Journal of Experimental Biology and Agricultural Sciences, June - 2018; Volume – 6(3) page 516 – 521

Journal of Experimental Biology and Agricultural Sciences

http://www.jebas.org

ISSN No. 2320 – 8694

MOLECULAR CHARACTERIZATION OF CODING REGION OF LACTOFERRIN GENE OF MALABARI AND ATTAPPADY BLACK GOATS OF KERALA

Anjusekar C.[1], Uma Radhakrishnan[2*], Shynu M.[2]

[1]MVSc Scholar, [2]Assistant Professor, Department of Veterinary Biochemistry, College of Veterinary and Animal Sciences, Mannuthy, Thrissur- 680651, Kerala, India

Received – February 24, 2018; Revision – May 16, 2018; Accepted – June 15, 2018
Available Online – June 20, 2018

DOI: http://dx.doi.org/10.18006/2018.6(3).516.521

KEYWORDS

Lactoferrin gene

Characterization

Nucleotide variations

Malabari goats

Attappady Black goats

ABSTRACT

Lactoferrin (Lf), an iron binding glycoprotein mainly found in the secretions of the body like milk, tears, saliva, pancreatic juice, bile, genital fluids etc is well noted for its role in innate immunity as well as biological functions like antibacterial, antifungal, anti-tumorigenic, antiviral and other metal binding properties. The present study unveils the molecular characterization of coding region of lactoferrin (*Lf*) gene of Malabari and Attappady Black goat breeds of Kerala, which are reputed for their sturdiness and resistance to diseases. Thirty minutes post milking milk samples from early lactating goats of both the breeds were collected for RNA isolation followed by cDNA synthesis and subsequent amplification of partial coding region of *Lf* gene. The amplicons were sequenced and the sequences were analyzed using various bioinformatics tools. A 1914 bp long partial coding region encoding 638 amino acids was obtained for Malabari goats while that of Attappady Black goat yielded 1975 bp encoding 657 amino acids. The sequences of both the breeds were 94-99% similar to *Lf* gene of other ruminant species. Eight nucleotide variations were observed in Malabari whereas ten variations were seen in the nucleotide sequences of Attappady Black when compared with *Lf* gene of *Capra hircus* (Gen Bank Acc. No. NM_001285548). The five non-synonymous amino acid variations observed in both the breeds as compared to *C. hircus* were p.Arg88Leu, p.Lys124Gln, p.Pro154Phe, p.Leu357Val and p.Gly414Asp. This is the first report of cDNA sequence and nucleotide variations of *Lf* gene of the indigenous goat breeds of Kerala.

* Corresponding author

E-mail: uma@kvasu.ac.in (Dr. Uma Radhakrishnan)

Peer review under responsibility of Journal of Experimental Biology and Agricultural Sciences.

Production and Hosting by Horizon Publisher India [HPI] (http://www.horizonpublisherindia.in/).

1 Introduction

Lactoferrin (Lf), a multifunctional single polypeptide glycoprotein, is composed of around 690 amino acids, having a molecular mass of 80kDa (Baker & Baker, 2005). It is secreted by epithelial cells and is found in almost all the secretions of the body (Alexander et al., 2012). Lf belongs to the transferrin family and is capable of binding to and transporting Fe^{3+} ions (Metz-Boutique et al., 1984). It exhibits a range of biological activities including antioxidant, antibacterial, antiviral, antifungal as well as metal binding properties (Adlerova et al., 2008). It has been found to modulate innate immune mechanism by reducing pro-inflammatory cytokines (Valenti et al., 2011). As it is present in specific granules of neutrophils (Bennett & Kokocinski, 1978), it is having an essential role in cell mediated immunity.

Lactoferrin (*Lf*) gene is composed of 17 exons and 16 introns (Kim et al.,1998; Seyfert et al., 1994; Kang et al., 2011) and has a size of 34.5 Kb. In cattle and goats, it is located on chromosome 22 (Schwerin et al., 1994). The intron- exon distribution pattern of *Lf* has been observed to be very similar in cattle, sheep and goat, where all these species were having same exon length but varied intron length (Kang et al., 2011).

Kerala, the southern state of India, has two native goat breeds *viz,* Malabari and Attappady Black. Malabari goats, more common in Northern Kerala, are well adapted to the hot humid climate of this region and are popular for their high prolificacy. Attappady Black goats are found in the hilly region of Attappady in Palakkad district and are reared mainly by tribal people for meat purpose. These indigenous goat breeds are lean built, eat tree leaves which are pungent and bitter, consume less water and do lot of physical activities. These breeds are well known for their disease resistance and their adaptability to extreme agro-climatic conditions. Reports on molecular level exploration of major and minor milk proteins of these breeds have been found scanty. Hence the present study was undertaken to characterize the coding region of *Lf* gene in these indigenous goat breeds of Kerala and to compare their sequences with *Lf* sequences available from the database.

2 Materials and Methods

2.1 Isolation of milk somatic cells

A volume of 50 mL milk samples (thirty minutes post milking) were collected from Malabari and Attappady Black goats (in early lactation period) maintained at University Goat and Sheep Farm, College of Veterinary and Animal Sciences, Mannuthy, Kerala and transported in ice. The fresh milk samples were processed immediately as per the protocol of Boutinaud & Jammes (2002) with slight modifications. The milk samples were centrifuged at 2000 x g for 15 min at 4ºC to remove the fat layer and the supernatant was discarded. The pellet containing milk somatic cells (MSC) was washed twice with ice-cold PBS (Phosphate buffered saline)(pH 7.2) supplemented with 0.5mM EDTA (Ethylenediaminetetraacetic acid) and 0.1% DEPC (Diethyl pyrocarbonate) and finally resuspended in 200 µL of PBS-EDTA.

2.2 Isolation of total RNA from MSC and cDNA synthesis

The total RNA from MSC was isolated by TRI-reagent (Sigma Aldrich) as per manufacturer's instructions and then treated with *DNase* 1 (Sigma Aldrich; amplification grade) to remove DNA contamination if any. RNA samples were quantified by Nano Drop spectrophotometer (Thermo Scientific, USA) and checked for the integrity on 1% agarose gel. Reverse transcription was performed to synthesize cDNA from the isolated RNA using Revert Aid First strand cDNA synthesis kit (Thermo Scientific) and oligo dT primers with 0.1 µg of RNA in a reaction volume of 20 µL and were stored at −80 °C until use.

2.3 PCR amplification of *Lf* gene

The oligonucleotides FgLf (5' TGCCGGAGTGGTCCAAATGCTA3') and RgLf (5'GCTTCTTTGCAGGCTTTACCT3') were designed based on the *C. hircus* sequence (Gen Bank Acc. No. NM_001285548) retrieved from database. The custom synthesized primer pair was used in a 25 µL PCR reaction containing 10 picomoles of each primer; 200µmol L^{-1}each of dATP, dCTP, dGTP and dTTP; 1.5 mmol $L^{-1}MgCl_2$; and 0.5 U Jumpstart AccuTaq LA DNA polymerase (Sigma Aldrich) for the amplification of *Lf* gene. The thermal cycling profile consisted of denaturation at 95ºC for 10sec, annealing at 60ºC for 30sec and extension at 68ºC for 90 sec for 35 cycles followed by a final extension at 68ºC for 10 min. The PCR products were electrophoresed in 1% agarose gel for 40 min.

2.4 Sequence analysis

Using FgLf and RgLf primer set, the amplicons were sequenced at the DNA sequencing facility at AgriGenom Pvt. Ltd, Kochi, Kerala. The sequence similarity search was performed using Basic Local Alignment Search Tool (BLASTn) provided by the National Centre for Biotechnological Information (NCBI). Using GeneTool Lite software, the pair wise identity matrix was derived with *Lf* sequences of different species obtained from GenBank. Multiple sequence alignment of the *Lf* gene sequences of Malabari (*MgLf*) and Attappady Black (*AgLf*) with *C. hircus* sequence (Gen Bank Acc. No. NM_001285548) was performed using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) program. The 'Translate' tool of the online portal ExPASy was used to predict

the amino acid sequences encoded by *MgLf* and *AgLf*. The amino acid sequences thus obtained were compared with different Lf protein sequences of mammalian origin present in the database using BLASTp tool of NCBI to find out the similarity between species. Multiple sequence alignment of Malabari and Attappady Black lactoferrin protein sequences with that of the database *C. hircus* Lf amino acid sequence was done using Clustal Omega program. The secondary structure was predicted by SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) (Geourjon & Deleage, 1995). The modeling of the tertiary structure of the protein was done by using SWISSMODEL server (https://swissmodel.expasy.org) (Biasini et al., 2014). The phylogenetic relationship of *Lf* gene with that of different species was analyzed using MEGA version 6.0 software.

3 Results

3.1 Amplification and sequencing of Lf cDNA

The coding region of *Lf* gene of Malabari and Attappady Black goat breeds were successfully amplified using custom-synthesized primer pair. On electrophoresis, the amplicons showed a single band of about 2Kb size (Figure 1). The gel purified amplicons from both the goat breeds were sequenced. The sequences obtained were of size 1917 bp for Malabari *Lf* gene (*MgLf*) (GenBank Acc. No: MG980401) and 1975 bp for Attappady Black *Lf* gene (*AgLf*) (GenBank Acc. No: MG980402).

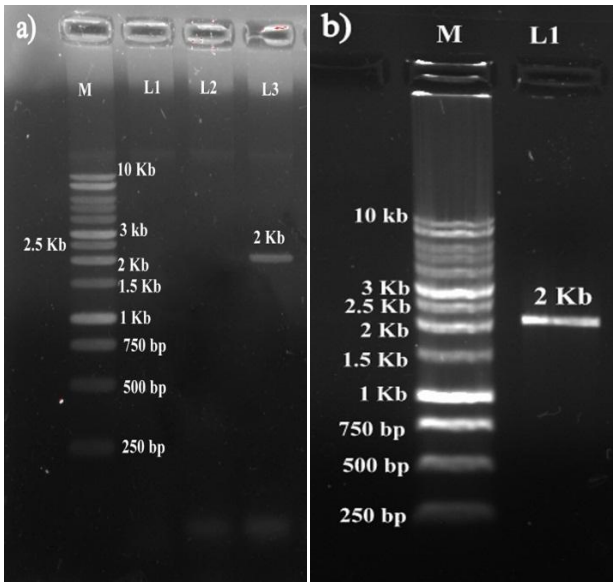

Figure 1 PCR amplified *Lf* gene of a) Malabari goat; b) Attappady Black goat. a) Lane M- Molecular marker (1kb); Lane L3- Malabari *Lf* gene amplicon; b)Lane M- Molecular marker (1kb);Lane L1- Attappady Black *Lf* gene amplicons

3.2 Sequence analysis

3.2.1 Nucleotide similarity analysis

The obtained nucleotide sequences, *MgLf* and *AgLf*, were subjected to BLASTn analysis to ascertain their identity as *Lf* gene. The pair wise identity matrix of *MgLf* and *AgLf* with *Lf* sequences of different mammalian species retrieved from the NCBI database showed 99% homology with the GenBank *C. hircus* sequence (NM_001285548.1) and 94-99% similarity with that of other ruminants.

3.2.2 Multiple sequence alignment

Using Clustal Omega program, the multiple sequence alignment of *MgLf*, *AgLf* and *C. hircus* sequence (NM_001285548) was obtained. On comparing with the GenBank sequence, *MgLf* sequence showed 8 nucleotide variations while *AgLf* showed a total of 10 variations.

3.2.3 Prediction of amino acid sequences

The sequences of the proteins encoded by *MgLf* and *AgLf* obtained by *in silico* translation were found to be of size 638 amino acids for Malabari Lf and 657 amino acid residues for Attappady Black Lf.

3.2.4 Protein similarity search

The predicted amino acid sequences of both the goat breeds were analyzed for similarity with the database Lf sequences belonging to different species using the BLASTp tool of NCBI. Both Malabari and Attappady Black Lf showed 99% similarity to *C. hircus* Lf protein in the database. The multiple sequence alignment of the amino acid sequences of Malabari and Attappady Black goats with that of the GenBank sequence was done using Clustal Omega program. Five non-synonymous variations were seen in both Malabari and Attappady Black goat breeds when compared with the *C. hircus* protein sequence (Table 1).

3.2.5 Protein structure prediction

The secondary structure prediction by SOPMA indicated that Lf protein of both the goat breeds consisted of α helices, extended strands, β turns and random coils. Comparison of the secondary structures of MgLf and AgLf with that of *C. hircus* is shown in Table 2.

Table 1 Nucleotide and amino acid variations of *Lf* gene of Malabari (MG980401) and Attappady Black (MG980402) goats on comparison with *C. hircus* (NM_001285548)

Malabari goat with *Capra hircus*			Attappady Black goat with *Capra hircus*		
Nucleotide variation	Codon change	Amino acid variation	Nucleotide variation	Codon change	Amino acid variation
-	-	-	G272A	CGG→CGA	Syn(Arg)
G337T	CGG→CTG	Arg88Leu	G337T	CGG→CTG	Arg88Leu
A444C	AAG→CAG	Lys124Gln	A444C	AAG→CAG	Lys124Gln
T464C	GTT→GGC	Syn(Gly)	T464C	GTT→GGC	Syn(Gly)
C534T C535T	CCC→TTC	Pro154Phe	C534T C535T	CCC→TTC	Pro154Phe
T1143G	TTG→GTG	Leu357Val	T1143G	TTG→GTG	Leu357Val
G1102A	GGT→GAT	Gly414Asp	G1102A	GGT→GAT	Gly414Asp
-	-	-	A1787G	ACA→ACG	Syn(Thr)
T1985C	TTT→TTC	Syn(Phe)	T1985C	TTT→TTC	Syn(Phe)

Table 2 Features of secondary structure of Lf protein of Malabari and Attappady Black goats and *C. hircus* (NM_001285548)

Protein parameters	Malabari	Attappady Black	*Capra hircus*
α helices	197	200	231
β turns	138	141	147
Extended strands	73	75	78
Random coils	230	241	252

The fully automatic procedure on the SWISS-MODEL server was used to construct a 3D structural model of Lf protein of both the goat breeds (Figure 2). The predicted 3D structure will provide the basis for further structure–function studies of Lf.

3.2.6 Phylogenetic analysis

Phylogenetic analysis of the coding sequences of *Lf* gene of different species was done using MEGA version 6.0 software and a phylogenetic tree was constructed using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) (Figure 3) which revealed that Malabari and Attappady Black breeds belonged to the same clade as that of goat and were very closely related to sheep and antelope. The goats, sheep and antelope were found to be related to the clade formed by cattle, yak and buffalo. This clarified that all the members of the family *Bovidae;* cattle, buffalo, goat, sheep, antelope and yak shared a recent common ancestor. All these members of *Bovidae* shared a distant lineage with pig, horse and camel. Humans formed a separate clade.

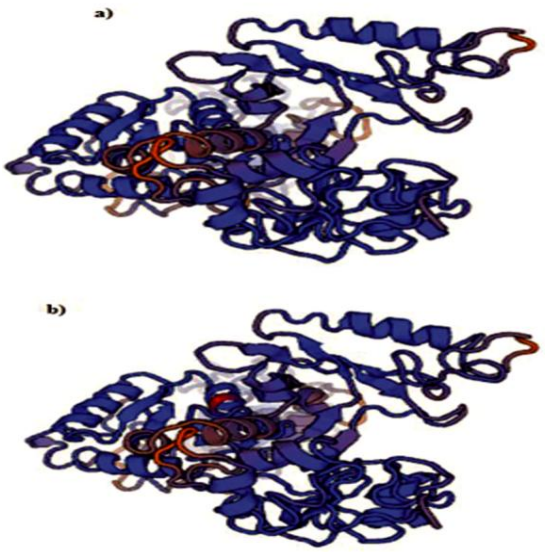

Figure 2 Three dimensional structural model of a) Malabari goat lactoferrin; b) Attappady Black goat lactoferrin

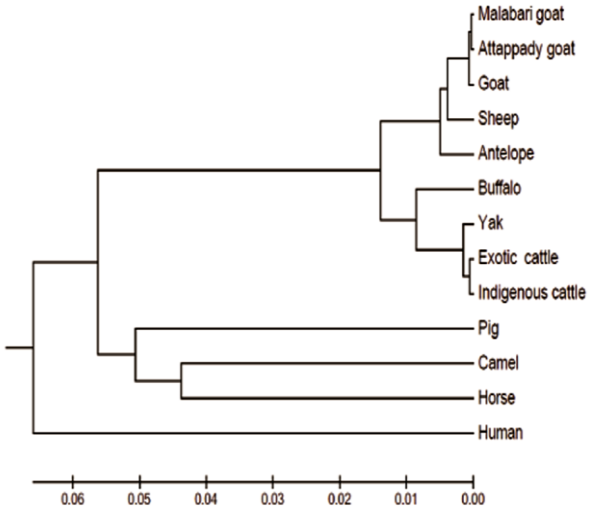

Figure 3 Phylogenetic analysis (UPGMA) of coding region of *Lf*gene of Malabari and Attappady Black goats with that of different mammalian species

4 Discussion

Lactoferrin is an iron binding glycoprotein with an array of molecular functions. Any significant variation in the nucleotide sequence of *Lf* gene could alter its biological properties. The native goat breeds of Kerala, Malabari and Attappady Black are disease resistant and adaptable to extreme agro-climatic conditions; hence their gene pool provides a valuable platform to explore the potentials of different bioactive peptides including Lf.

The total RNA extracted from MSC of Malabari and Attappady Black goat breeds was utilized for cDNA synthesis followed by amplification of *Lf* gene (about 2Kb size) and sequencing of the amplicons. The sequences of *MgLf* and *AgLf* which were of size 1914 bp and 1975 bp respectively were compared with *C. hircus Lf* nucleotide sequence retrieved from GenBank. The sequences of coding region of *Lf* gene of both Malabari and Attappady Black goat breeds revealed more than 99% identity with goat lactoferrin mRNA sequences in the database retrieved by BLASTn. In *MgLf* there were 8 nucleotide variations, out of which 3 were transversions and 5 were transitions while *AgLf* sequence showed a total of 10 variations comprising of 3 transversions and 7 transitions. Conceptualized translation of nucleotide sequence revealed 5 non-synonymous amino acid changes (p.Arg88Leu, p.Lys124Gln, p.Pro154Phe, p.Leu357Val, p.Gly414Asp) in both the breeds. At position 88 of the protein, arginine, a basic amino acid was replaced by leucine, a hydrophobic non-polar amino acid. At position 124 of the protein, lysine, a polar positively charged amino acid residue was replaced by glutamine, a polar amino acid with no charge in its side chain. The 154th residue proline, a non-polar amino acid was found to be replaced by phenylalanine, another non-polar amino acid. At position 357, leucine, a non-polar amino acid was found to be replaced by valine, another non-polar amino acid. The 414th residue, glycine, a polar amino acid with no charge on its side chain was found to be replaced by aspartic acid, a polar amino acid with negative charge.

Le Provost et al. (1994) reported the characterization of caprine *Lf* coding region of 2411bp size. Chen et al. (2007) expressed and purified goat Lf in *Pichia pastoris* and their amplified product was 2235 bp in size. In the present study the PCR amplified products were sequenced by primer walking technique and the end regions of the products could not be sequenced properly, that led to sequences shorter than the expected product size of 2048 bp. Pauciullo et al. (2010) sequenced the full reading frame (2127 bp) of *Lf* cDNA of Italian Nicastrese goat breed which was famous for its disease resistance and compared with Saanen goat breed *Lf* sequences. They found 11 nucleotide variations responsible for 5 amino acid changes which were the same as observed in the present study. Kang et al. (2008) noted 6 novel amino acid variations while analyzing the sequences of goat *Lf* gene.

The evolutionary relationship of *Lf* nucleotide sequences with 11 other species was generated. Analysis of the phylogenetic tree of *Lf* gene confirmed the presence of a common ancestor for the members of *Bovidae* family i.e., cattle, buffalo, goat, sheep, yak and antelope. Horse, camel and pig were distantly related to *Bovidae* while human beings formed a separate clade. Similar results regarding the phylogenetic relationship of *Lf* gene in cattle, goat, sheep, horse and camel were reported by many researchers (Yakubu et al., 2014; Akumbugu & Olusegun, 2017) by conducting *in silico* analysis of the sequences retrieved from the database.

Salient changes in the primary structure of proteins can alter their 3D structure and consequently change their functional properties. Kang et al. (2008) considered Lf amino acid variations within species to be related to antibacterial property or other biological activities. The non-synonymous amino acid variations revealed by the present study, at 88th and 414th positions of Lf seems to be highly relevant regarding to its biological functions as they are contributed by amino acids with totally different physical and chemical properties. These changes could be significant to the disease resistance exhibited by these autochthonous goat breeds. Moreover the synonymous amino acid variations though do not lead to altered protein structure, could alter the substrate specificity to mRNA binding and thereby down-regulate the translation process resulting in modified conformation of protein (Kimchi-Sarfaty et al., 2007). Lf is considered as one of the house keeping genes modulating iron homeostasis and immune

responses; hence special emphasis has to be given to the two unique synonymous variations of Attappady Black goats to rule out their contributions with respect to biological functions. Detailed population level studies on these nucleotide variations could throw light on the relevance of Lf in the herd immunity and disease resistance of these goats.

Conclusion

Lf is attributed with multiple biological functions besides taking part in iron homeostasis in the intestine. The present study reveals that the Lf of native goat breeds of Kerala, though very similar in genetic makeup to other members of the *Bovidae* family, possesses unique variations in its amino acid sequences which could be relevant for its antimicrobial and other biological properties.

Acknowledgements

The authors thank the Kerala Veterinary and Animal Sciences University and its Director, Academic and Research for the financial support and facilities provided.

Conflict of Interest

Authors would hereby like to confirm that there is no conflict of interests that could possibly arise.

References

Adlerova L, Bartoskova A, Faldyna M (2008) Lactoferrin: a review. Veterinary Medicine53: 457-468.

Akumbugu FE, Olusegun OA (2017) Genetic diversity of lactoferrin gene in silico on selected mammalian species. Biotechnology in Animal Husbandry 33: 171-180.

Alexander DB, Iigo M, Yamauchi K, Suzui M, Tsuda H (2012) Lactoferrin: an alternative view of its role in human biological fluids. Biochemistry and Cell Biology 90: 279–306.

Baker EN, Baker HM (2005) Molecular structure, binding properties and dynamics of lactoferrin. Cellular and Molecular Life Sciences62: 2531-2539.

Bennett RM, Kokocinski T (1978) Lactoferrin content of peripheral blood cells. British Journal of Haematology 39: 509–521.

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research 42: 252-258

Boutinaud M, Jammes H (2002) Potential uses of milk epithelial cells: a review. Reproduction Nutrition Development 42: 133-147.

Chen GH, Yin LJ, Chiang IH, Jiang SH (2007) Expression and purification of goat lactoferrin from *Pichia pastoris* expression system. Journal of Food Sciences 72: 67-71

Geourjon C, Deleage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Computer Applications in the Biosciences 11: 681-684.

Kim SJ, Yu DY, Pak KW, Jeong S, Kim SW, Lee KK (1998) Structure of the human Lactoferrin gene and its chromosomal localization. Molecules and Cells 8: 663-668.

Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM (2007) A silent polymorphism in the MDR1 gene changes substrate specificity. Science 315: 525–528.

Kang JF, Li XL, Zheng HQ, Zhou RY, Li LH, Zhao HY (2011) Sequence analysis of the lactoferrin gene and variations of G7605C T in 10 chinese goat breeds. Biochemical genetics 49: 63-72.

Kang JF, Li XL, Zhou RY, Li LH, Feng FJ, Guo XL (2008) Bioinformatics analysis of lactoferrin gene for several species. Biochemical Genetics 46:312–322.

Le Provost F, Nocart M, Guerin G, Martin P (1994) Characterization of the goat lactoferrin cDNA: assignment of the relevant locus to bovine U12 synteny group. Biochemical and Biophysical Research Communications 203:1324-1332.

Metz-Boutigue MH, Jolles J, Mazurier J, Schoentgen F, Legrand D, Spik G, Montreuil J, Jolles P (1984) Human lactotransferrin: amino acid sequence and structural comparisons with other transferrins. European Journal of Biochemistry145: 659–676.

Pauciullo A, Cosenza G, Nicodemo D, Gallo D, Mancusi A, Crepaldi P, Di Berardino D, Ramunno L (2010) Molecular cloning, promoter analysis and SNP identification of Italian Nicastrese and Saanen lactoferrin gene. Veterinary Immunology and Immunopathology 134: 279–283.

Schwerin M, SolinasToldo S, Eggen A, Brunner R, Seyfert HM, Fries R (1994) The bovine lactoferrin gene (LTF) maps to chromosome 22 and syntenic group U12. Mammalian Genome 5:486-489.

Seyfert HM, Tuckoricz A, Interthal H, Koczan D, Hobom G (1994) Structure of the bovine lactoferrin-encoding gene and its promoter. Gene 143:265-269.

Valenti P, Catizone A, Pantanella F, Frioni A, Natalizi T, Tendini M, Berlutti F (2011) Lactoferrin decreases inflammatory response by cystic fibrosis bronchial cells invaded with Burkholderia cenocepacia iron-modulated biofilm. International Journal of Immunopathology and Pharmacology 24: 1057–1068.

Yakubu A, Faith EA, Peters SO, Takeet MI, De Donato M, Imumorin IG (2014) In silico molecular analysis of the evolution and differentiation of lactoferrin gene in some ruminant and non-ruminant animals. Proceedings of the 37th annual conference of genetics society of Nigeria (gsn), Federal University, Lafia, 21st - 24th October, 2013.Pp.539-548.