

Mutations in CD14 gene causes mastitis in different breeds of buffalo as confirmed by in silico studies and experimental validation

Aruna Pal (✉ arunachatterjee@gmail.com)

West Bengal University of Animal and Fishery Sciences <https://orcid.org/0000-0001-5681-2065>

Abantika Pal

Indian Institute of Technology Kharagpur

Arjava Sharma

National Bureau of Animal Genetic Resources

Tarun Kumar Bhattacharya

Directorate of Poultry Research

Research article

Keywords: CD14, cytokine, bio-informatics, mutation, SNP, deleterious, I-mutant, Provean, mastitis

Posted Date: July 1st, 2019

DOI: <https://doi.org/10.21203/rs.2.10779/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background CD 14 is an important pattern recognition receptor having innate immune function and has antibacterial activity. It binds with LPS of gram-negative bacteria, arachidonic acid, and lipoteichoic acid. Being a receptor, it binds with the pathogen with the help of other cytokines. Mutations in CD14 affect the binding ability which in turn affects the biological potentiality. Method The present study was conducted on 228 nos. of buffaloes pertaining to four different breeds as Murrah, Mehsana, Surti and Bhadawari. CD14 gene was characterized and polymorphism was detected through Single nucleotide conformation polymorphism. Association study was conducted for different variants of CD14 with mastitis in buffalo, detected through somatic cell count, california mastitis test. Result Eight variants of CD14 were detected and mutational hotspots were detected in bubaline CD14 with 58 number of non-synonymous SNP, out of which 18 were observed to be deleterious and 34 as thermodynamically unstable. In the present study, we had detected the mutations in CD14 gene and its association with the somatic cell score and other indicators for mastitis. In-silico studies were conducted to understand the molecular mechanism how the mutations affect the biological potentiality by analyzing different domains and structural analysis along with various post-translational modification sites. Conclusion Deleterious mutations were observed in CD14 gene which have significant effect on mastitis of buffalo. It may be employed for marker assisted selection, therapeutic application of recombinant CD14, gene therapy, transgenic animal production with wild type CD14 resistant to mastitis as future strategy.

Background

Mastitis is an important economic disease of livestock. It is also considered as one of the most costly production related disease in a dairy industry. (Oviedo-Boyso [et.al.2007](#)). Economic losses due to mastitis may reach upto 35 million dollars per year world wide(Giraudo et al,1997; Pereira et al.2011). Subclinical mastitis is of greater concern than clinical mastitis, since incidences were 15 to 40 times more than that of clinical mastitis and it forms the basis of herd outbreaks (Boichard et al 2003; Detilleux et al, 2012). As no gross abnormality in milk and udder is noticed, subclinical mastitis goes unnoticed to the farmers. So, subclinical mastitis is considered more important due to a negative impact on the economy throughout the world. The figure for subclinical mastitis cases are approx.18.40% (Joshi and Gokhle 2006).The prevalence of subclinical mastitis with respect to different etiological agents have been identified (Ayano et al., 2013).

Economic losses due to mastitis may be due to depressed milk production, deterioration of milk quality, cost due to health management as treatment and vaccination, milk disposal due to antibiotic residues, advanced culling due to permanent damage to udder, addition labour requirement. There are certain other limitations as evolvment of antibiotic resistant bacteria, vaccinations are not cent percent due to evolvment of multistain organism, due to frequent mutation , by genetic shift and genetic drift. Other constraint are milk disposal due to its residual effect of antibiotic on the end consumers, worldwide concern for organic food production and ethical considerations of sufferings of animals has limited the antibiotic therapy. Study has been attempted with bovine mastitis prevention by inoculation of lactic acid

bacteria at the dry off period ([Pellegrino et al., 2017](#)). Thus host immune response and its immunomodulation may be considered as an alternative for disease resistance.

Detection of bacterial lipopolysaccharide is mediated by TLR4 and MD2 with CD 14 as co-receptor ([Kitchens et al., 2000](#)), only in the presence of LBP (lipopolysaccharide-binding protein) ([Muta et al., 2001](#)). CD14 also binds with lipoteichoic acid. ([Ranoa et al., 2013](#)). CD14 is expressed mainly by [macrophages](#) , [neutrophils](#) and [dendritic cells](#) ([Funda et al., 2001](#)).

In case of mastitis, upon the entrance of bacteria, there is neutrophil infiltration in the mammary gland . Cell surface receptors as CD14 and TLRs stimulate cytokines and other immune effector molecules. LPS signalling is a crucial step mediated by LR4, CD14 and MyD2 in presence of LPB, leading to release of Cytokine ([Schepper et al., 2008](#)). The concentration of sCD14 is highest within a few day of parturition and significantly increases parallel with increase in SCC, challenged with *E.coli* LPS. ([Lee et al., 2003](#)). It is reported that Buffaloes is more resistant than other livestock species and have least number of LRR i.e, leucine rich repeat, it has been observed that lesser no of LRR are less susceptible to disease ([Pal et al., 2010](#)).

Manipulation of CD14 gene may be conducted in disease management in various ways, which aid in decreasing disease incidences. The future use includes use as a therapeutic agent; somatic gene therapy and transgenic disease resistant animal production.

The present investigation was aimed for cloning and sequencing of the CD14 gene of buffalo and identification of the mutant variants. The current study aims to identify the molecular mechanism how the mutations in CD14 causes diseases as mastitis.

Methods

Animals :

Four different breeds of riverine buffalo (n=228), Bhadawari (n=44), Mehsana (n=71), Murrah (n=57) and Surti (n=56) of age 4-5 years of age were studied. Bhadawari buffalo samples were collected from Govt. Livestock Farm, Etawah (U.P.), Mehsana buffalo samples were collected from Gujrat Agricultural University, S. K. Nagar (Gujrat), Murrah samples were collected from LPM Section, IVRI (Indian Veterinary Research Institute), Izatnagar (U.P.), whereas Surti samples were collected from Gujrat Agricultural University, Navsari, Anand (Gujrat).

This study was carried out in accordance with the recommendations of IAEC (Institute Animal Ethics Committee, Indian Veterinary Research Institute). The protocol was approved by the IAEC, IVRI.

DNA preparations:

10 ml blood was collected from each animal in aseptic condition and genomic DNA was isolated from blood samples by phenol-chloroform extraction method as described by Sambrook and Russel (1989)

with few modifications. After the whole process, the DNA was taken in an Eppendorf tube and has been dissolved in TE buffer.

DNA amplification:

All PCR amplifications were performed in 25 µl reaction volume. Each reaction contained 3.0 µL 10X PCR assay buffer, 0.5 µL of 10 mM dNTP, 60 ng of each primer, 0.5 units of Taq DNA polymerase and nuclease-free water to bring the total volume to 25 µl. Around 100 ng of cDNA was used as the template. The details of primers for different fragments of GH gene are Forward: CD14-F ATGGTCTGCGTGCCCTACCTG and Reverse CD14-R GGAGCCCGAGGCTTCGCGTAA. Thermal cycling was performed on an ABI PCR machine initial denaturation at 94°C for 3 min, denaturation at 94°C for 30 sec, annealing at 61°C for 35 sec, and extension at 72°C for 3 min were carried out for 35 cycles followed by final extension at 72°C for 10 min.

Gene sequencing and sequence analysis:

Automated sequencer was used for DNA sequencing by dideoxy chain termination method of CD14 amplified fragment. The nucleotide sequence obtained was analyzed with different Bioinformatics tools and software (I-mutant, PROVEAN, SIFT, and others for analysis of post-translational modification). Three-dimensional modeling and model quality assessment was carried out with Pymol software. An analysis of several post-transcriptional modifications in the novel sequences were observed.

SNP detection of CD14 gene in riverine buffalo

Molecular sequence analysis of CD14 gene through sequence alignment had lead to identification of Single nucleotide polymorphism (SNPs). Alignment was conducted through MAFFT (Kato and Standley, 2016) and MegAlign Programme of DNASTAR package. Non-synonymous SNPs present in the protein coding region were analyzed for their thermodynamic stability and deleterious nature through I-Mutant3.0 (<http://folding.biofold.org/cgi-bin/i-mutant3.0>) and PROVEAN (Protein Variation Effect Analyzer, <http://provean.jcvi.org/index.php>) softwares respectively.

The novel sequences were submitted to the National Centre for Biotechnology Information Genbank.

Three dimensional structure prediction and Model quality assessment

The PDB structures of different variants of CD14 was developed through homology modeling using PHYRE2 server (Kelley et al , 2015). The three dimensional structures of both wild and mutant variant were visualized by PyMOL (<http://www.pymol.org/>). Different moieties in mutant variant and domain for CD14 were also visualized through Pymol.

Protein-protein interaction network depiction

Proteins cannot function alone, rather works through a biochemical pathway. Hence it is important to understand the proteins which are interacting for an functional outcome. Protein interaction was

estimated through STRING 9.1. Interactions with score < 0.3 are considered as low confidence, scores ranging from 0.3 to 0.7 are classified as medium confidence and scores > 0.7 yield high confidence. The functional partners were depicted.

Prediction of post-translational modification sites in mutant CD14

In-silico analysis were employed for the detection of post translational modification sites of bubaline CD14. Protein dynamics and functioning greatly depends on these PTM. The protein phosphorylation was analyzed using a tool, specifically, Ser, Thr and Tyr residues which required for catalyzing its role. The prediction of glycation site was carried out with NetGlycate 1.0 server (<http://www.cbs.dtu.dk/services/NetGlycate/>). The serine (Ser), threonine (Thr) and tyrosine (Tyr) residues with a score of >0.5 were depicted as to be glycated. NetPhos 2.0 server (<http://www.cbs.dtu.dk/services/NetPhos/>) was used for detection of protein phosphorylation sites.

Analysis of samples for clinical incidences for mastitis

The symptoms for clinical mastitis were udder inflammation, with the usual signs of heat, pain, redness and swelling, color alteration, and abnormal milk with flakes. The presence or absence of clinical mastitis was established by estimation of somatic cell count (SCC) and the California mastitis test (CMT).

Milk samples were collected aseptically in strip cup in Right fore (RF), Left fore (LF), Right hind (RH) and Left hind (LH) chambers for the detection of sub-clinical mastitis.

Estimation of California mastitis test

California mastitis test is also an important test for the detection of mastitis following the protocol as described by Dhakal (2006). Milk is collected in plastic paddle with four chambers. Equal amount of mastitis reagent was used in each cup and mixed gently. A score of 1 or more was considered as positive. The mastitis reagent was prepared initially by suspension of sodium lauryl sulphate (3g) in 100 ml of warm distilled water at pH 8.0. The reagent was finally prepared by mixing with bromocresol purple (1: 10000)

Estimation of rennet coagulation time

The rennet coagulation time (RCT) of the milk was estimated with 0.2% Rennet solution ([Bittante, 2011](#)).

Estimation of somatic cell count

Somatic cells in milk is a positive indicator for mastitis. 0.01 ml of milk was spreaded evenly over the slide. Staining of the slides were conducted with Newman Lampert stain (containing 1.2 g methylene blue, 54 ml 95% ethyl alcohol, 40 ml tetrachloroethane, 6 ml glacial acetic acid). Microscopic view of the slides reveal deep blue nuclei against light blue background. The total number of cells were counted to be total number of cells in 25 fields x working factor (Guha et al., 2012).

Statistical analysis

Identified SNPs were associated with the phenotypic traits in terms of mastitis specific marker as Somatic cell count. The effect of genotype on mastitis related traits were estimated through ANOVA. The model used for analysis was:

$$Y_{ijkl} = \mu + G_i + A_j + F_k + B_l + e_{ijklm}$$

where Y_{ijkl} = lth observation of the target trait, μ = Overall mean,

G_i = Fixed effect of ith genotype, A_j = Fixed effect of jth animals,

F_k = Fixed effect of kth farm, B_l = Breed effect, e_{ijklm} = Random error

The observations were collected in same season, hence seasonal effects were not considered. Suitable data transformations were utilized for data in scores. DNASTAR software package was employed for BLAST analysis (MegAlign Programme of Lasergene Software) and prediction of peptide sequence (Edit sequence Programme of Lasergene Software). Translate tool from Expasy software package was employed for confirmation of peptide sequence.

t-test and Duncan multiple range test (DMRT) were employed for the analysis. Computer software "SPSS (1999)" was employed for the analyses (Petrie and Watson, 2008).

3. Result

3.1. Characterization of CD 14 gene of buffalo

The buffalo CD14 coding sequence was of 1122 bp size with 62.3% GC content. The open reading frame of CD14 gene was of 1116 nucleotide, with a natural start codon and TAA stop codon. The total length of Bubaline CD14 molecule was observed to be 373 amino acids with signal peptide of 20 amino acid. Cattle CD14 also contains 373 amino acids and revealed 98% identity with buffalo CD14. The comparative analysis of derived amino acid sequence between cattle and buffalo revealed slightly higher molecular weight (39705.07 Da) of buffalo sequence as compared to cattle (39679.96 Da). The buffalo CD14 peptide sequence was characterized by one extra strongly basic amino acid and had two polar amino acids less than that of cattle. The nucleotide sequence comparison of cattle with buffalo revealed 22 nucleotide substitutions, out of which eleven were synonymous codon without any amino acid change. The peptide sequence comparison of buffalo CD14 with that of cattle revealed eleven amino acid substitution at 14, 62, 131, 134, 139, 143, 154, 209, 235-236, 277, 337th position. In the derived peptide sequence of CD14 of buffalo, six leucine-rich repeats (LRR) were observed (Fig 1), whereas a cattle has ten LRR. Six LRR have been identified and depicted in different colours (Fig 1), amino acid site 114-138 as wheat, 168-176 as pale green, 192-218 as light blue, 244-256 (pale yellow), 266-283 (light pink), 296-309 (blue-white). Buffalo CD14 molecule was predicted to contain 17.2% leucine, which is similar to mouse (17.66%) and higher than that of human (15.5%). Analysis of derived peptide sequence

revealed 373 amino acids which is GPI anchored at C-terminus near 353 amino acid position, about 7 amino acid after hydrophobic tail. Four putative N-linked glycosylation sites were observed in bubaline CD14.

3.2. Comparison of CD14 with related genes.

Comparison of CD14 gene with other related genes containing leucine-rich repeats, viz. MD2, TLR2, TLR4, and TLR9 of cattle revealed very low similarity of 22.4, 21.8, 18.8 and 23.4%, respectively. This reveals the distinct nature of CD14 gene from that of related gene sequences.

The molecular interaction of CD14 gene with the related genes are depicted in Fig2. It is evident from the figure that CD14 is related to TLR2, TLR4, TLR6, IL10, TNF, MYD88, LY96 by String analysis (Fig 2). The result was confirmed through analysis of biochemical pathway that these genes are related in TLR signalling pathway through KEGG analysis (Fig 3). Moreover, it is evident that each molecule responsible for innate immunity, either receptor or immune mediator acts through a cascade of mechanisms with interaction with other.

3.3. Identification of CD14 variants with SNPs

SSCP followed by sequencing has identified eight variants of CD14 gene. cDNA of CD14 gene was amplified for different variants from Variant A to H and sequenced. Sequences were designated as CD14 A (MF 092866), CD14 B (MF 092866), CD14 – C (MF 092866), CD14-D (MF 092866), CD14-E (MF 092866), CD14-F (MF 092866), CD14-G (MF 092866) and CD14-H (MF 092866). Variant A of CD14 gene was observed to be the wild type (Fig1 and Fig4). The allelic frequencies for different variants of CD14 as CD14-A, CD14-B, CD14-C, CD14-D, CD14-E, CD14-F, CD14-G and CD14-H were observed to be 0.47, 0.09, 0.08, 0.18, 0.022, 0.13, 0.018 and 0.018 respectively

The 3D structure of Variant A has been depicted in Fig 2 with disulfide bond as a blue colored sphere, LPS binding site as green, leucine zipper as yellow.

Mutations in CD14 affect the binding ability which in turn affects the biological potentiality. Mutational hotspots were detected in bubaline CD14 with 58 number of non-synonymous SNP, out of which 18 were observed to be deleterious and 34 as thermodynamically unstable. Eight different variants of CD14 gene have been identified in all the four breeds of buffalo with 58 SNPs, revealing a high degree of polymorphism. AA (genotypic frequency 0.468) as most frequent and AG and AH as the least frequent genotypes were identified as, with very low frequencies (0.0174) for both the genotypes. High degree of variability was observed for alleles B (81.3%), C (97.0%), D (86.0%), E (97.4%) and F (85.8%) compared to A allele of amplified CD14 nucleotide of buffalo, which may be due to coding for leucine-rich repeats, that confers the recognition and binding ability for a wide range of pathogens. However, an absence of any change of amino acid at the sites of glycosylation confirms that the vital functions of the CD14 molecule remained unaltered.

Breed wise differences for allelic frequencies for CD14 gene were observed.

None of the breed of buffalo under current study was observed to be resistant to mastitis. However the differences in susceptibility was observed for Breeds of buffalo. Among the four breeds under consideration, Surti was observed to be most susceptible to mastitis, while Bhadawari was observed to be comparatively better resistant to mastitis under current study. The resistance of buffalo breed as Mehsana and Murrah was observed to be intermediate. Surti was observed to be mostly susceptible to mastitis.

Comparisons were attempted for identification of mutant type of CD14 variant with the wild type.

The various functional domains of CD14 were studied with respect to wild type variant A as follows:

Variant A versus Variant B

A number of SNPs were identified out of which were observed to be synonymous and non-synonymous (Table 1). G210 C was observed to be deleterious mutation by Provean software, leading to altered function. I-mutant software depicts decreased stability at amino acid position G207A, T209A, G210C, 218, 230, 233, 234.

3D structure of variant A and Variant B have been depicted in fig 4 and fig 5. Disulphide bridges in blue, Alpha helix as pink and Beta pleated sheet as red, coil in cyan. The site for 209 as orange, 210 as green, 218 purple-blue.

Variant A versus variant C

Three non-synonymous SNPs were identified out of which T235 K was observed to be deleterious by Provean. I-mutant has predicted Y225C and T235K with reduced thermodynamic stability (Table 1). The 3D structure for CD14 variant C (Fig 6), with the mutations, have been depicted in figure 4. The secondary structure of CD 14 has been depicted in an alpha helix as cyan, beta sheet as red, loop in pink. Disulphide bonds in blue, Y225C in green and E243D and T235K as yellow.

Variant A versus Variant D

24 synonymous mutations were observed out of which 12 were observed to be deleterious (L215P, L253R, D254L, L255H, S256N, N258C, S259P, T263C, P273A, R277M, N280D, S282A, F283L) as depicted in Table 1. This high degree of mutations has an effect on its structure also as revealed by TM-align software. Structural alignment of CD14A with CD14 D has revealed a greater degree of variation, RMSD was observed to be 1.87. The structural alignment with the differing domain of CD14 had been highlighted in Fig 7.

Sites for disulphide bond has been depicted as blue coloured sphere. Site for mutation L215P as red. L253R, D254L, L255H, S256N, N258C, S259P as yellow.

T263C Orange, P273A Cyan, R277M Grey, N280D Warm pink, S282A Purple blue.

The domain of CD14 present at site L215P is leucine-rich repeat. It has already been reported that LRR in extracellular domain is responsible for the recognition of pathogens, hence any mutation at this site may lead to defects in coded protein causing impairment of the function of pathogen recognition.

The mutation of CD14 at site 253-256 codes for leucine-rich repeat, may cause a defect in pathogen recognition and the ultimate function may be impaired. It is to be kept in mind that CD14 is a pattern recognition receptor containing leucine-rich repeats (LRR). It is to be taken into consideration that whenever there is any mutation causing alteration of leucine moiety, the function of LRR is effected. The mutations at P273A, R277M, N280D, S282A, F283L site of CD14 gene of buffalo cause alteration of LRR domain.

The 3D structure depicting 3 D structural alignment of Variant A and variant D (Fig 7).

Variant A versus Variant E

The mutations identified in CD14 E are listed in Table 1. Seven synonymous mutations were observed out of which K296 L was predicted to be deleterious by Provean. 3D structure of Variant E of a CD14 molecule (Fig 8). The sites for disulphide bond was predicted in blue spheres. K296L was predicted in red. T209S in green, R277S yellow, V289L orange.

Results in I-mutant revealed a large decrease of thermodynamical stability for T209S, R277S, V289L, R337H. Hence the mutation at these sites will also lead to structural instability and ultimately decreased function.

T209S, R277S moieties are present in LRR region, which in turn is present for pathogen recognition. V289L moiety is present in leucine zipper site. Leucine zipper is responsible for dimerization and binding with TLR4, MD2 and its functioning. R337H is the site for the alpha helix.

Mutation of CD14 buffalo at the site K296L indicates an alteration in the major function of some of the domain as LRR, domain linker, and leucine zipper. This zipper site is very important for dimerization and pathogen binding.

Variant A versus Variant F

This high degree of mutations has an effect on its structure also as revealed by TM-align software. Deleterious mutations as revealed were L255P, S256N, H257P, S259L, L260P, C270K, W272H (Table 1). Structural alignment of CD14A with CD14 F has revealed a greater degree of variation, RMSD was observed to be 1.81. The structural alignment has been observed in Fig 9. The site for mutation as L255P, S256N, H27P had been depicted as red sphere. S259L and L260P had been depicted in magenta. C270K as cyan. W272H as grey. 255 to 257 aa of CD14 Comprises for LRR. Both C270K, W272H sites comprise of LRR domain. Site C270K encodes for a di-sulphide bond. Alteration in amino acid moiety at site 270, impairs di-sulphide bonding, which affects protein folding. Impaired protein folding, in turn, effects its function. Much structural dissimilarity has been observed as RMSD 1.81.

Variant A versus G

Seven non-synonymous mutations were observed out of which K296 L was predicted to be deleterious by Provean (Table 1). Structural alignment revealed differences with RMSD was observed to be 0.22 (Fig 10). Variant A as green, G by Cyan. The sites for disulfide bond was predicted in blue spheres. K296L was predicted in red.

Results in I-mutant revealed a large decrease of thermodynamical stability for T209S, R277S, V289L, R337H (Table 1). Hence the mutation at these sites will also lead to structural instability and ultimately decreased function.

T209S, R277S moieties are present in LRR region, which in turn is present for pathogen recognition. V289L moiety is present in leucine zipper site. Leucine zipper is responsible for dimerization and binding with TLR4, MD2 and its functioning. R337H is the site for the alpha helix.

Mutation of CD14 buffalo at the site K296L indicates an alteration in the major function of some of the domain as LRR, domain linker, and leucine zipper. This zipper site is very important for dimerization and pathogen binding. Structural dissimilarity with RMSD 0.22 has also been observed.

Variant A versus variant H

Three non-synonymous mutations were detected as T209S, T235S, R277S (Table 1). However, the mutations were observed to be thermodynamically stable as revealed by I-mutant and none of the mutations were observed to be deleterious.

The alignment of all the variants of CD14 had been depicted in Fig 11, which clearly reveals the mutations, both synonymous and non-synonymous for bubaline CD14.

3.2. Association of SNPs with phenotypic traits related to mastitis

Mutations in CD14 effect the binding ability which in turn affects the biological potentiality. Mutational hotspots were detected in bubaline CD14 with 58 number of non-synonymous SNP, out of which 18 were observed to be deleterious and 34 as thermodynamically unstable.

Eight variants of CD14 were identified, where variant A was observed to be the wild type. Phenotypic association with different indicator traits for mastitis were identified and listed in Table 2.

Association of SNP of CD14 with Somatic cell count

Polymorphism or SNPs of CD14 gene were observed to be significantly associated with somatic cell count. Significant differences somatic cell count were observed between wild-type Variant A with variant D and variant F

Association of SNP of CD14 with California mastitis test

Polymorphism or SNPs of CD14 gene were observed to be significantly associated with California mastitis test. Variant D and variant F of CD14 was observed to be strongly positive in contrast to a negative response in wild-type Variant A and variant H. Variant G was also observed to have a distinct positive response for California mastitis test (Table 2). As evident in the Table 2, the data for California mastitis test was observed to be discrete variable. For the analysis, scoring was used for qualitative data, then employed log transformation in order to maintain continuity of data.

Association of SNP of CD14 with rennet coagulation time

Polymorphism or SNPs of CD14 gene were observed to be significantly associated with rennet coagulation time. Significant differences were observed between wild type Variant A with variant D and variant F of CD14 protein (Table 2).

Mutated CD14 with inactive biological function was observed to be responsible for mastitis.

Discussion

CD14 is an important molecule conferring innate immunity. CD 14 is a pattern recognition receptor-rich in leucine-rich repeat (LRR) with N-terminal hydrophobic pocket or groove responsible for binding with pathogens. It can bind with a wide range of substances as LPS, lipotechoic acid, mannuronic acid, which are an integral part of pathogens and releases a series of cytokines, which ultimately destroy the pathogen. Extensive variability among CD14 has been observed between the ruminant species (Pal et al., 2018). Most of the variations have been observed in LRR region, LPS binding and LPS recognition site. Since LRR region is the site for pathogen recognition, the variability among the LRR leads to recognition of the wide range of pathogens.

CD14 acts through TLR 4 signalling pathway and interacts with a series of genes as revealed by String analysis and KEGG analysis. CD14 shRNA was observed to have effect on gene expression of TLR4 signal pathway in monocyte and macrophages of buffalo (Li et al., 2014).

Six LRR had been identified in bubaline CD14. CD 14 is basically a receptor molecule with the N-terminal pocket having the receptor binding groove. There is wide variability in CD14 variants with changes in LPS binding site and LRR site.

In our current study, we have detected wide variability as eight variants of CD14. Most of the variants were observed in six LRR regions. Bubaline CD14 with nucleotide sequence from 587th to 854th position of the second exon of buffalo (coding for amino acid sequence 197th to 285th) was observed to be highly variable among the variants of bubaline CD14 and designated as Mutational Hotspot. This fragment codes for the leucine-rich repeats (LRR) starting from 195-218, 220-245, 247-270, 273-285th codon from the start codon. Leucine rich repeat (LRR) are important domain of CD14 responsible for pathogen binding and pathogen recognition. CD14 is basically a receptor molecule with the ability to bind with a wide range of ligands as lipopolysaccharides (from gram negative bacteria),mannuronic acid

(from *Pseudomonas* spp.) lipoarabinomannans (from *Mycobacterium* spp.), peptidoglycans (from *Staphylococcus aureus*). Hence the CD14 receptor site must be flexible enough to accommodate these ligands. Within the leucine-rich repeats, the particular leucine moiety was observed to be almost unchanged and the variations were observed for other amino acids.

A total of 59 non-synonymous SNPs were identified, which exceeds the synonymous substitutions indicating positive selection. For nine SNPs, the wild-type amino acid had been substituted by different sets of amino acids corresponding to different variants of CD14. 24 sites were observed to be only thermodynamically unstable as revealed by I-mutant software, seven sites as only deleterious by Provean software. Eleven sites were identified as both of unstable and deleterious mutations.

In our earlier study, differences were observed among the CD14 molecule of different breeds of buffalo. In Bhadawari breed, only three patterns were identified; AA genotype having the highest frequency (0.818), whereas the lowest frequency was identified as AB. In Mehsana, the most frequent genotype was observed to be AA genotype (0.465), whereas AB (0.056) and AG (0.056) were found to be the least frequent genotypes. Murrah breed, the most frequent genotype was observed to be AA (0.596), whereas AB (0.105) and AD (0.105) were the least frequent genotypes. In contrast to other buffalo breeds, in Surti breed, the highest frequency was observed for AD genotype, and AA (0.071) and AH (0.071) genotypes were the least frequent. Since the frequency of homozygote (AA) was less compared to that of heterozygote (AD), a higher degree of heterozygosity among the population of Surti breed of buffalo was predicted. As Bhadawari possess the highest frequency homozygote (AA), the population was predicted to be more homozygous.

The frequency of A allele ranged from 0.909 (Bhadawari) to 0.5 (Surti). The allelic frequencies for both G and H were found to be too low (0.008) to be regarded as polymorphic (Pal et al., 2014).

There was no change at the position of N-linked glycosylation sites of buffalo CD14 predicted peptide sequence. N-linked glycosylation is essential for the polarity and solubility of protein or protein folding. It determines if the CD14 molecule will be available either in membranous and soluble form.

We had reported for the first time the polymorphism with 52 SNPs identified at the coding region of CD14 gene of buffalo. In our earlier studies, we could detect polymorphism of CD14 gene in buffalo (Pal et al., 2014), cattle (Pal et al., 2011) and goat (Pal et al., 2009, Pal et al., 2013) with mutational hot spot detected. Variations were also observed for CD14 among different ruminant species.

In the current study, we could study the molecular mechanism of a depressed phenotype of a mutant variant of CD14, which is reported for the first time. Two basic mechanisms were responsible, the first is the deleterious nature of the mutation and second is the decreased thermodynamic stability of the mutant CD14. The mutational hot spot was detected and variant CD14 D and F were mostly affected as they contain the large number of identified non-synonymous mutation, which was deleterious in nature and was of reduced thermodynamic stability. As evident from the study that most of the mutations were confined to LRR region or effecting the vital regions for post translational modification. The binding site

for CD14 resides at N-terminal region, composed mostly of LRR region. The deleterious mutations of CD14 affecting the LRR sites of the binding site may lead to the reduction of the pathogen binding ability of CD14. Some other set of deleterious mutations were observed, which affect LRR, Domain linker, leucine zipper, and disulphide bonding. These are the vital functions of CD14 molecule. Leucine zipper is important for dimerization and pathogen binding. Deleterious mutations in this site cause disrupted function of leucine zipper, which may ultimately cause functional reduction of CD14. The functional form of CD14 is the dimer form. Disulphide bond is responsible for protein folding. Any protein including CD14 is active only when it is in 3D structural form.

The second factor as thermodynamical stability is also equally important. Decreased thermodynamic stability causes reduced half life of the biologically active CD14 molecule. Thus the variant of CD14 with more number of non-synonymous mutations with deleterious or unstable mutations, will definitely have reduced phenotypic value.

Considering the polymorphic nature of the gene, a critical association study encompassing all regions of the CD14 gene in a large population might reveal its functional importance in Buffalo. Thus in this study, we had conducted association study of variants of CD14 gene with the traits for udder health or mastitis as somatic cell count, California mastitis test and rennet coagulation time and significant differences were observed.

From the *in silico* study and also from experimental validations, Variant A was observed to be the wild type. Association study of CD14 gene with the indicator traits for mastitis revealed significant differences and confirmed the above findings. Variant H was also observed to be the resistant variant. Out of three non-synonymous SNPs identified, all were observed to be neutral. CD14 variants as D and F were observed to be highly deleterious and have high somatic cell count, rennet coagulation time and strongly positive for California mastitis test and Variant G as deleterious to some extent.

Genetic polymorphism of CD14 gene and its polymorphism with disease genetics (mastitis and others) had been studied by a number of workers. Report indicated that CD14 SNPs regulate the innate immune response (Liu et al., 2012).

However, most of the studies were in the promoter region of CD14 gene. CD14 gene variants at C-260T were observed to be associated with coronary heart disease risk in human (Dai et al., 2016). CD14 gene polymorphism has also been associated with a variety of diseases as cancer (Wang et al., 2014), sarcoidosis (Fridlender et al., 2010), pulmonary tuberculosis (Rosas-Taraco et al., 2007), asthma (Wang et al., 2005), ischemic heart disease (Zhang et al., 2009), mastitis (Selvan et al., 2014, Baro, D. 2012), alcoholic liver disease and chronic hepatitis C infection (Meiler et al., 2005), coronary artery disease (Li et al., 2015). CD14 has been identified as genetic risk factor for restenosis after percutaneous coronary intervention (Zholdybayeva et al., 2016). C-159T SNP of the CD14 gene promoter was observed to have effect on lung dysfunction in smokers (Zhou et al., 2009).

Characterization of bovine CD14 and association studies were conducted to explore the surface expression on monocytes and polymorphonuclear neutrophils, an indicator of immune status (Ibeagha-Awemu, 2008). The effect of CD14 in inflammatory responses during mastitis had been depicted (De Schepper et al., 2008; Buitenhuis, et al., 2011). Reports indicate association of CD14 SNPs with SCC (Sharma, et al. 2011; Ibeagha-Awemul et al. 2008). SNPs in the CD14 Promoter decreased the ability of Sp Protein Binding and increases Transcriptional Activity (LeVan et al., 2001). Meta-analysis was employed to study the association of CD14 with cancer risk (Wang et al.,2014), coronary heart disease risk (Pu et al., 2013) inflammatory bowel disease (Wang et al., 2012) and coronary artery disease (Li et al., 2015).

It is understood from the current study that the mutation in CD14 gene causes unstable and deleterious mutations, which ultimately causes depressed function as immunity leading to the occurrence of diseases, mastitis in the current case. The variability identified in the CD14 gene sequence and SNPs and variants may be identified at the day old stage and employed for marker-assisted selected to retain the wild-type for the production of a future disease-free stock of the farm. Wild-type CD14 identified may be useful for recombinant CD14 production for therapeutic use. Wild-type variant of CD14 may be employed for somatic gene therapy. Ultimately transgenic animal may be produced from wild variant CD14.

Conclusion

CD14 is an important pattern recognition receptor with ligand binding activity for lipopolysaccharide, lipoteichoic acid, mannuronic acid. Bubaline CD14 had been observed to be hypervariable and mutational hot spot have been detected, particularly at LRR region, responsible for pathogen recognition. Variant A has been observed to be the wild-type by both in- silico study and experimental validation with somatic cell count, California mastitis test and rennet coagulation time. Mutations in CD14 gene have SNPs and Indel, causing thermodynamically unstable and deleterious CD14 peptide with reduced immunological properties and diseases. Eight variants of bubaline CD14 were identified, out of which variant D and F were identified to have most of the deleterious and thermodynamically unstable mutations, leading to more susceptibility to mastitis. Proper characterization may lead to future study for therapeutic application of recombinant CD14 wild type variant A, somatic gene therapy, transgenic animal production with gene insert from Variant A and marker-assisted selection for variant A for production of future stock of livestock with reduced susceptibility of diseases, particular mastitis.

Abbreviations:

CD14 (Cluster of Differentiation 14), SNP (Single nucleotide Polymorphism), LRR (Leucine rich repeat), DMRT (Duncan multiple range test), LPB (Lipopolysaccharide binding protein), SCC (Somatic cell count), LPS (Lipopolysaccharide), CMT (California mastitis test), TLR(Toll like receptor), ANOVA (Analysis of variance)

Declaration

Ethics approval : The current study ad ethical clearence from Institute Animal Ethics Committee, Indian Veterinary Research Institute, Izatnagar, UP

Competing interests : There exists no competing interest.

Funding: Funding for the research work was provided by IVRI as institute fund (during PhD work)

Authors' contributions:

Aruna Pal has designed the project, conducted the research work, analyze data and wrote the manuscript. Abantika Pal has conducted the bioinformatics analysis. Arjava Sharma has conceptualized the project and corrected the manuscript. T.K.Bhattacharya has conceptualized the project, designed it, analyzed data and corrected the manuscript.

Acknowledgements:

I do hereby acknowledge Director, Indian Veterinary Research Institute, for providing the laboratory facilities, and the fund for carrying out the PhD work. I would like to thank the In Charge, Govt. Livestock Farm, Etawah (U.P.), Gujrat Agricultural University, S. K. Nagar (Gujrat), LPM Section, IVRI (Indian Veterinary Research Institute), Izatnagar (U.P.), Gujrat Agricultural University, Navsari, Anand (Gujrat) for providing the samples for carrying out the research work..

Conflict of Interest:

There exists no conflict of interest.

References

Rosas-Taraco , A.G., Revol , A., Salinas-Carmona , M.C., Rendon , A., Caballero-Olin , G., Arce-Mendoza , A.Y . 2007. CD14 C(-159)T Polymorphism Is a Risk Factor for Development of Pulmonary Tuberculosis . The Journal of Infectious Diseases, 196 (11): 1698–1706.

Ayano,A. A., Hiriko ,F., Simyalew, A. M. and Yohannes, A. 2013. Prevalence of subclinical mastitis in lactating cows in selected commercial dairy farms of Holeta district. Journal of Veterinary Medicine and Animal Health Vol. 5(3), pp. 67-72.

Bazil V, Horejsi V, Hilgert I, and McMichael M J, eds. (1982) Leucocyte typing III. White cell differentiation antigens. In The Workshop: Myeloid Panel Antibodies Recognizing the 53-kDa Molecular Weight Monocyte Antigen (CD14). Oxford University Press, Oxford , UK, p. 611–613.

Bittante, G. 2011. Modeling rennet coagulation time and curd firmness of milk. Journal of Dairy Science 94(12):5821-32

Boichard et al 2003; CD14 deduced from cDNA clones. *Indian J. of Animal Science*. 83:10, 1062-1067.

Buitenhuis, B., C.M. Rontved, S. M Edwards, K.L. Ingvarsten and P. Sorensen, 2011. In depth analysis of genes and pathways of the mammary gland involved in the pathogenesis of bovine *Escherichia coli* mastitis. *BMC Genomics*, 12: 130-130. DOI: 10.1186/1471-2164-12-130

Meiler, C., Mühlbauer, M., Johann, M., Hartmann, A., Schnabl, B., Wodarz, N., Schmitz, G., Schölmerich, J., and Hellerbrand, C. 2005. Different effects of a CD14 gene polymorphism on disease outcome in patients with alcoholic liver disease and chronic hepatitis C infection. *World J Gastroenterol*. 11(38): 6031–6037.

Dai X, Wiernek S, Evans JP, Runge MS. Genetics of coronary artery disease and myocardial infarction. *World J Cardiol* 2016; 8(1): 1-23.

Baro, D..2012. Polymorphism of CD14 gene and its association with mastitis in Murrah buffaloes. Thesis submitted to the Lala Lajpat Rai University of Veterinary and Animal Sciences in partial fulfillment of the requirement for the degree of M.V.Sc in Animal Biotechnology.

De Scheppe, S., De Ketelaere, A., Bannerman, D.D., Paape, M.J., Peelman, L., Burvenich, C. 2008. The toll-like receptor-4 (TLR-4) pathway and its possible role in the pathogenesis of *Escherichia coli* mastitis in dairy cattle. *Vet. Res.* (2008) 39:05.

Detilleux, J.C. 2012. A mathematical model to study resistance and tolerance to infection at the animal and population levels: application to *E. coli* mastitis. *Front Genet*. 2012; 3: 146.

Dhokal, I.P . 2006. Normal Somatic Cell Count and Subclinical Mastitis in Murrah Buffaloes. 53 (2): 81-86.

Erridge C, Bennett-Guerrero E, Poxton IR (2002) Structure and function of lipopolysaccharides. *Microbes Infect*. 4: 837–851.

Fridlender, Z.G., Schwartz, A., Kohan, M. Amir G., Glazer, M., Berkman, N. 2010. Association between CD14 gene polymorphisms and disease phenotype in sarcoidosis. *Respiratory Medicine*. 104 (9) , : 1336-1343

Gioannini TL, Teghanemt A, Zhang D, Coussens NP, Dockstader W, Ramaswamy S, Weiss JP (2004) Isolation of an endotoxin-MD-2 complex that produces Toll-like receptor 4-dependent cell activation at picomolar concentrations. *Proc. Natl. Acad. Sci. USA* 101: 4186–4191.

Giraud, J.A., Calzolari, A., Rampone, H., Rampone, A., Giraud, A.T., Cristina Bogni, C., Larriestra, A., Nagel, R. 1997. Field Trials of a Vaccine Against Bovine Mastitis. 1. Evaluation in Heifers. *Journal of Dairy Science*. 80 (5): 845-853

Goldsby RA, Kindt TJ, Osborne BA (2000) *Kuby Immunology*. 4th Edn. W.H. Freeman and Company, New York.

Goyert SM, Ferrero E, Rettig WJ, Yenamandra AK, Obata F, Le Beau MM (1988) The CD14 monocyte differentiation antigen maps to a region encoding growth factors and receptors. *Science* 239: 497–500.

Guha , A., Guha, R. and Gera, S. 2012. Comparison of somatic cell count, California mastitis test, chloride test and rennet coagulation time with bacterial culture examination to detect subclinical mastitis in riverine buffalo (*Bubalus bubalis*) . *African Journal of Agricultural Research* Vol. 7(41), pp.5578-5584.

Hailman E, Lichenstein HS, Wurfel MM,. Miller DS, Johnson DA, Kelley M, Busse LA, Zukowski MM, Wright S D (1994) Lipopolysaccharide (LPS)-binding protein accelerates the binding of LPS to CD14. *J. Exp. Med.* 179: 269–277.

Haziot A, Chen S, Ferrero E, Low MG, Silber R, Goyert SM (1988) The monocyte differentiation antigen CD14, is anchored to the cell membrane by a phosphatidyl inositol linkage . *J. Immunol.* 141: 547–552.

Ibeagha-Awemu, E. M., Lee, J. W., Ibeagha, A. E. (2008). Bovine CD14 gene characterization and relationship between polymorphisms and surface expression on monocytes and polymorphonuclear neutrophils”. *BMC Genetics*, 9: 50-60

Joshi S, Gokhale S (2006). Status of mastitis as an emerging disease in improved and periurban dairy farms in India. *Ann. N. Y. Acad. Sci.* 1081:74-83.

Kelley ST, Lukk T, Nair SK and Tapping RI (2013) The crystal structure of human soluble CD14 reveals a bent solenoid with a hydrophobic amino terminal pocket *J Immunol*190:1304-1311.

Kim JI, Lee CJ, Jin MS, Lee CH, Paik SG, Lee H, .Lee JO (2005) Crystal structure of CD14 and its implications for lipopolysaccharide signaling. *J. Biol. Chem.* 280: 11347–11351.

Lee JW, Paape MJ, Elsasser TH, Zhao X (2003) Recombinant soluble CD14 reduces severity of intramammary infection by *E.coli*. *Infection and Immunity* 71(7), 4034-4039.

Li,Y.Y., Xiang-Ming Wang, Chuan-Wei Zhou, Jian Xu, Yun Qian. 2015. CD14 gene-159C/T polymorphism and coronary artery disease: a meta-analysis involving 4467 subjects *Int J Clin Exp Med*;8(8):12149-12160

Liu HH, Hu Y, Zheng M, Suhoski MM, Engleman EG, Dill DL, Hudnall M, Wang J, Spolski R, Leonard WJ, Peltz G.2012. CD14 SNPs regulate the innate immune response. *Mol Immunol.* 51(2):112-27.

Pellegrino , M., Berardo , N., Giraudo, J., Nader-Macías , M.E.F. Bogni, C.2017. Bovine mastitis prevention: humoral and cellular response of dairy cows inoculated with lactic acid bacteria at the dry-off period. *Beneficial Microbes*: 8 (4) - Pages: 589 – 596

Meiler C, Muhlbauer M, Johann M, Hartmann A, Schnabl B, Wodarz N, Schmitz G, Scholmerich J, Hellerbrand C.2005. Different effects of a CD14 gene polymorphism on disease outcome in patients with alcoholic liver disease and chronic hepatitis C infection. *World J Gastroenterol.* 2005 Oct 14;11(38):6031-7.

Núñez Miguel R, Wong J, Westoll JF, Brooks HJ, O'Neill LA, Gay NJ, Bryant CE, Monie TP (2007) A dimer of the Toll-like receptor 4 cytoplasmic domain provides a specific scaffold for the recruitment of signaling adaptor proteins. *PLoS ONE* 2: e788.

Opal SM, Cohen J (1999) Clinical gram-positive sepsis: does it fundamentally differ from gram-negative bacterial sepsis? *Crit. Care Med.* 27: 1608–1616.

Oviedo-Boyso J, Valdez-Alarcón JJ, Cajero-Juárez M, Ochoa-Zarzosa A, López-Meza JE, Bravo-Patiño A, Baizabal-Aguirre VM. 2007. Innate immune response of bovine mammary gland to pathogenic bacteria responsible for mastitis. *J Infect.*;54(4):399-409.

Pal, A., Pal, A Jr, Chatterjee, P.N. Sharma, A. and Biswas, P. 2018. Structural - functional domain analysis and molecular phylogeny of CD14 gene and derived peptide in domestic ruminants. *Gene*. communicated.

Pal, A. and Chatterjee, P.N. 2009. Molecular cloning and characterization of CD14 gene in goat. *J. of Small Ruminant Research.* 82: 84-87.

Pal, A., Chatterjee, P.N. and Sharma, A. 2013. Molecular evolution and structural analysis of Caprine CD14 deduced from cDNA clones. *Indian J. of Animal Science.* 83:10, 1062-1067

Pal, A., Chatterjee, P.N., and Sharma, A. 2014. Sequence characterization and polymorphism detection in bubaline CD14 gene. *Buffalo Bulletin.* 32:2, 138-156

Pal, A., Sharma, A., Bhattacharya, T. K., Chatterjee, P. N, and Chakravarty A. K. (2011). Molecular characterization and SNP detection of CD14 gene of crossbred cattle. *SAGE-Hindawi Access to Research, Molecular Biology International*, (2011), article ID 507346, 13 pages, doi: 10.4061/2011/507346

Pereira UP, Oliveira DG, Mesquita LR, Costa GM, Pereira LJ 2011. Efficacy of *Staphylococcus aureus* vaccines for bovine mastitis: a systematic review. *Vet Microbiol.* 2011 Mar 24;148(2-4):117-24.

Petrie A., & Watson P. 2008. *Statistics for veterinary and animal science.* Blackwell Publishing; London:

Prohinar P, Re F, Widstrom R, Zhang D, Teghanemt A, Weiss JP, Gioannini TL (2007) Specific high affinity interactions of monomeric endotoxin. protein complexes with Toll-like receptor 4 ectodomain. *J. Biol. Chem.* 282: 1010–1017.

Pu, H., Yin, J., Wu, Y. 2013. The association between CD14 gene C-260T polymorphism and coronary heart disease risk: A meta-analysis. *Molecular Biology Reports* 40(6).

Pugin, J., I. D. Heumann, A. Tomasz, V. V. Kravchenko, Y. Akamatsu, M. Nishijima, M. P. Glauser, P. S. Tobias, and R. J. Ulevitch (1994) CD14 is a pattern recognition receptor. *Immunity* 1: 509–516.

Raetz CR, Whitfield C (2002) Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* 71: 635–700.

Rietschel ET, Brade H, Brade L, Kaca W, Kawahara K., Lindner B, Luderitz T, Tomita T, Schade U, Seydel U, et al. 1985. Newer aspects of the chemical structure and biological activity of bacterial endotoxins. *Prog. Clin. Biol. Res.* 189: 31–51.

Salomao R, Brunialti MK, Rapozo MM, Baggio-Zappia GL, Galanos C and Freudenberg M (2012) Bacterial sensing, cell signaling, and modulation of the immune response during sepsis. *Shock* 38: 227–242.

Sambrook, J.F. and Russell, D.W. 2001. *Molecular Cloning: A Laboratory Manual*, 3rd ed., Vols 1,2 and 3 ed., Cold Spring Harbor Laboratory Press, 2001, 2100 pp. ISBN-13 978-0-87969-577-4

Schroder NW, Opitz B, Lamping N, Michelsen KS, Zahringer U, Gobel UB, Schumann RR .2000. Involvement of lipopolysaccharide binding protein, CD14, and Toll-like receptors in the initiation of innate immune responses by *Treponema glycolipids*. *J Immunol.* 1;165(5):2683-93.

Selvan,S., Gupta, I.D., Verma,A., Chaudhari, M., V., Kanungo, S., and Kumar, R.S. 2014. Characterization of promoter region of CD14 gene and association with clinical mastitis in Karan Fries cattle. Indian Journal of Animal Research. Volume 48 Issue 6: 545-547

Sharma, N., Singh, N.K. and Bhadwal, M.S. 2011. Relationship of Somatic Cell Count and Mastitis: An Overview Asian-Aust. *J. Anim. Sci.* Vol. 24, No. 3 : 429 - 438

Shimazu R, Akashi S, Ogata H, Nagai Y, Fukudome K, Miyake K , Kimoto M (1999) MD-2, a molecule that confers lipopolysaccharide responsiveness on Toll-like receptor 4. *J. Exp. Med.* 189: 1777–1782.

Simmons DLS, Tan DG, Tenen A, Weller N, Seed B (1989) Monocyte antigen CD14 is a phospholipid anchored membrane protein. *Blood* 73: 284–289.

SPSS (1999)", <https://www.ibm.com/in-en/marketplace/spss-statistics>

Teghanemt A, Prohinar P, Gioannini TL, Weiss JP (2007) Transfer of monomeric endotoxin from MD-2 to CD14: characterization and functional consequences. *J. Biol. Chem.* 282: 36250–36256.

Teghanemt, A., Widstrom RL, Gioannini TL, and Weiss JP (2008) Isolation of monomeric and dimeric secreted MD-2. Endotoxin.sCD14 and Toll-like receptor 4 ectodomain selectively react with the monomeric form of secreted MD- 2. *J. Biol. Chem.* 283: 21881–21889.

Tobias PS, Soldau K, Gegner JA, Mintz D., Ulevitch RJ (1995) Lipopolysaccharide binding protein-mediated complexation of lipopolysaccharide with soluble CD14. *J. Biol. Chem.* 270: 10482–10488.

LeVan , T.D., Bloom , J.D., Bailey , T.J., Karp , C.L., Halonen , M., Martinez, F.D. and Vercelli, D. 2001. A Common Single Nucleotide Polymorphism in the CD14 Promoter Decreases the Affinity of Sp Protein Binding and Enhances Transcriptional Activity. *J Immunol.* 167 (10) 5838-5844.

Wang J, Guo X, Yu S, Song J, Zhang J. (2014) Association between CD14 Gene Polymorphisms and Cancer Risk: A Meta-Analysis. PLoS ONE 9(6): e100122. doi:10.1371/journal.pone.0100122

Wang PY, Kitchens RL, Munford RS (1998). Phosphatidylinositides bind to plasma membrane CD14 and can prevent monocyte activation by bacterial lipopolysaccharide. J. Biol. Chem. 273: 24309–24313.

Wang Z, Hu J, Fan R, Zhou J, Zhong J (2012) Association between CD14 Gene C-260T Polymorphism and Inflammatory Bowel Disease: A Meta- Analysis. PLoS ONE 7(9): e45144. doi:10.1371/journal.pone.0045144

Wang, J.Y., Wang, L.M., Lin, C.G.J., Chang, A.,C.W., Wu, L.S.H. 2005.

Association study using combination analysis of SNP and STRP markers: CD14 promoter polymorphism and IgE level in Taiwanese asthma children. *Journal of Human Genetics* 50 :36–41

Wurfel MM, Wright SD (1997) Lipopolysaccharide-binding protein and soluble CD14 transfer lipopolysaccharide to phospholipid bilayers: preferential interaction with particular classes of lipid. J. Immunol. 158: 3925–3934.

Li, X.P., Li, M., Huang,S., Qiao, S., Qin, Z., Kang,C., Shi, D. 2014. The effect of Buffalo CD14 shRNA on the gene expression of TLR4 signal pathway in buffalo monocyte, macrophages. 19, 623-637.

Yoon HJ, Shim JH, Yang SH, Chae DW, Kim H, Lee DS, Kim HL, Kim S, Lee JS, Kim YS (2003) Association of the CD14 gene -159 polymorphism with progression of IgA nephropathy. *Journal of Medical Genetics* 40 (2), 104-108.

Yu B, Hailman E, Wright SD (1997). Lipopolysaccharide binding protein and soluble CD14 catalyze exchange of phospholipids. J. Clin. Invest. 99: 315– 324.

Zhang, H. F., Zhong,B. L., Zhu,W.L. , Xie, S.L., Qiu, L. X., Zhu L.G., Yan Wang, Y. & Lei. L 2009. CD14 C-260T gene polymorphism and ischemic heart disease susceptibility: A HuGE review and meta-analysis. *Genetics in Medicine* 11, 403–408

Zholdybayeva, E.V., Talzhanov, Y.A., Aitkulova, A.M. , Tarlykov, P.V. ,Kulmambetova, G.N., Iskakova,A.N., Aliya U. Dzholdasbekova, A.U., Olga A. Visternichan,O.A., Taizhanova, D.Z., Ramanculov, Y.M.2016 . Genetic risk factors for restenosis after percutaneous coronary intervention in Kazakh population *Hum Genomics* (2016) 10: 15. <https://doi.org/10.1186/s40246-016-0077-z>

Zhou, H., Neil E. Alexis, N.E., Almond, M., Donohue, J., Force, C.L., Philip A. Bromberg, P.A., Peden, D.B. 2009. Influence of C-159T SNP of the CD14 gene promoter on lung function in smokers. *Respiratory Medicine*. 103 (9),: 1358-1365

Tables

Table 1: Identified SNPs for CD14 variants

Sl no.	SNP	Var B	Var C	Var D	Var E	Var F	Var G	Var H	Status of mutation
1	G207A	P							Unstable
2	D208Q	P							Neutral
3	T209A	P			P		P		Unstable
4	T209S					P		P	Neutral
5	G210C	P							Both unstable and deleterious
6	L215P			P					Both unstable and deleterious
7	Y225C		P						Neutral
8	A227P	P							Neutral
9	L228V	P							Unstable
10	N230H	P							Unstable
11	A231P	P							Neutral
12	M233L	P							Unstable
13	E234K	P							Unstable
14	E234D		P						Neutral
15	T235S							P	Neutral
16	R247K			P					Unstable
17	V248L			P					Unstable
18	Q249P			P					Neutral
19	P250A			P					Unstable
20	Q251H			P					Unstable
21	S252R			P	P				Neutral
22	S252T					P			Neutral
23	L253R			P					Both unstable and deleterious
24	D254L			P					Both unstable and deleterious
25	L255H			P					Both unstable and deleterious
26	L255P					P			Both unstable and deleterious
27	S256N			P		P			Deleterious
28	H257R			P					Neutral
29	H257P								Deleterious
30	N258C			P					Deleterious
31	S259P			P					Deleterious
32	S259L					P			Deleterious
33	L260P					P			Both unstable and deleterious
34	R261H			P					Unstable
35	R261N					P			Unstable
36	V262P					P			Unstable
37	T263C			P					Both unstable and deleterious
38	T263P					P			Unstable
39	A264T					P			Unstable
40	G266A			P		P			Unstable
41	A267Q					P			Unstable
42	R269G					P			Unstable
43	C270K					P			Unstable
44	V271A			P					Unstable
45	W272H					P			Unstable
46	P273A			P					Both unstable and deleterious
47	S274P			P					Neutral
48	A275V					P			Neutral
49	R277M			P					Both unstable and deleterious
50	R277S							P	Neutral
51	N280D			P					Deleterious
52	S282A			P					Both unstable and deleterious
53	F283L			P	P		P		Unstable
54	E287P				P		P		Neutral

55	Q288R				P		P		Neutral
56	V289L				P		P		Unstable
57	K296L				P		P		Deleterious
58	R337H				P		P		Unstable

Table 2: Association of different variants of CD14 with indicator traits for mastitis

	Somatic cell count (10 ⁵ cells per ml)	California mastitis test	Rennet coagulation time (sec)
CD14 A	0.89 ± 0.23*	Negative	128.70 ± 3.80*
CD14B	1.26±0.34*	Weak	168.56± 3.67*
CD14C	1.15± 0.25*	Trace	135.60 ± 2.89*
CD14D	2.17±0.37**	Strong Positive	286.07 ± 2.87**
CD14E	1.29±0.28*	Weak	175.45± 3.24*
CD14F	1.8 ±0.38**	Strong Positive	265.07± 3.56**
CD14G	1.32±0.39*	Distinct Positive	189.04± 3.23*
CD14H	0.88 ±0.27*	Negative	129.20 ± 3.9*

*Superscript a and b indicate the differences at P < 0.01% of significance.

Figures

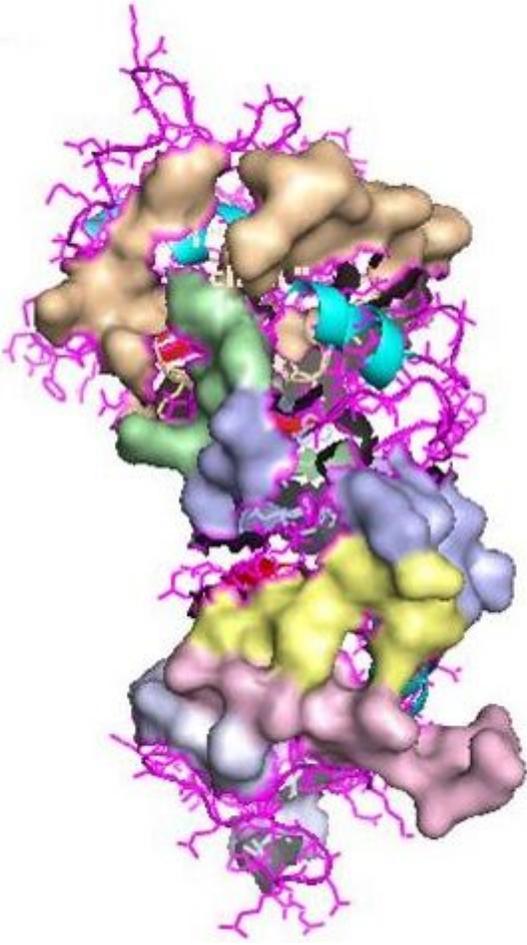


Figure 1

Wild type variant A of bubaline CD14 with LRR (leucine rich repeat)

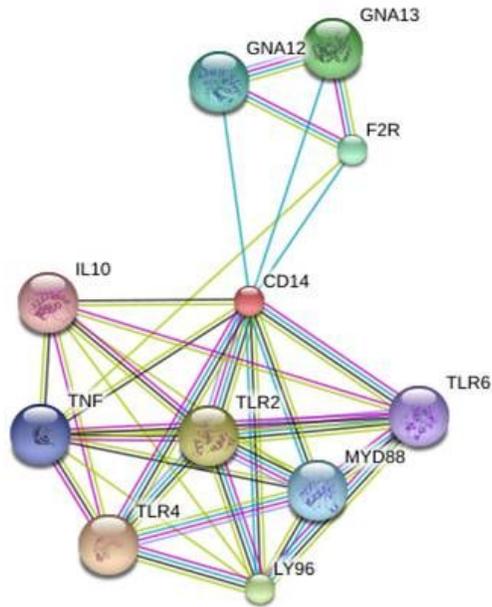


Figure 2

String analysis for the depiction of molecular interaction of CD14 with other related proteins

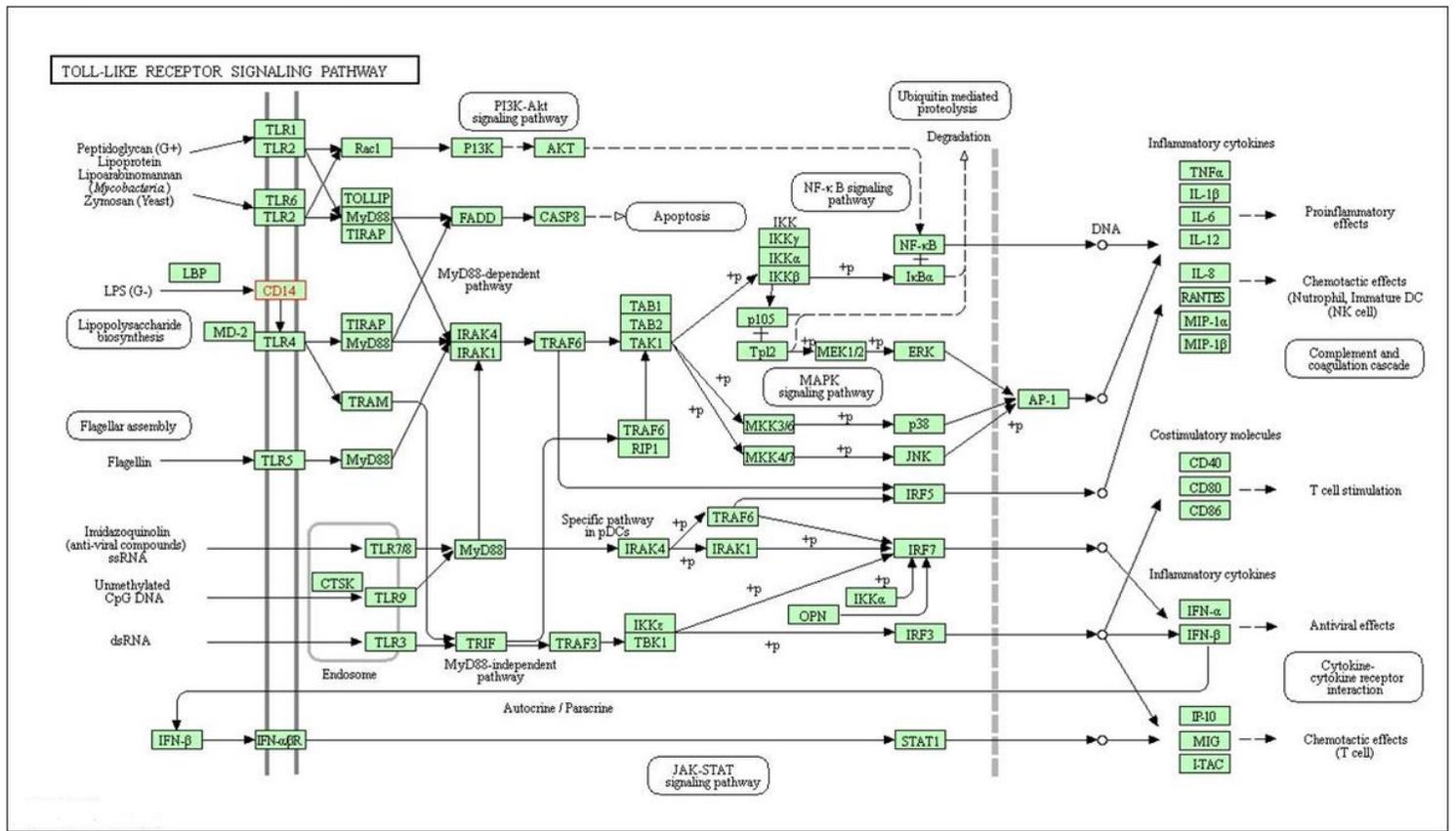


Figure 3

KEGG analysis for the depiction of molecular pathway analysis of CD14

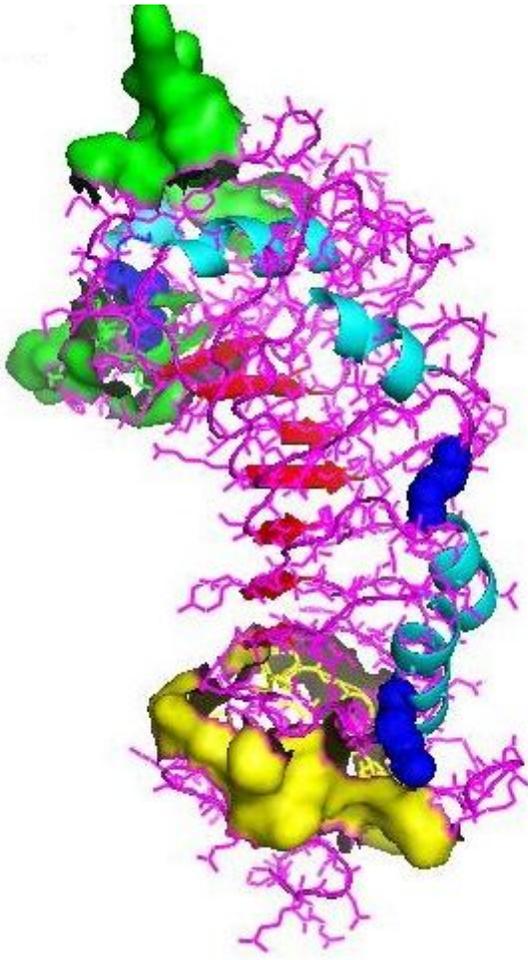


Figure 4

Wild type CD14 A of *Bubalus bubalis*

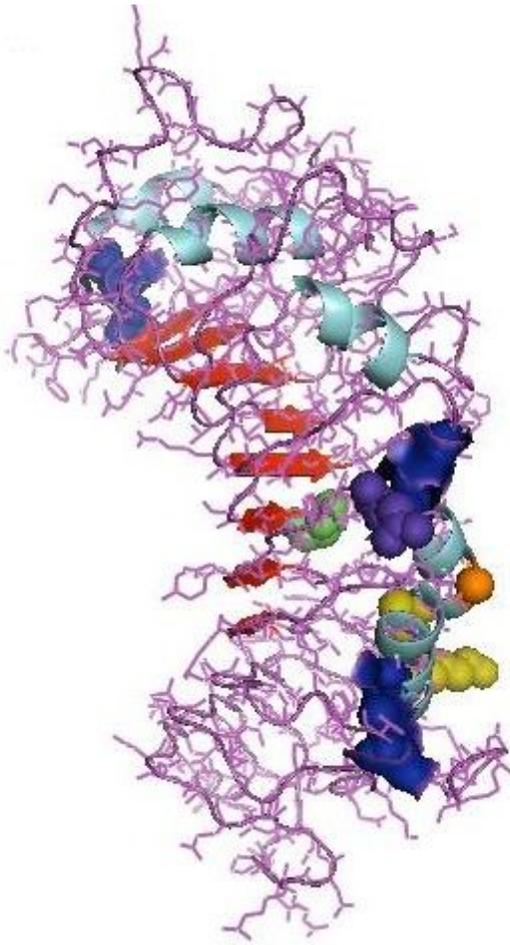


Figure 5

CD14 variant B of *Bubalus bubalis*

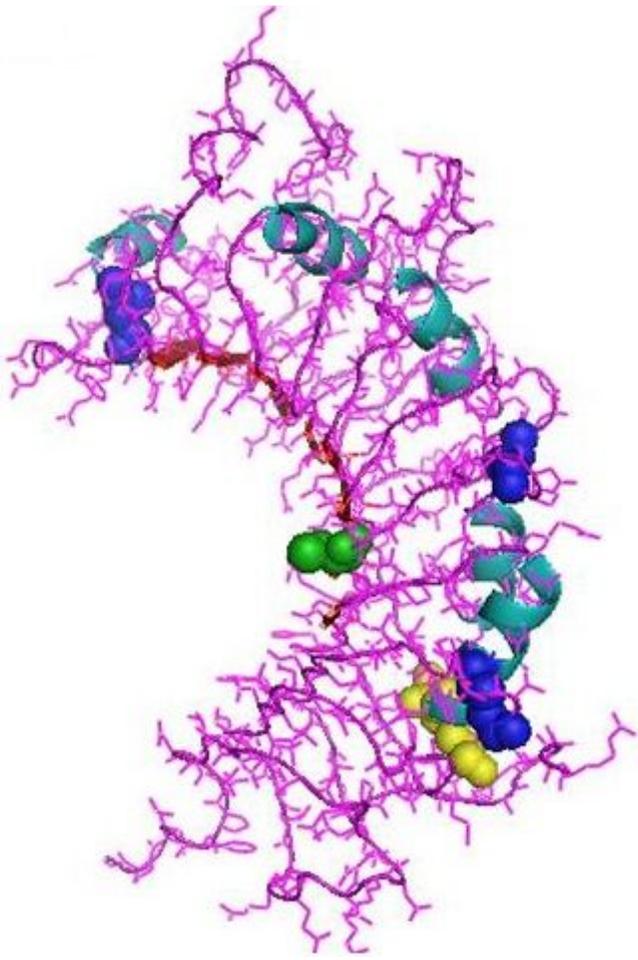


Figure 6

CD14 variant C of Bubalus bubalis

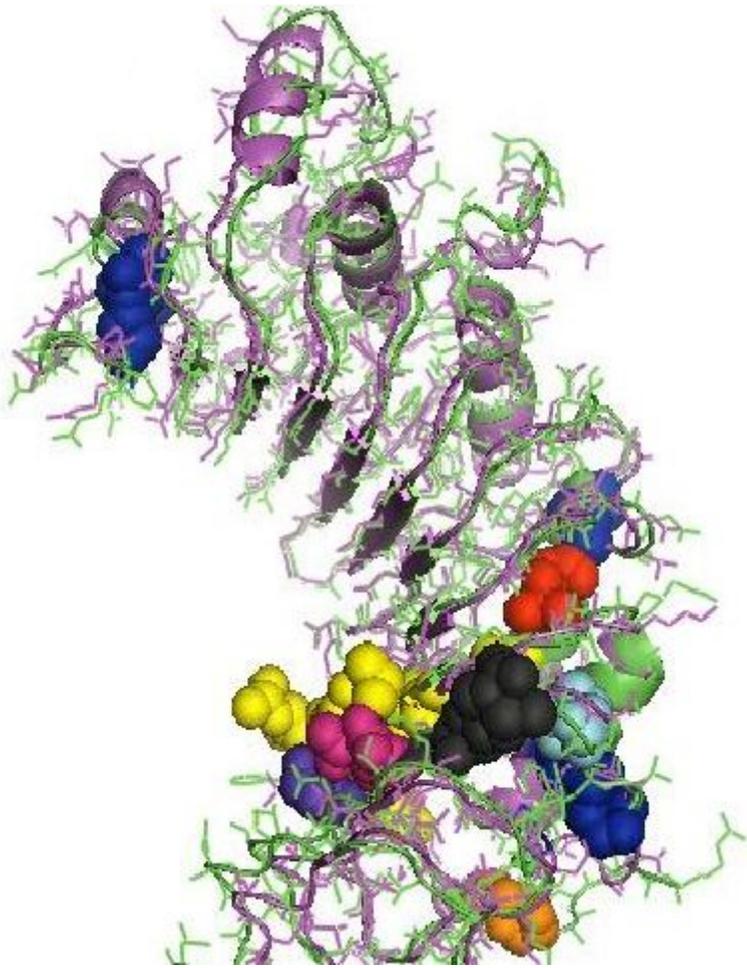


Figure 7

Structural alignment of Var A with Variant D of *Bubalus bubalis*

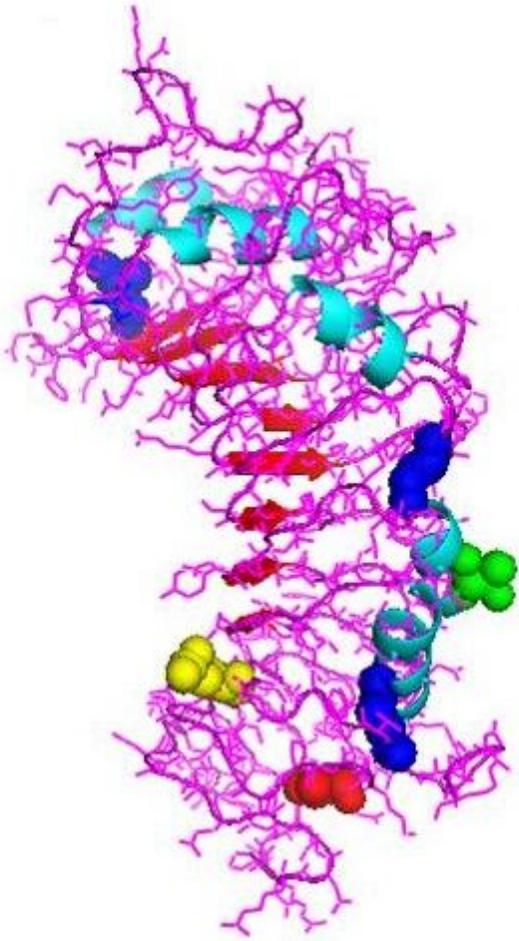


Figure 8

3 D structure of CD14 E of Bubalus bubalis

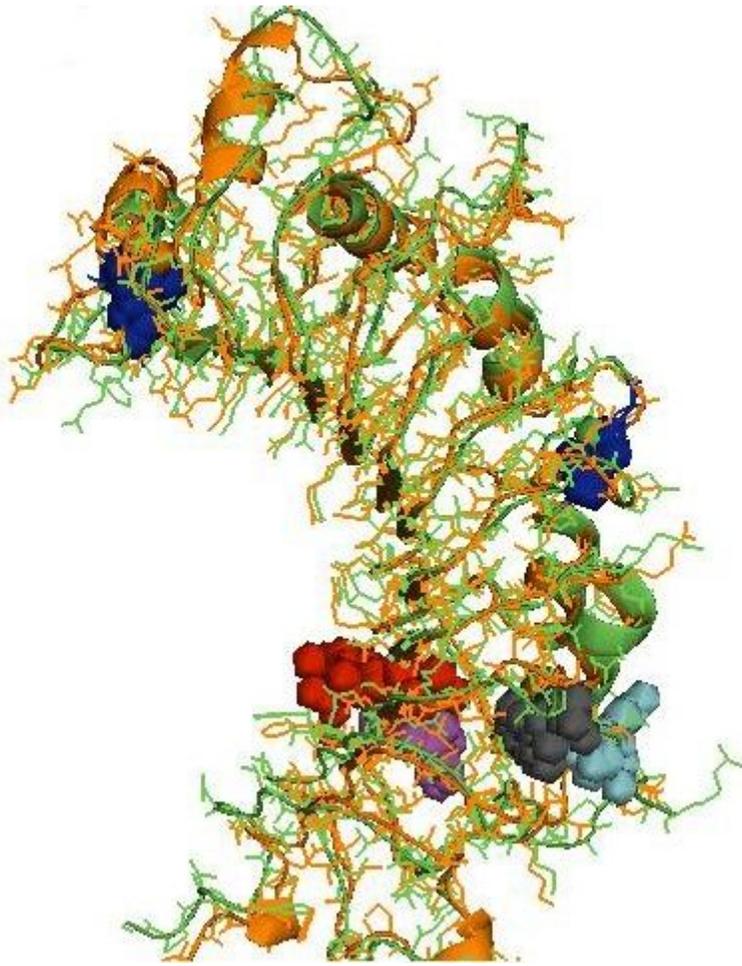


Figure 9

Structural alignment for CD14 A (green) with CD14F (orange) of *Bubalus bubalis*

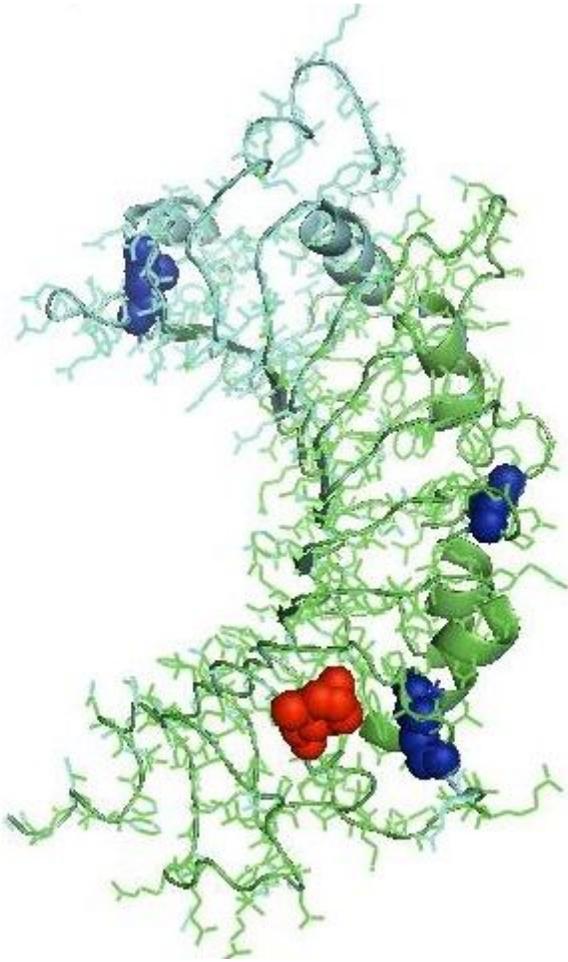


Figure 10

Alignment report for CD14 A with CD14 G of *Bubalus bubalis*

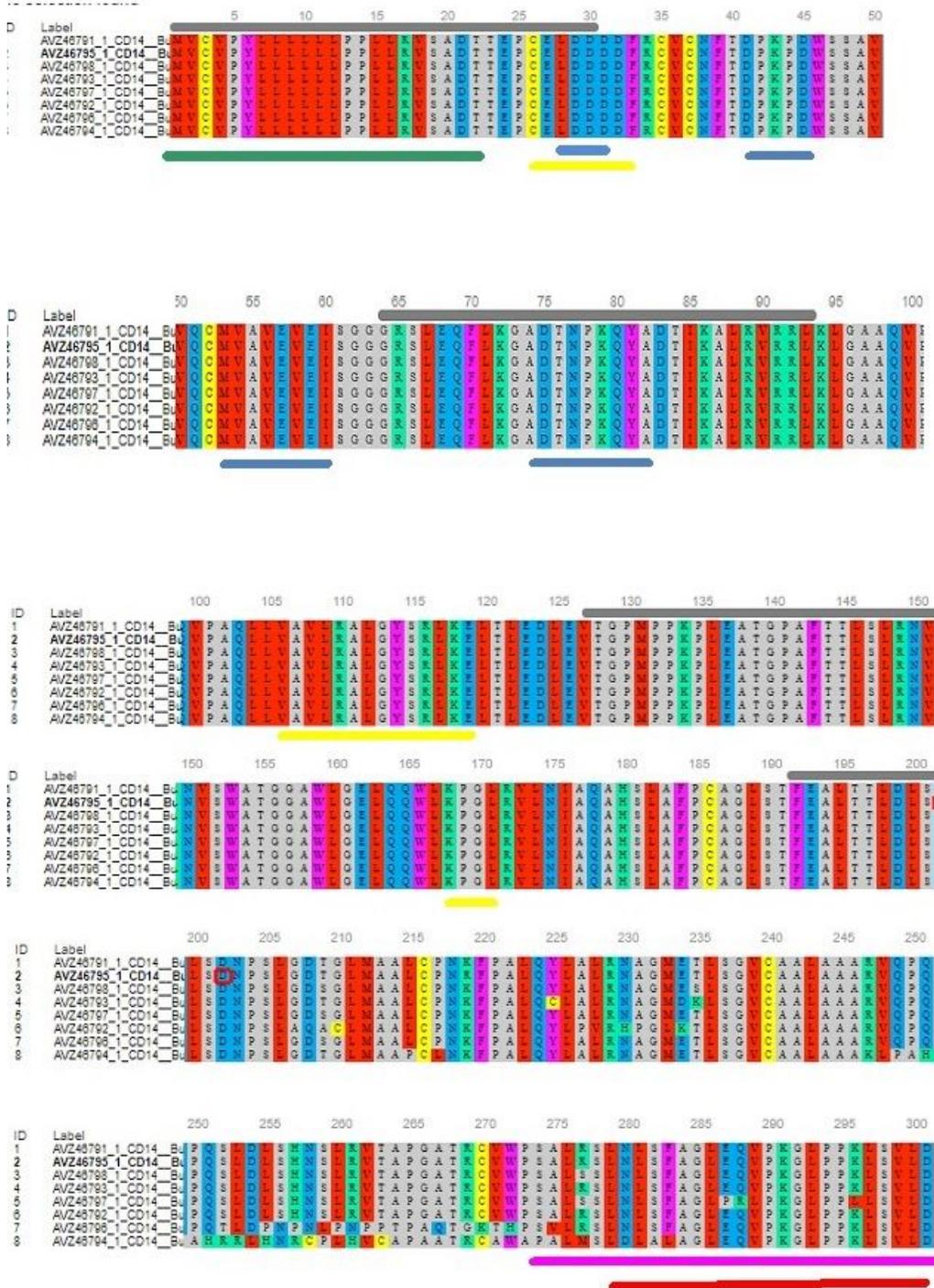


Figure 11

Alignment report of different CD 14 variants of *Bubalus bubalis* Id 1,2,3,4,5,6,7, 8 represents Variant A, B,C,D,E,F,G,H respectively. Green line: Signal peptide Blue line: LPS binding site Yellow line : LPS signaling site Red line : Leucine zipper Black overline: Leucine rich repeat Red line: Leucine zipper Pink line: Domain linker prediction site