Journal of Animal Research: v.7 n.6, p. 1123-1128. December 2017

DOI: 10.5958/2277-940X.2017.00168.1

*Usp*A Gene Based Characterization of *Escherichia coli* Strains Isolated from Different Disease Conditions in Goats

Anil K. Mishra[1], Desh Deepak Singh[2*], Gururaj Kumarsen[1], Geetika Gupta[1], Nitika Sharma[1], Naveen Kumar[3], Shivasharanappa Nayakvadi[4] and Souvik Paul[1]

[1]Division of Animal Health, ICAR-Central Institute for Research on Goats, Makhdoom, Farah, Mathura, Uttar Pradesh, INDIA
[2]Department of Veterinary Pathology, College of Veterinary Sciences, NDUA&T, Faizabad, Uttar Pradesh, INDIA
[3]Veterinary Type Culture Collection, ICAR- National Research Centre on Equines, Bagla Road, Hisar, Haryana, INDIA
[4]Division of Animal Science, ICAR-Central Coastal Agricultural Research Institute, Old Goa, Goa, INDIA

**Corresponding author: DD Singh; Email: drdd2005@gmail.com*

Received: 13 June, 2017 **Revised:** 26 Nov., 2017 **Accepted:** 29 Nov., 2017

ABSTRACT

Escherichia coli (*E. coli*) carry six universal stress protein (*usp*) genes: A, C, D, E, F and G, and the expression of these genes are triggered by various environmental stresses. The *usp*A gene is important for survival of *E. coli* during cellular growth, adhesion and motility. The present study was conducted to characterize three pathogenic *E. coli* strains isolated from the cases of diarrhoea, pneumonia and mastitis in goats. A polymerase chain reaction (PCR) was performed to amplify 884 bp open reading frame (ORF) of the *usp*A gene from the *E. coli* strains. The uspA amplicons of the three *E. coli* strains were sequenced, and compared with the published sequences in NCBI GenBank, and their phylogenetic relationships were analysed. The diarrheic strain showed significant variation in the nucleotide composition as compared to pneumonia and mastitis associated strains. In the ORF of *usp*A gene, silent mutations were noticed in the nucleotide sequence positions 27, 33, 207 and 316, which were not reflected phenotypically. Among the peptides, 'KHILIAVDLS' could be a putative candidate for use as epitope in diagnostics. Further, comprehensive studies on sequence analysis of the *usp*A gene will help us to understand the distribution, variability, and phylogenetic relationships of different pathogenic *E. coli* isolated from different disease conditions in goats.

Keywords: *E. coli*, *usp*A gene, goats, PCR, phylogenetic analysis

Escherichia coli (*E. coli*) is a large and diverse group of bacteria of the family *Enterobacteriaceae* associated with various infections in animals and humans. *E. coli* can contaminate, colonize, and subsequently cause infection at various extra-intestinal sites, and is a major cause of septicemia, peritonitis, abscesses, meningitis, and urinary tract infections in humans (Cavalieri *et al.,* 1984; Acharya, 1992). In animals, it is responsible for a variety of infections such as mastitis, metritis, colisepticemia, neonatal diarrhoea (white or yellow scours), urinary tract infections and pneumonia etc. *E. coli* scours is the single major cause of death in neonatal goat kids (Sharma and Dutta, 1964; Sherman, 1987; Radostits *et al.*, 1999). *E. coli* is also considered to be one of the most frequent causes of clinical mastitis in dairy animals including goats (Sipka, 2014; Hegde *et al*., 2013). It is also reported as predominant cause of pneumonia in goat and sheep (Sambyal *et al*., 1980; Sharma *et al*., 1991). Universal stress protein (*usp*) superfamily comprises of a conserved group of proteins that are found in the genomes of bacteria, archaea, fungi, protozoa and plants (Nachin *et al.,* 2005; Siegele, 2005). The biological and biochemical functions of the majority of these proteins are not known (Siegele, 2005). *Escherichia coli* has six *usp* genes namely A, C, D, E, F and G, and the *usp*A gene is important for the survival of *E. coli* during cellular growth, adhesion and motility (Nachin *et al.,* 2005). Synthesis of *usp* protein is induced in response to stresses such as heat shock, nutritional starvation, osmotic pressure and the presence of toxic agents etc. (Nyström and Neidhardt, 1992), however,

cold shock doesn't induce synthesis of uspA (Nyström and Neidhardt, 1993). In the present study, *E. coli* were isolated from goats affected with diarrhoea, pneumonia and mastitis, and were characterized by sequencing and phylogenetic analysis of their *usp*A genes.

MATERIALS AND METHODS

Samples collection

Samples were collected aseptically in ice from a total of 249 goats including 149 kids with diarrhoea, 85 goats with clinical or subclinical mastitis and 15 pneumonic lungs during clinical examination and necropsy of animals at the goat farms of ICAR- Central Institute for research on Goats, Makhdoom, Mathura, Uttar Pradesh. The bacteriological isolation from all samples was done as per the method described by Cowan and Steel (1975).

PCR amplification of *usp*A gene of *E. coli*

For PCR amplification of *usp*A gene of *E. coli*, the genomic DNA was extracted by hot-chill method as described by Yang *et al.* (2008). Briefly, the cell pellet was suspended in 200 µl of Tris-EDTA buffer, and then placed in a boiling water bath for 2 min followed by freezing at -70° for 3 min. Three freeze-thaw cycles were performed, and then centrifugation was done at 10000 g for 5 min. ~~A~~ 50 µl of the supernatant was transferred to a sterile tube, and stored at -20° till further use. Species specific primers (F-5'-CCGATACGCTGCCAATCAGT-3' and R-5'-ACGCAGACCGTAAGGGCCAGAT-3') were used to amplify the *usp*A genes of *E. coli* (Osek, 2001; Rajput *et al*., 2014). PCR reaction was performed in a volume of 25 µl containing 1.5 units of *Taq* DNA polymerase, 1.5 mM $MgCl_2$, 200 µM of each deoxynucleoside triphosphate, 10 pmol of each primer, 50 ng of template DNA and nuclease free water up to 25 µl.PCR amplifications were done in a thermal cycler (GMI, USA) with initial DNA denaturation step at 95 °C for 2 min followed by 30 cycles beginning with 0.5 min of denaturation at 94 °C, 1 min of primer annealing at 58°C, and 1 min of extension at 72 °C. The final extension step was performed at 72 °C for 5 min. The amplified PCR products (10 µl aliquots) were analyzed by electrophoresis in 1.5% agarose gels in tris acetate EDTA buffer at 100 V. The gels were stained with ethidium bromide and photographed under ultraviolet light using a commercial documentation system (Alpha Innotech, USA).

Sequence analysis and epitope prediction

DNA sequencing was carried out for *usp*A gene of three virulent *E. coli* isolates which were isolated from goats affected with diarrhoea, pneumonia and mastitis respectively. The amplified PCR amplicons of the above isolates were gel purified using Gen Elute Gel Extraction Kit (Sigma, USA), and submitted for an automated sequencing from both directions (Merck Bioscience, India) using the PCR primers as the sequencing primers. The sequence data were submitted to NCBI GenBank for the accession numbers. The *usp*A gene sequences of *E. coli* and their deduced amino acid sequences were aligned with the other sequences from GenBank database using ClustalW Sequence Alignment Module. Sequences were compiled and analyzed using Bio-Edit (Hall *et al*., 2011), and phylogenetic tree was generated using MEGA 6.0 (Tamura *et al*., 2013). The protein structures were predicted using the web-based RaptorX online tool (Källberg *et al.*, 2012). The ORF of uspA gene was translated and the linear B-cell epitopes were predicted using Tongaonkar and Kolaskar antigenicity prediction method (Kolaskar and Tongaonkar, 1990) computed with the online tool http://immunax.dfci.harvard. edu/tools/antigenic.html.

RESULTS AND DISCUSSION

E. coli has been reported as important causative agents of mastitis, pneumonia and diarrhoea in animals including goats (Sharma and Dutta, 1964; Sharma *et al*., 1991; Hegde *et al*., 2013). Isolation and identification *E. coli* was done on the basis of cultural, morphological and biochemical characteristics (Hedge *et al*., 2013). The organisms were Gram negative short rods, catalase positive, oxidase negative and IMViC reactions as +ve, +ve, -ve,-ve. The organisms showed lactose fermentation on MacConkey agar (Fig. 1) and characteristic green metallic sheen on EMB agar (Fig. 2). In the current study, three pathogenic *E. coli* strains one each from diarrhoea, pneumonia and mastitis affected goats were amplified and sequenced. The *usp*A genes from the *E. coli* strains were successfully amplified using species specific primers. PCR Amplification resulted in to a single amplicon of 884-bp (Osek, 2001; Rajput *et al*., 2014) as illustrated (Fig. 3).

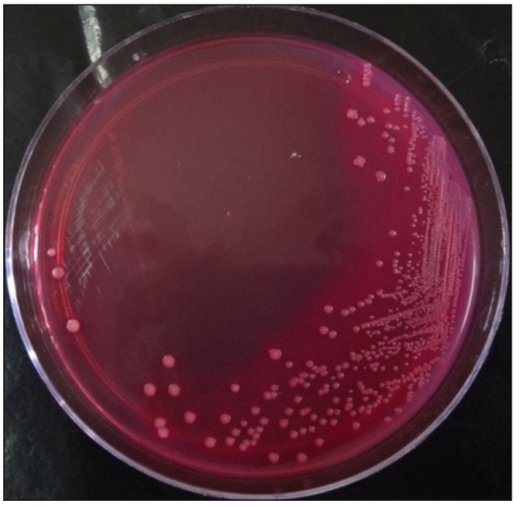

Fig. 1: *E. coli* on MacConkey agar showing lactose fermentation

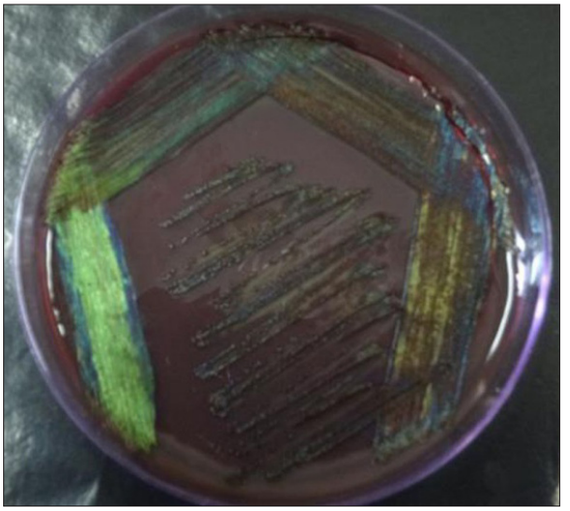

Fig. 2: *E. coli* on EMB agar showing characteristic green metallic sheen

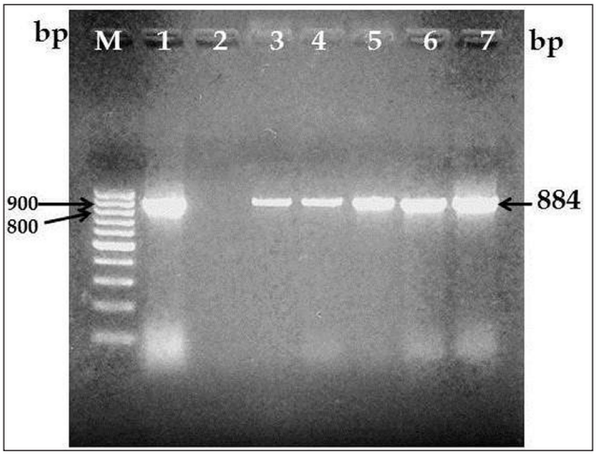

Fig. 3: PCR amplification of UspA gene of *E. coli*: Lane M: DNA ladder; Lane 1: positive control; Lane 2: negative control; Lane 3, 4, 5, 6, 7: amplified PCR products

Nucleotide sequences of *usp*A genes from diarrheic (ECD1), pneumonic (ECP) and mastitic (ECM) *E. coli* were submitted to NCBI GenBank for the accessions, and subsequently assigned as KF765738, KF765740 and KF765739 respectively. The sequenced information was further analysed for phylogeny, and compared with the existing databases of other *E. coli* strains to identify potential changes in the nucleotide composition (Fig. 4).

Tajima's test of molecular hypothesis was conducted to identify evolutionary divergence in the *uspA* gene coding regions sequenced, and 756 identical sites were found. The diarrheic strain showed significant variation in the nucleotide composition as compared to pneumonic and mastitic strains. Significant differences were noticed with ECD1 showing 9 unique differences in the coding region, followed by ECM with 4 and ECP with 1 unique difference respectively. This is due to the fact that, the diarrheic strain was in a different clade as compared to the O157:H7 strain which means it is not verotoxic but highly pathogenic as observed clinically and pathologically. Protein structure prediction was done based on template based tertiary structure modelling (Fig. 5 & 6).

Protein structure models of universal stress protein of ECD1 and ECM strains were designed based on the translation of *uspA* gene coding region (Fig. 7) and structure predicted using the web-based RaptorX online protein prediction tool (Källberg *et al.,* 2012). The findings of present study corroborated with the amino acid composition reflecting changes in positions 142 and 144, with change from lysine to glutamic acid or *vice versa* respectively, between the strains (Fig. 8). The amino acid composition was deduced based on the primary aligned sequence information from open reading frame of *usp*A gene. In the ORF of *usp*A gene, silent mutations were noticed in the nucleotide sequence positions 27, 33, 207 and 316, which were not reflected phenotypically.

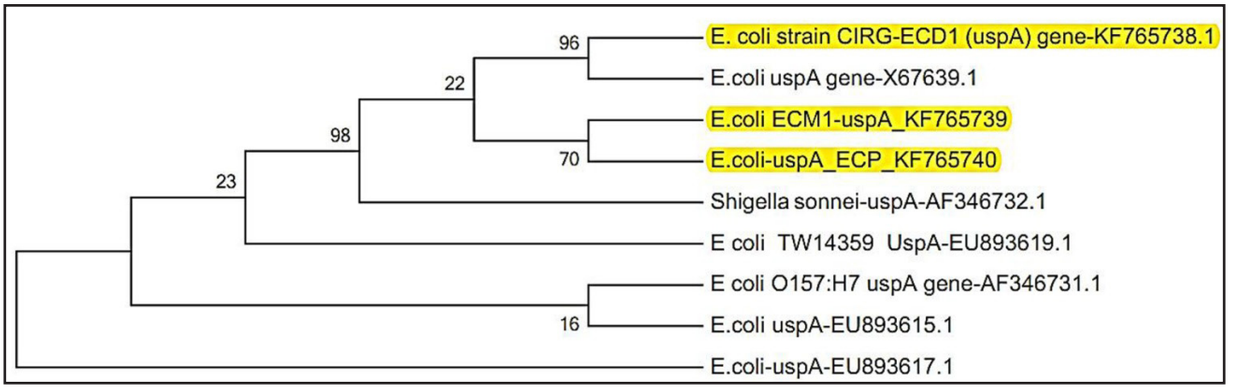

Fig. 4: Phylogenetic tree constructed for universal stress protein based on the coding sequences by neighsbour joining method

	1	2	3	4	5	6	7	8	9
1. E. coli strain CIRG-ECD1 (uspA) gene-KF765738.1									
2. E.coli ECM1-uspA_KF765739	0.014								
3. E.coli-uspA_ECP_KF765740	0.012	0.002							
4. E.coli uspA gene-X67639.1	0.012	0.021	0.019						
5. Shigella sonnei-uspA-AF346732.1	0.012	0.007	0.005	0.019					
6. E coli O157:H7 uspA gene-AF346731.1	0.019	0.014	0.012	0.026	0.012				
7. E coli TW14359 UspA-EU893619.1	0.019	0.014	0.012	0.026	0.012	0.000			
8. E.coli-uspA-EU893617.1	0.019	0.014	0.012	0.026	0.012	0.000	0.000		
9. E.coli uspA-EU893615.1	0.019	0.014	0.012	0.026	0.012	0.000	0.000	0.000	

Fig. 5: Pair wise distance matrix of ECD1 with the other isolates of *E. coli*

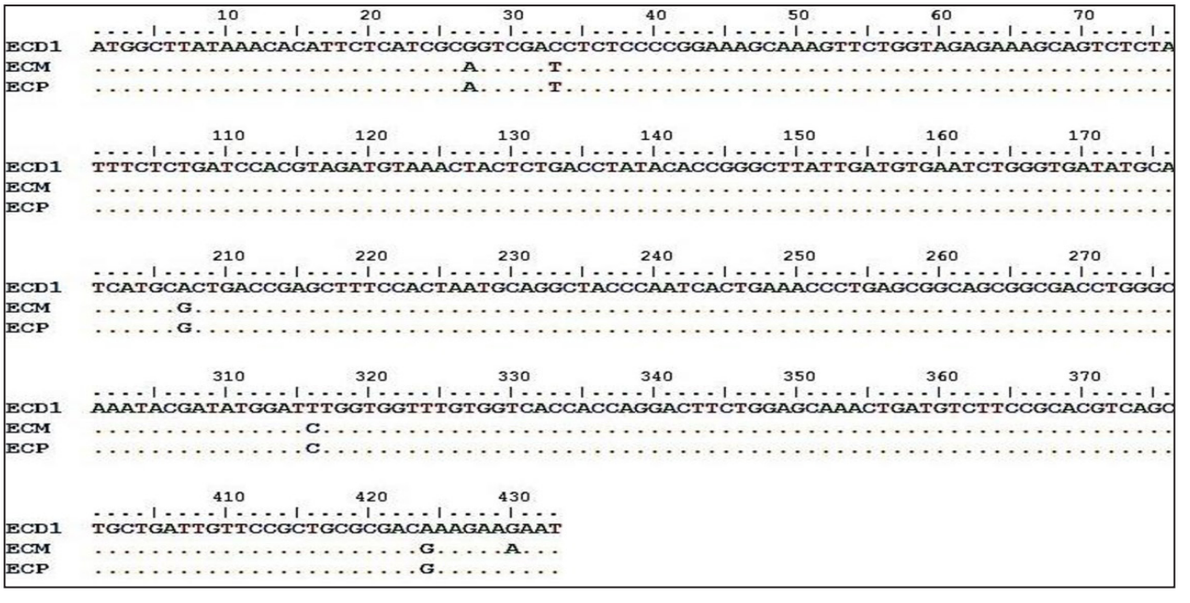

Fig. 6: Dot plot of open reading frame of uspA gene of three isolates of *E. coli*

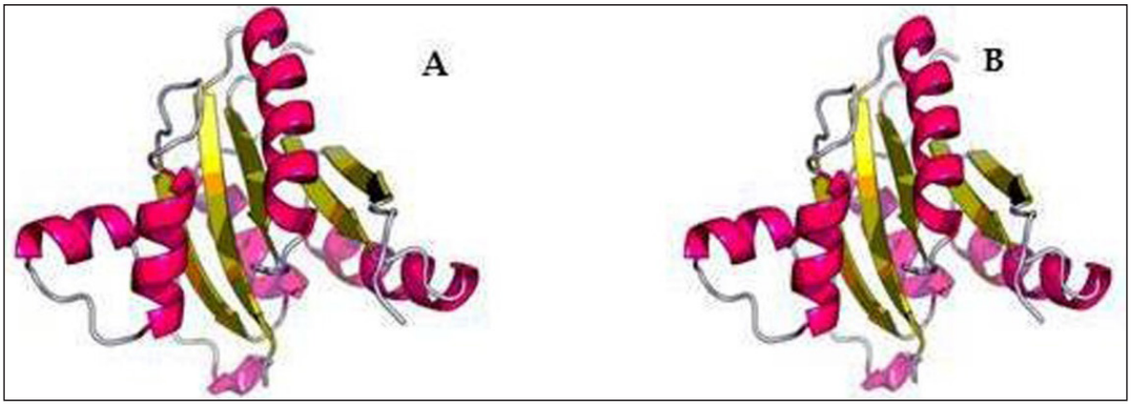

Fig. 7: Protein structure model of *uspA* of ECD1 (A) and ECM (B)

```
          10        20        30        40        50        60        70        80        90        100
     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ECD1 MAYKHILIAVDLSPESKVLVEKAVSMARPYNAKVSLIHVDVNYSDLYTGLIDVNLGDMQKRISEETHHALTELSTNAGYPITETLSGSGDLGQVLVDAIK
ECP  ....................................................................................................
ECM  ....................................................................................................

          110       120       130       140
     ....|....|....|....|....|....|....|....|....
ECD1 KYDMDLVVCGHHQDFWSKLMSSARQLINTVHVDMLIVPLRDKEE
ECP  ..........................................E..
ECM  ..........................................E.K
```

Fig. 8: Dot plot for amino acid residues of uspA protein of three isolates of *E. coli*

However, two nucleotide positions including 424 and 430 showed mutation in codons reflected phenotypically as amino acids KEE, EEE and EEK in ECD1, ECP and ECM respectively. The amino acid positions 142 showed this change from Lysine to glutamic acid in ECP and ECM, whereas 144th position showed a change from glutamic acid to lysine in ECM. The predicted epitopes with their antigenicity is given in Table 1.

Table 1: Antigenic epitope prediction by using Kolaskar and Tongaonkar method showing peptide sequences

No.	Start	End	Peptide	Length
1	4	13	KHILIAVDLS	10
2	15	27	ESKVLVEKAVSMA	13
3	31	54	NAKVSLIHVDVNYSDLYT GLIDVN	24
4	67	72	HHALTE	6
5	91	99	LGQVLVDAI	9
6	104	112	MDLVVCGHH	9
7	122	140	SARQLINTVHVDMLIVPLR	19

Variation at the nucleotide level may not directly affect the function as evidenced by the protein structure prediction that showed no significant difference with respect to the folding and conformation, which are more conserved. Further, epitope mapping of immunodominant epitopes on the universal stress protein and contact binding could establish functional properties with respect to the virulence. Protein folding is more conserved in nature although differences do exist in the nucleotide and amino acid composition. The *Usp*A proteins have multifaceted function with overlapping activities with the other paralogs viz., *Usp*C, *Usp*D, *Usp*G, *Usp*E etc. Although the functions of *Usp*A family was not defined unequivocally, but it is evident from the earlier findings that it is essential for the bacteria to cope with the cellular defense mechanism and oxidative stress (Nachin *et al.,* 2005; Siegle, 2005). The *Usp*A gene can be targeted for many immunological assays, and can be included as a candidate for multiple subunit vaccines. Among the peptides, 'KHILIAVDLS' could be a putative candidate for use as epitope in diagnostics. This would make an ideal

epitope for preparation of multiple subunit vaccines which could augment its efficacy by including it with other peptides of virulence factors identified from pathogenic *E.coli*. The immunodominant epitopes can also be used for developing peptide vaccine for control of neonatal diarrhoea after conducting *in silico* docking experiments. Moreover, these peptides could be used as an antigen for detection of specific antibodies against *E.coli* by lateral flow assay techniques.

ACKNOWLEDGEMENTS

Authors are thankful to the Director, ICAR-Central Institute for Research on Goats, Makhdoom, Farah, Mathura-281122, Uttar Pradesh (India) for providing necessary facilities to carry out the research.

REFERENCES

Acharya, V.N. 1992. Urinary tract infection--a dangerous and unrecognized forerunner of systemic sepsis. *J. Postgrad. Med.,* **38**(2): 52-4.

Cavalieri, S.J., Bohach, G.A. and Snyder, I.S. 1984. *Escherichia coli* α-hemolysin: Characteristics and probable role in pathogenicity. *Microbiol. Rev*., **48**: 326-343.

Cowan, S.T. and Steel, K.J. 1975. Manual for the identification of medical bacteria. Cambridge University Press, Cambridge.

Hall, T. 2011. Bio Edit: An important software for molecular biology. *GERF Bull. Biosci*., **2**: 60-61.

Hegde, R., Isloor, S., Prabhu, K.N., Shome, B.R., Rathnamma, D., Suryanarayana, V.V.S., Yatiraj, S., Prasad, C.R., Krishnaveni, N., Sundareshan, S., Akhila, D.S., Gomes A.R. and Hegde, N.R. 2013. Incidence of subclinical mastitis and prevalence of major mastitis pathogens in organized farms and unorganized sectors. *Indian J. Microbiol.,* **53**(3): 315-320.

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H. and Xu, J. 2012. Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc*., **7**(8): 1511-1522.

Kaper, J.B., Nataro, J.P. and Mobley, H.L. 2004. Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol*., **2**: 123–140.

Kolaskar, A.S. and Tongaonkar, P.C. 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*., **276**(1): 172-174.

Nachin, L., Nannmark, U. and Nyström, T. 2005. Differential roles of the universal stress proteins of *Escherichia coli* in oxidative stress resistance, adhesion, and motility. *J. Bacteriol*., **187**(18): 6265-6272.

Nystrom, T. and Neidhardt, F.C. 1993. Isolation and properties of a mutant of *Escherichia coli* with an insertional inactivation of the UspA gene, which encodes a universal stress protein. *J. Bacteriol*., **175**:3949–3956.

Nyström, T. and Neidhardt, F.C. 1992. Cloning, mapping and nucleotide sequencing of a gene encoding a universal stress protein in *Escherichia coli*. *Mol. Microbiol*., **6**: 3187–3198.

Osek, J., 2001. Multiplex polymerase chain reaction assay for identification of enterotoxigenic *Escherichia coli* strains. *J. Vet. Diagn. Invest*., **13**(4): 308-311.

Radostits, O.M., Gay, C.G, Blood, D.C. and Hinchcliff, K.W. 1999. Diseases caused by bacteria. In: Veterinary Medicine: A textbook of the diseases of cattle, sheep, pigs, goats and horses. 9th Ed., Saunders, London, pp. 753–822.

Rajput, S.K., Gururaj, K., Tiwari, U. and Singh, G. 2014. Study of the characterization of *E. coli* isolates in goat kids. *Indian Res. J. Genet. Biotech.,* **6**(1): 324-329.

Sambyal, D.S., Banerjee, M., Baxi, K.K. and Gupta, P.P. 1980. Bacteriology of pneumonia of sheep and goats. *J. Res.,* **17**: 89-91.

Sharma, P.K., Boro, B.R. and Borah, P. 1991. Incidence of caprine pneumonia and associated bacterial species.*Indian J. Anim. Sci*., **64**(1): 54-55.

Sharma, R.M. and Dutta, S.K. 1964. Calf mortality: Its incidence and etiology in India and possible method of control. *Indian J.* Anim. Sci., **11**: 65-66.

Sherman, D.M. 1987. Causes of kid morbidity and mortality: An overview. In Proceedings of the IV International Conference on Goats, pp. 335-349.

Siegele, D.A. 2005. Universal stress proteins in *Escherichia coli*. *J. Bacteriol*., **187**(18): 6253-6254.

Sipka, A., Klaessig, S., Duhamel, G.E., Swinkels, J., Rainard, P. and Schukken, Y. 2014. Impact of intramammary treatment on gene expression profiles in bovine *Escherichia coli* Mastitis. *PLoS One*, **9**(1): e85579.

Sousa, C.P.D. 2006. *Escherichia coli* as a specialized bacterial pathogen. *Revista De Biologia E Ciências Da Terra*., **6**(2): 341-352.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol*., **30**(12): 2725-2729.

Yang, J.L., Wang, M.S., Cheng, A.C., Pan, K.C., Li, CF and Deng, SX 2008. A simple and rapid method for extracting bacterial DNA from intestinal microflora for ERIC-PCR detection. *World J. Gastroenterol*., **14**(18): 2872.