestible feeds better than cattle and, therefore, can be maintained on low quality fodder and crop [1]. Impor-tantly, buffaloes are resistant to common diseases, ticks and external parasites that commonly afflict cattle[5].

Little is known about the structural and functional features of the immune system of this economically important species. The immunoglobulin genetics of other domestic species has been extensively studied [6] including humans and mice[7]. Limited sequence di-vergence is noted in phylogenetically close cattle [8-10] where somatic hypermutations [11,12] and generation of exceptionally long third complementarity determin-ing regions of heavy chains (CDR3H) [12-14] provide the required antibody diversity. Within the preponder-ant λ-light chain expression in cattle, a restricted Vλ1-Jλ3-Cλ3 recombination encodes the most λ-light chain 66 repertoire in cattle [15,16]. Immunoglobulin heavy chain constant regions genes that encode IgM, IgD, IgG, IgA and IgE isotypes, have been analyzed in many species 67 [6], including cattle [16-20]. The immuno-globulin gamma heavy chain gene has been mapped to buffalo chromosome 20q23-q25 by in-situ hybridi-zation [21]. Buffalo IgG, IgM and IgA immunoglobu-lin isotypes have been serologically characterized [22] where two sub-classes of buffalo IgG (IgG1 and IgG2) are identified[23]. To advance genetic and structural understanding about buffalo immunoglobulins, we partially characterized buffalo germline Cμ gene that encodes IgM, an immunoglobulin that appeared first during vertebrate evolu-

tion and is the fist to be expressed on developing B-lymphocytes. The buf-falo germline Cμ gene sequence from Niliravi breed shares high amino acid sequence similarity with cattle and, also, the predicted distinct C1q-binding structural characteristics.

## Materials And Methods

### Genomic DNA

Peripheral blood collected from a water buffalo of Nil-iravi breed, kept at the dairy farm of Punjab Agricul-tural University, Ludhiana, India, was used to extract genomic DNA as described[9].

### PCR And Sequencing

The buffalo germline Cμ gene, spanning codons 201 to 550, was PCR amplified using sense (5'-GTGT-GCGAAGTCCAGCA-3') and anti-sense (5'-AGAC-TAGTTACCGGTGGACTTGTCC-3') primers from conserved Cμ1 and Cμ4 exon sequences, respectively [17,24] under conditions that did not permit PCR ar-tifact[18]. The PCR steps involved a hot start at 95 0C for 2 min, denaturation at 95 0C for 1 min, annealing at 650C for 1 min and extension at 720C for 1 min up to a total of 30 cycles. The PCR conditions included 1.5 mM MgCl2, 0.8 μM of each primer and 2.5 U of Taq polymerase (Perkin-Elmer, Branchburg, NJ) in a 100 μl volume. The PCR product (~ 1.5 kb) was gel fraction-ated and purified using GeneClean II (Bio 101, Vista,CA) and subjected to automated DNA sequencing in both directions (MOBIX, McMaster University, Ham-ilton, ON, Canada). The internal sequencing primers were synthesized from the determined buffalo Cμ gene sequence (5'-TGAGGCCTCGGTCTGCT-3'), corresponding to codons 401 to 407. The buffalo Cμ codons are numbered according to[7] following Ou index [25]. The DNA sequence was analyzed using Geneious Pro 5.6.4 program (www.geneious.com) and the predicted protein secondary structure determined using the original Garnier Osguthorpe Robson algo-rithm (GOR I) provided by the EMBOSS suite [26].

## Results And Discussion

The nucleotide sequence and the deduced amino acid sequence of water buffalo germline Cμ gene, spanning between codons 201 and 550, is presented in Fig. 1.

The water buffalo Cμ gene shares a high nucleotide (95.52%) and amino acid (94.28%) sequence similarity with Cμ of cattle, closest ruminant species of family bovidae. Analysis of the buffalo germline Cμ gene se-quence revealed that it encodes part of Cμ1 domain (codons 201-221) and all of Cμ2 (codons 221- 333), Cμ3 (codons 334-438) and Cμ4 (codons 439-549) do-mains of IgM. When compared with other species, the overall amino acid identity of water buffalo IgM was most similar to sheep (91.71%) followed by pig (64.00%), rabbit (63.14%), human (61.71%), horse (60.57%) and mouse (56.28%). High amino acid se-quence similarity of buffalo IgM with cattle (94.28%) and sheep (91.71%) is expected given the close phylo-genetic relationship in ruminant species. Similar to cat-tle and sheep, buffalo IgM has unique amino acid sub-stitutions at 10 positions (Leu-239, Ser-246, Ile-274, Glu-279, Arg-303, Lys-319, Ser-367, Gly-370, Ala-421 and Lys-442) noted to be conserved in non-ruminant species (Fig. 2). Buffalo IgM has four distinct amino acid replacements (Met-301, Val-310, Asn-331 and Thr-432) spread across Cμ2, Cμ3 and Cμ4 that diverge from conserved amino acids in cattle and sheep IgM. As compared to cattle, buffalo Cμ gene has a codon deletion at position 507 (GTG encoding valine present in cattle) and insertion of GGC encoding glycine at position 532 in the Cμ4 domain (Fig. 1). Nucleotide deletions, insertions and substitutions are also noted in the intron sequences between the buffalo Cμ exons.

The conserved cysteines in buffalo IgM, essential for domain structure formation via intra-chain disulfide bridge, is noted at position 202 (Cμ1 domain which would interact with another cysteine residue within the Cμ1 domain; not investigated here); 252-313 (Cμ2 do-main), 360-418 (Cμ3 domain) and 466-528 (Cμ4 do-main; Fig. 2). Similarly, cysteine amino acids responsible for inter-chain disulfide bridges between the heavy chains of the monomeric (position 330) or polymeric (position 406) IgM [25] are conserved. Like most other species, buffalo IgM has two tryptophan residues in each of the Cμ2, Cμ3 and Cμ4 domains (Fig. 2). These findings are consistent with the critical role of conserved cysteine and tryptophan amino acids in main-taining the domain structure of immunoglobulin [27].

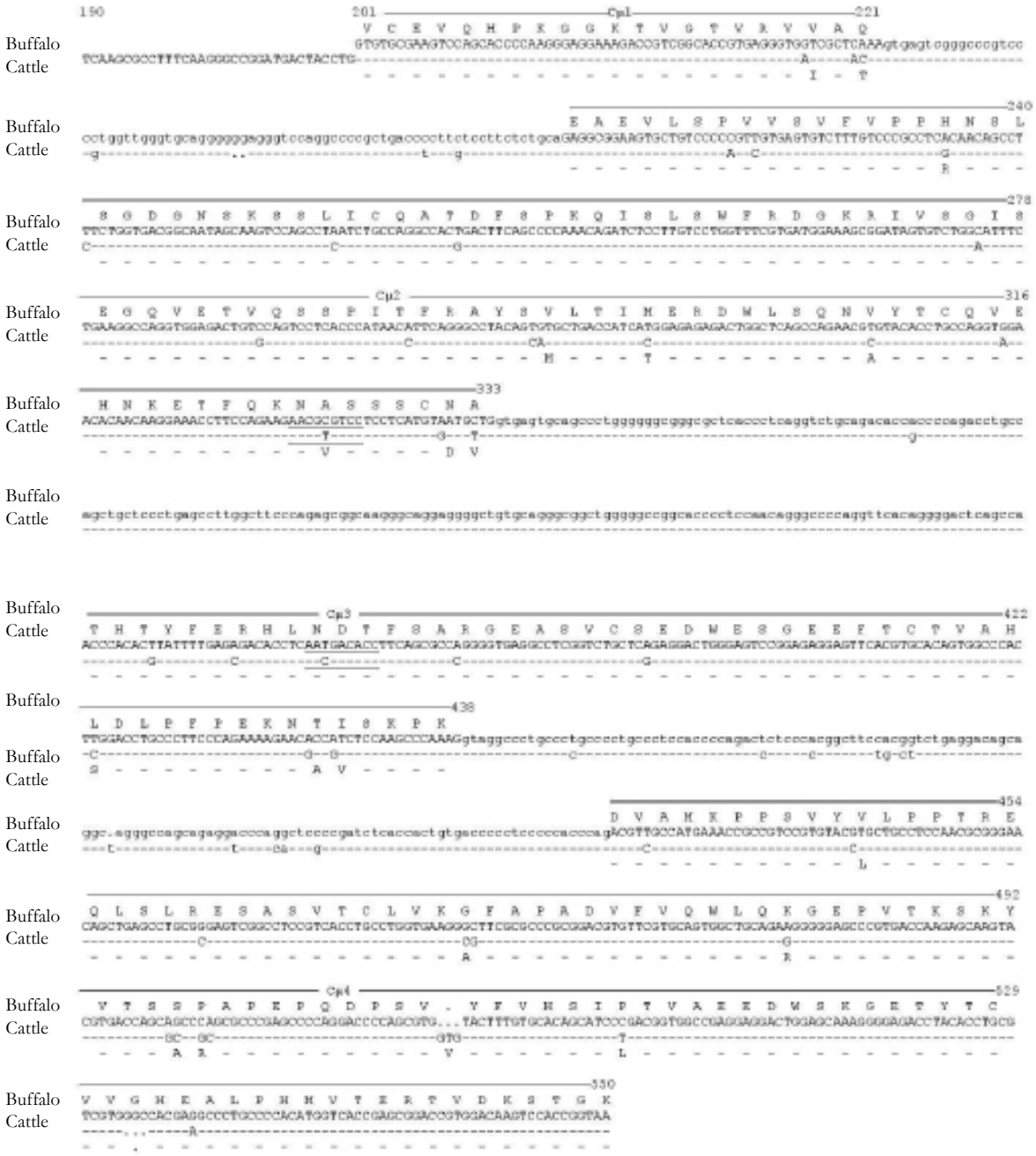
Buffalo IgM has three potential N-linked glycosylation (Asn-X-Thr/Ser) sites: one at position 325-327 in the Cμ2 domain and two at positions 372-374 and 394-396 in the Cμ3 domain (Fig.2). While other species have either one (cattle, sheep, pig, human) or two (rabbit, horse and mouse) N-linked glycosylation sites in Cμ1 domain. As for Cμ3 domain, there exist one (pig and mouse), two (cattle, sheep, human and horse) or three (rabbit) N-linked glycosylation sites (Fig.2). This sug-gests significant variability in the number of N-linked glycosylation sites in both Cμ2 and Cμ3 domains across species. Such differences in the N-linked glycosylation sites in the Cμ2 and Cμ3 domains may be of func-tional significance as these could influence functional configuration of IgM [28], especially movement of Fab arms or accessibility od C1q binding site. No N-linked glycosylation site exists in the buffalo Cμ4 do-main, a characteristic shared with other mammalian species but not with lower vertebrates [29-31].

Similar to cattle, sheep and goat, buffalo IgM has five prolines in the Cμ2 domain, the lowest number of prolines in this region that acts as hinge in contrast to other species, such as, pig and rabbit (7), human (8) and horse and mouse (9). The buffalo IgM has only six hydrophilic threonine amino acids in the Cμ2 domain, the lowest number in this region, unlike cattle (7), sheep (8), rabbit (9), pig and human (10), horse (12) and mouse (13). Similar to cattle (19) and sheep (20), buffalo IgM is rich in serine (19) in the Cμ2 domain.

Other species like mouse (9), human (12) and horse (13), have fewer serine residues in the Cμ2 domain, however. It seems that fewer hydrophilic threonine in Cμ2 domain of ruminant species is compensated by higher number of hydrophilic serine residues. Pres-ence of fewer proline in the Cμ2 domain will provide structural rigidity that may restrict segmental flexibility of Fab arms. Higher number of hydrophilic threonine and serine amino acid residues in Cμ2 of buffalo IgM is likely to augment its ability to extend into the sol-vent, however. We earlier reported similar findings for cattle IgM where structural constraints imposed by the restricted segmental flexibility of Fab arms is compen-sated by exceptionally long CRD3H (>50 amino ac-ids) region [158 13, 18]. It is possible that such a long CDR3H exists in buffalo antibodies as well.

The C1q-binding site buffalo IgM, spanning positions 408-428 in Cμ3 domain, has 12 conserved residues across species (positions 408-Glu, 409-Asp, 410-Trp, 411-Ser, 418-Cys, 419-Thr, 420-Val,

**Figure 1. The nucleotide and the deduced amino acid sequence of water buffalo germline Cμ gene from an animal of Niliravi breed. The exons (CH1, CH2, CH3 and CH4) are shown in block letters and the intervening intron sequences are indicated in small letters. The deduced amino acids are shown as single letter code above each codon. The Kabat numbering system [7] is used for the codons encoding the domains. The potential N-linked glycosylation sites are underlined. The cattle germline Cμ gene sequence [17] is included for comparison. The dash (-) indicates identical nucleotide or amino acid and the dot (.) indicates a lack of nucleotide or amino acid.**




422-His, 424-Asp, 425-Leu, 426-Pro and 428-Pro). Of the twelve con-served amino acids, three notable exceptions exist in humans (Glu replaced by Asp at position 408), mouse (Ser replaced by Asn at position 408) and goat (Trp replaced by Arg at position 409; Fig. 3). The conserved protein motif 'Thr-Cys-Thr-Val-Ala-His' provides protein signatures of C1q binding region in ruminant species. The predicted protein secondary structure of C1q binding site reveals its distinct structural features in buffalo and cattle IgM where a long alpha-helical structure is predominant (Fig. 3), unlike other species, followed by a short turn together with a coiled struc-ture common to all species. The C1q binding site in cattle and buffalo IgM also lacks beta strand altogether unlike other species. These structural features deviate from IgM of other ruminant species, like sheep and goats, where turns and/or coils are evident in this re-gion similar to other species. By contrast, the alpha-helical structure is altogether absent in C1q binding site of human IgM. These configurational differences in the conserved C1q binding region of IgM across species appear to be relevant to complement fixation and activation by classical pathway. It is possible that increased structural flexibility in the C1q-binding site compensates for the structurally rigid Cμ2 domain of buffalo and cattle IgM.

Overall, buffalo Cμ domain shares high amino acid sequence similarity with Cμ of other ruminant spe-cies like cattle and sheep. The buffalo IgM has fewer proline residues in the Cμ2 acting as hinge that would restrict the segmental flexibility of Fab arms. High hydrophilic threonine and serine amino acid content in Cμ2 domain will likely enhance its ability to extend into the solvent. The secondary protein structure of C1q binding site reveals its distinct structural features in buffalo and cattle IgM where a long alpha-helical structure is predominant (Fig. 3), unlike other spe-

**Figure 2. Comparison of amino acid sequence of water buffalo germline Cμ gene from Niliravi breed with predicted amino acid sequence of Cμ gene from cattle [17], sheep [24], goat (GenBank accession EU182621), pig [29], rabbit [32], human [33], horse [34] and mouse [35]. The dash (-) indicates identical nucleotide or amino acid and the dot (.) indicates a lack of nucleotide or amino acid. The gaps are inserted to optimize homology.**

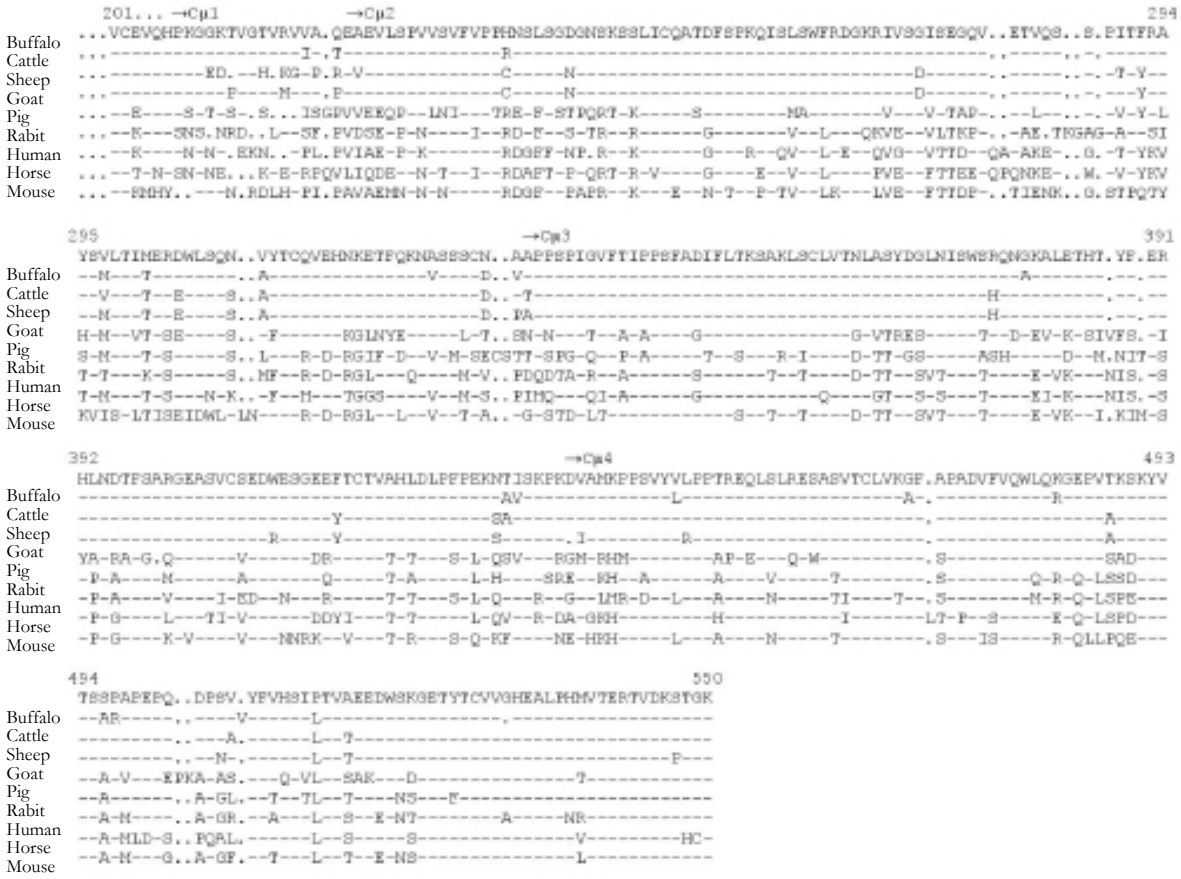
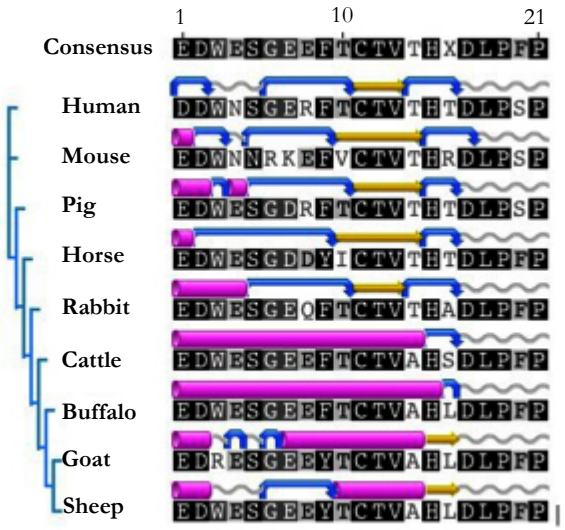



**Figure 3. Predicted tructural characteristics of C1q binding site of IgM across species using the original Garnier Osguthorpe Robson algorithm (GOR I) provided by the EMBOSS suite [26]. Alpha helix - ; Beta strand - ; Coil - ; Turn -**




cies which seems to be of functional significance.

## Acknowledgement

This research was supported by NSERC Canada Dis-covery grant to Dr. Azad K. Kaushik.

## References

[1]. Borghese A, Mazzi M. Buffalo population and strategies in the world. FAO REU Technical series 67, FAO Rome. 2005;Buffalo Production and Research:1-40.

[2]. Singh CV, Barwal RS. Buffalo breeding research and improve-ment strategies in India. 9th World buffalo Congress, Argentina. 2010:1024-31.

[3]. Tripaldi C. Buffalo milk quality. FAO REU Technical series 67, FAO Rome. 2005;Buffalo Production and Research:173-84.

[4]. Borghese A. Buffalo cheese and milk industry. . FAO REU Tech-nical series 67, FAO Rome. 2005;Buffalo Production and Research:185-96.

[5]. Fagiolo A, Roncoroni C, lai O, Borghese A. Buffalo Pathologies. FAO REU Technical series 67, FAO Rome. 2005;Buffalo Produc-tion and Research:249-98.

[6]. Butler JE. Immunoglobulin diversity, B-cell and antibody repertoire development in large farm animals. Rev Sci Tech. 1998 Apr;17:43-70.

[7]. Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C. Sequences of Proteins of Immunological Interest. 5th ed. Bethesda, MD: Pub-lic Health Ser-

vice, U. S. Department of Health and Human Ser-vices, National Institutes 205 of Health, 1991.

[8]. Berens SJ, Wylie DE, Lopez OJ. Use of a single VH family and long CDR3s in the variable 206 region of cattle Ig heavy chains. Int Im-munol. 1997 Jan;9:189-99.

[9]. Saini SS, Hein WR, Kaushik A. A single predominantly expressed polymorphic immunoglobulin VH gene family, related to mamma-lian group, I, clan, II, is identified in cattle. Mol Immunol. 1997 Jun;34:641-51.

[10]. Sinclair MC, Gilchrist J, Aitken R. Bovine IgG repertoire is domi-nated by a single diversified VH gene family. J Immunol. 1997 Oct 15;159:3883-9.

[11]. Kaushik AK, Kehrli ME, Jr., Kurtz A et al. Somatic hypermutations and isotype restricted exceptionally long CDR3H contribute to an-tibody diversification in cattle. Vet Immunol Immunopathol. 2009 Jan 15;127:106-13.

[12]. Koti M, Kataeva G, Kaushik AK. Novel atypical nucleotide inser-tions specifically at VH-DH junction generate exceptionally long CDR3H in cattle antibodies. Mol Immunol. 2010 Jul;47:2119-28.

[13]. Saini SS, Allore B, Jacobs RM, Kaushik A. Exceptionally long CDR3H region with multiple cysteine residues in functional bovine IgM antibodies. Eur J Immunol. 1999 Aug;29:2420-6.

[14]. Shojaei F, Saini SS, Kaushik AK. Unusually long germline DH genes contribute to large sized CDR3H in bovine antibodies. Mol Immu-nol. 2003 Sep;40:61-7.

[15]. Saini SS, Farrugia W, Ramsland PA, Kaushik AK. Bovine IgM anti-bodies with exceptionally long complementarity-determining region of the heavy chain share unique structural properties conferring restricted V222 H+V-lambda pairings. International Immunology. 2003 Jul;15:845-53.

[16]. Pasman Y, Saini SS, Smith E, Kaushik AK. Organization and genomic complexity of bovine lambda-light chain gene locus. Vet Immunol Immunopathol. 2010 Jun 15;135:306-13.

[17]. Mousavi M, Rabbani H, Pilstrom L, Hammarstrom L. Characteriza-tion of the gene for the membrane and secretory form of the IgM heavy-chain constant region gene (C mu) of the cow (Bos taurus). Immunology. 1998 Apr;93:581-8.

[18]. Saini SS, Kaushik A. Origin of bovine IgM structural variants. Mol Immunol. 2001 Sep;38:389-96.

[19]. Zhao Y, Kacskovics I, Pan Q et al. Artiodactyl IgD: the missing link. J Immunol. 2002 Oct 15;169:4408-230 16.

[20]. Xu B, Wang J, Zhang M et al. Expressional Analysis of Immuno-globulin D in Cattle (Bos taurus), a Large Domesticated Ungulate. PLoS One. 2012;7:e44719.

[21]. Hassanane MS, Gu F, Chowdhary BP, Andersson L, Gustavsson I. In situ hybridization 233 mapping of the immunoglobulin gamma heavy chain (IGHG) gene to chromosome 20q23-q25 in river buf-faloes. Hereditas. 1993;118:285-8.

[22]. Satija KC, Rajpal S, Pandey R, Sharma VK. Electrophoresis of buf-falo (bos bubalis) serum proteins including immunoglobulins. In-fect Immun. 1979 May;24:567-70.

[23]. Svasti MR, Prawatmuang P, Vajanamarhutue C, Kajadphai A, Wangthammang ST, N. The presence of two IgG subclasses in wa-ter buffalo immunoglobulins. J Sci Soc Thailand. 1976;2:56-66.

[24]. Patri S, Nau F. Isolation and sequence of a cDNA coding for the immunoglobulin mu chain of the sheep. Mol Immunol. 1992 Jul-Aug;29:829-36.

[25]. Putnam FW, Florent G, Paul C, Shinoda T, Shimizu A. Complete amino acid sequence of the Mu heavy chain of a human IgM im-munoglobulin. Science. 1973 Oct 19;182:287-91.

[26]. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 2000 Jun;16:276-7.

[27]. Kehry M, Sibley C, Fuhrman J, Schilling J, Hood LE. Amino acid sequence of a mouse immunoglobulin muchain. Proc Natl Acad Sci U S A. 1979 Jun;76:2932-6.

[28]. Jefferis R, Lund J. Glycosylation of antibody molecules: structural and functional significance. ChemImmunol. 1997;65:111-28.

[29]. Bosch BL, Beaman KD, Kim YB. Characterization of a cDNA clone encoding for a porcine immunoglobulin mu chain. Dev Comp Immunol. 1992 Jul-Aug;16:329-37.

[30]. Dahan A, Reynaud CA, Weill JC. Nucleotide sequence of the con-stant region of a chicken mu heavy chain immunoglobulin mRNA. Nucleic Acids Res. 1983 Aug 25;11:5381-9.

[31]. Ghaffari SH, Lobb CJ. Cloning and sequence analysis of channel catfish heavy chain cDNA indicate phylogenetic diversity within the IgM immunoglobulin family. J Immunol. 1989 Feb 15;142:1356-65.

[32]. Bernstein KE, Alexander CB, Reddy EP, Mage RG. Complete se-quence of a cloned cDNA encoding rabbit secreted mu-chain of VHa2 allotype: comparisons with VHa1 and membrane mu se-quences. J Immunol. 1984 Jan;132:490-5.

[33]. Harindranath N, Donadel G, Sigounas G, Notkins AL. Comparison of complete nucleotide sequence of the human IgM heavy chain constant region of polyreactive and monoreactive antibodies. Mol Immunol. 1993 Jan;30:111-2.

[34]. Schrenzel MD, King DP, McKnight ML, Ferrick DA. Characteriza-tion of horse (Equus caballus) immunoglobulin mu chain-encoding genes. Immunogenetics. 1997;45:386-93.

[35]. Kawakami T, Takahashi N, Honjo T. Complete nucleotide se-quence of mouse immunoglobulin mu gene and comparison with other immunoglobulin heavy chain genes. Nucleic Acids Res. 1980 Sep 11;8:3933-45.