

Association of SNPs in the coding regions of *CD4* gene with mastitis susceptibility and production traits in dairy cattle

Saqib Zeb^{1,3} Nawab Ali^{1,3} Sadaf Niaz³ Arsalan Rasheed³ Irfan Khattak²

Naimat Ullah Khan² Yachun Wang^{1*} Tahir Usman^{2*}

Abstract

CD4 gene is reported to have significant role in various inflammatory conditions including mastitis. Therefore, the current study was designed to screen out single nucleotide polymorphisms (SNPs) in the bovine *CD4* gene and to evaluate its effects on mastitis indicator and production traits. Total 201 milk and blood samples were collected from dairy cattle maintained at government dairy farms of Khyber Pakhtunkhwa, Pakistan. Milk samples were used for DHI test (milk composition and somatic cell count) and the blood for DNA extraction. SNPs were initially detected in the pool DNA samples followed by screening in the whole cohort. The association of SNPs with the studied phenotypic traits was analysed using the GLM procedure of SAS 9.4. Total four SNPs including three in exon 2 (SNP 1 A>C rs110955838, SNP 2 T>G rs134722546, SNP 3 T>C rs135440143) and one in exon four (SNP4 G>A, rs134369392) were revealed. The results showed that all the SNPs were significantly associated with frequency of clinical mastitis incidence ($P<0.05$). Only SNP4 (G>A) was found to have significant association with annual milk yield ($P<0.05$). Furthermore, breed and herd showed significant association with mastitis indicator and production traits ($P<0.005$). The parity wise milk production results revealed highest annual milk yield in cattle with first parity (3649.2 L) followed by 2nd (3412.21 L), 3rd and above parities (3112.32 L). The present study suggests that *CD4* gene should be considered as a candidate gene and the identified SNPs could be worthy molecular markers for selection of dairy cattle with genetic resistance to mastitis development.

Keywords: *CD4* gene, dairy cattle, mastitis, molecular markers, production traits

¹College of Animal Sciences and Technology, China Agriculture University, 100193 Beijing, PR China; Key Laboratory of Animal Genetics, Breeding and Reproduction.

²College of Veterinary Sciences and Animal Husbandry, Abdul Wali Khan University Mardan

³Department of Zoology, Abdul Wali Khan University Mardan

*Correspondence: tahircau@gmail.com, tahirusman@awakum.edu.pk, wangyachun@cau.edu.cn (T. Usman and Y. Wang)

Introduction

Bovine mastitis is the inflammatory disease of cow udder and teats, in which udder becomes inflamed and swollen and the milk color and texture is changed. This infection results in the alteration of milk quantity and quality (Zhao and Laccase, 2008). It is considered as the most significant commercial syndrome of dairy cattle across the world with annual commercial losses of £300 million recorded in the United Kingdom (UK). Researches carried out by different scientists have revealed significant association of this inflammatory condition with milk performance traits and mastitis indicator traits. Besides many genes, *CD4* gene is discovered to play key role in inflammatory condition like mastitis. Udder tissue of cattle possesses both cellular and humoral immune responses against the pathogens that results in mastitis (Sordillo *et al.* 1997; Rainard *et al.* 2006). T lymphocytes are reported to play key role in cell mediated immunity in inflammatory conditions (Fabbri *et al.* 2003). On the basis of differences in their cell surface marker molecules, T lymphocytes are classified into four subpopulations i.e. $CD4^+CD8^-$, $CD4^+$, $CD8^+$, $CD4^+CD8^+$ cells (Summerfield *et al.* 1996). The purpose of the *CD4* proteins produced by *CD4* gene in $CD4^+$ lymphocytes cells is to augment the activation and cloning of T cells. *CD4* molecules are chiefly exposed on the cell membrane of mature $CD4^+$ cells. The most important role of $CD4^+$ cells is to help other immune cells to perform functions in immune reactions, including the maturation of B cells and activation of cytotoxic T lymphocytes (CTL). *CD4* protein combine with MHC class II molecules hence take part in cell mediate immunity. Hence both *CD4* protein and $CD4^+$ T cells play central role in developing immune resistance against mastitis (Perez-Diez *et al.* 2007).

Different researches show that the SNPs exert significant effect on the production, reproduction, health and difference performance traits. Studies have shown that immune response against bacteria and other pathogens is affected by single nucleotide polymorphism in *CD4* gene (Shelley *et al.* 2003; Leyva-Baca *et al.* 2008). SNPs in *CD4* gene have significant association with mastitis related traits and production traits (Oyugi *et al.* 2009; He *et al.* 2011). Polymorphism in *CD4* gene were found to have significant association with fat percentage ($P < 0.01$) and were found non-significant with SCC and SCS ($P > 0.05$) (Usman *et al.* 2016; Usman *et al.* 2017). Based on previous studies, the present study was designed with the objective to screen out SNPs in bovine *CD4* gene and evaluate its association with milk production and mastitis indicator traits in dairy cattle maintained dairy farms of Khyber Pakhtunkhwa, Pakistan.

Materials and Methods

The study was approved by the institutional ethical committee and care was taken to comply with the 3R concept.

Samples collection: Samples were collected according to the animal welfare and care guidelines of the college of veterinary sciences and animal husbandry. Initially

214 blood and milk samples were collected from various dairy cattle farms i.e. Khazana, Surezi, Agriculture University Peshawar (AUP) and cattle breeding and dairy farm Harichand in Khyber Pakhtunkhwa, Pakistan. Finally 201 samples having complete records were processed for further analysis. These samples were collected from Holstein-Friesian (HF), Jersey (J), an indigenous breed named Achai (A), few crossbred cows (HF x Jersey and Achai x Jersey). Records of each cow were obtained from these dairy farm including the record of total milk yield, incidence of clinical mastitis in last three years and production record. The incidence of clinical mastitis was converted to frequency of clinical mastitis (FOM) per year in each cow.

Milk composition and somatic cell count: Milk composition (percentage of proteins, fats and lactose) was determined through EcoMilk analyser. The SCC was calculated from the milk using direct microscopic cell count method using Geimsa stain. Each SCC was then converted into somatic cell score (SCS) by the following formula:

$SCS = \log_2 [SCC/100] + 3$, the unit of SCC is 1,000 cells/ml

Processing of DNA: Extraction of DNA was performed from each blood sample by means of TianGen Kit following manufacturer's protocol. Presence of DNA was confirmed through gel electrophoresis technique. A pool DNA of 30 randomly selected animals was constructed. Primers for exon one, two and four of *CD4* gene were designed using the reference sequences from NCBI through primer3 software. Primers designed for the study had the following sequence.

Exon 1 and 2 Forward 5'- AGATGAACGCCTGCTAACTG -3/
Reverse 5'- TCCCTGAGACTCTGGCTGAT -3/
Exon 4 Forward 5'- ATGGACGACAGAGCGATACC -3/
Reverse 5'- GCGTGTCTCTTAGGGCAGA -3/

The selected exons of *CD4* gene were amplified through polymerase chain reaction (PCR) using the following conditions. Denaturation was done at 95 °C for 7 minutes, followed by 35 cycles of melting at 95 °C for 30 seconds, annealing at 56 °C for 30 seconds, elongation at 72 °C for 30 seconds, and final extension at 72 °C for 10 minutes.

Sequencing and statistical analysis: PCR products for *CD4* gene were sent to Tsingke Biological Technology, Beijing (China) for sequencing. The sequencing results were analysed by CodonCode aligner (V 8.0.2) for identification of SNPs. The DNA samples of the whole population (n=201) under the study were sent to Generay Biotechnology, Beijing, China. The following Generalized linear model (GLM) in SAS (V 9.4) was employed for association analysis of SNPs with production traits and mastitis related traits (SCC, SCS, and FOM):

$$Y_{ijklm} = \mu + a_i + \beta_j + \gamma_k + \delta_l + \lambda_m + e$$

Where, Y_{ijklm} show the phenotype SCC, SCS, FOM, fat %, Proteins %, or sugar %; μ shows overall mean; a_i shows fixed genotypic effect; β_j shows fixed herd effects, γ_k shows fixed parity effects; δ_l shows fixed calving season effect; λ_m shows fixed year of calving effect and e shows random residual error.

Linkage disequilibrium (LD): LD for the identified SNPs were analysed by Haploview (V 4.2).

Results

SNPs discovery and genotypes: Primers were used for polymerization of exon 1, exon 2 and exon 4 regions of *CD4* gene. DNA sequencing revealed three SNPs in

exon 2 and one in exon 4. Information related to the SNPs and gene is provided in the table 1.

Allele and genotype frequencies and Hardy-Weinberg equilibrium (HWE) test: Hardy and Weinberg equilibrium (HWE) test and allele and genotype frequencies revealed that genotypic frequencies of the studied SNPs of *CD4* gene are consistent with HWE ($P>0.05$) (Table 2).

Table 1 Information of the SNPs found in *CD4* gene

SNP #	Gene name	SNP location	Base Position*	Variable type	Variant ID	Consequence type
SNP 1	CD4	Exon 2	103647315	A>C	rs110955838	missense
SNP 2	CD4	Exon 2	103647216	T>G	rs134722546	missense
SNP 3	CD4	Exon 2	103647213	T>C	rs135440143	missense
SNP 4	CD4	Exon 4	103638641	G>A	rs134369392	missense

*According to *Bos taurus* ARS-UCD 1.2

Table 2 Allele and genotype frequencies and HWE test of SNPs in *CD4* gene

SNPs	Genotype	N	Genotype frequency	Allele	N	Allele frequency	HWE test *
SNP 1	AA	43	0.21	A	176	0.43	0.19
	AC	90	0.44	C	226	0.56	
	CC	68	0.33				
SNP 2	TT	43	0.21	T	175	0.43	0.15
	TG	89	0.44	G	227	0.56	
	GG	69	0.34				
SNP 3	TT	43	0.21	T	175	0.43	0.15
	TC	89	0.44	C	227	0.56	
	CC	69	0.34				
SNP 4	GG	47	0.23	G	186	0.46	0.25
	GA	92	0.45	A	216	0.53	
	AA	62	0.30				

* $P>0.05$ means consistent with HWE

Linkage Disequilibrium: The Haploview results showed strong LD for SNP 1 with SNP2, SNP 1 with SNP 3 and SNP 2 and SNP 3 (Figure 1).

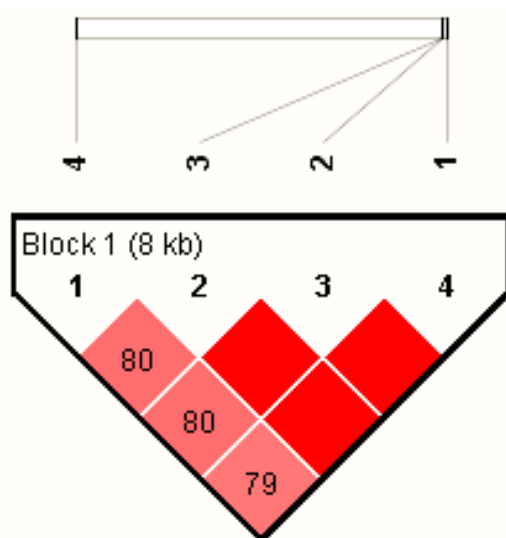


Figure 1 Linkage Disequilibrium plot of the SNPs in *CD4* gene (SNP 1 and SNP 2 $r^2=0.99$, SNP 1 and SNP 3 $r^2=0.99$, SNP 1 and SNP 4 $r^2=0.57$, SNP 2 and SNP 3 $r^2=1$, SNP 2 and SNP 4 $r^2=0.58$, SNP 3 and SNP 4 $r^2=0.58$)

Association and SNPs effect: Results of association analysis of SNPs with the studied phenotypes showed that all SNPs found in exon 2 and exon 4 regions were significantly associated ($P<0.05$) with FOM in overall studied population while SNP 4 found in exon 4 revealed significant association with AMY (Table 3).

Effect of breed, herd and parity on phenotypes: Association analysis among different breed showed significant differences for mastitis-related traits (SCC and FOM) and production related traits (Lactose percentage, protein percentage and annual milk yield). The results revealed that HF breed showed highest SCC and FOM and Achai breed was found with lowest values. In production traits the HF breed was found to have highest annual milk yield and Achai breed showed lowest annual milk yield (Table 4).

Herd-wise association analysis revealed significant differences among herds for mastitis-related traits (SCC, SCS and FOM) and production traits (Protein percentage, lactose percentage and annual milk yield) (Table 4).

The statistical analysis for the effect of number of parities revealed no significant differences among cows with different parities (Table 4).

Table 3 Effect of SNPs in *CD4* gene on mastitis related and production traits

Marker	Genotype	SCC	SCS	FOM	Fat%	Lactose%	Protein%	AMY
SNP 1	AA	191.28±82.59	3.26±0.36	0.54±0.13^a	3.79±0.47	4.47±0.13	2.98±0.08	4164.52±228.48
	AC	230.28±68.27	3.18±0.30	0.25±0.11^b	3.63±0.39	4.63±0.10	3.15±0.06	3346.92±188.89
	CC	219.12±69.28	3.18±0.31	0.19±0.11^b	3.97±0.39	4.74±0.11	3.15±0.06	3004.65±191.67
	<i>P</i> value	0.36	0.77	0.02	0.52	0.55	0.49	0.10
SNP 2	TT	191.28±82.48	3.26±0.36	0.54±0.14^a	3.79±0.47	4.47±0.13	2.98±0.08	2999.12±228.40
	TG	231.74±68.29	3.18±0.30	0.26±0.11^b	3.58±0.39	4.63±0.11	3.15±0.08	3356.05±189.1
	GG	217.39±69.27	3.17±0.31	0.18±0.14^b	4.04±0.39	4.75±0.11	3.14±0.07	4164.53±191.81
	<i>P</i> value	0.27	0.69	0.02	0.38	0.43	0.48	0.09
SNP 3	TT	191.28±82.48	3.26±0.36	0.54±0.14^a	3.79±0.47	4.47±0.13	2.98±0.08	2999.12±228.40
	TC	231.74±68.29	3.18±0.30	0.26±0.11^b	3.58±0.39	4.63±0.11	3.15±0.08	3356.05±189.1
	CC	217.39±69.27	3.17±0.31	0.18±0.14^b	4.04±0.39	4.75±0.11	3.14±0.07	4164.53±191.81
	<i>P</i> value	0.27	0.69	0.02	0.38	0.43	0.48	0.09
SNP 4	AA	264.92±71.90	3.42±0.32	0.20±0.12^a	3.79±0.41	4.74±0.11	3.14±0.07	3188.73±197.05^b
	AG	204.35±66.92	3.06±0.29	0.23±0.11^b	3.71±0.38	4.64±0.10	3.16±0.06	3304.59±183.42^b
	GG	183.51±80.96	3.18±0.36	0.54±0.13^a	3.91±0.46	4.49±0.13	2.98±0.07	3891.28±221.87^a
	<i>P</i> value	0.64	0.99	0.02	0.27	0.63	0.54	0.03

*SCC means Somatic cell count, SCS means Somatic cell score, FOM means frequency of mastitis occurrence, AMY means annual milk yield

*Values with bold character are significantly different ($P < 0.05$)

Table 4 Effects of various breed, herd and parity on studied mastitis related and production traits

		SCC	SCS	FOM	LP	PP	FP	AMY
Breeds	Holstein-Frisian	239.9±39	3.35±0.15	0.42±0.07	4.53±0.06	3.03±0.04	2.69±0.12	4531±151
	Jersey	297.62±62	3.66±0.24	0.15±0.04	4.82±0.08	3.15±0.04	2.35±.20	2512±100
	Achai	74.14±13	1.97±0.25	0.15±0.05	4.67±0.01	3.26±0.07	2.44±0.23	1378±59
	Crossbred 1	207.14±55	3.61±0.48	0.14±0.10	4.79±0.18	3.35±0.09	2.85±0.46	3922±429
	Crossbred 2	147.36±28	3.05±0.29	0.19±0.06	4.80±0.10	3.16±0.06	2.15±0.24	2127±129
	<i>P</i> value	<0.01	0.11	<0.01	<0.01	<0.01	0.30	<0.01
Herd	Peshawar	429.37±106	3.92±0.27	0.53±0.10	4.37±0.10	3.04±0.07	2.8±0.19	4057±152
	Harichand	169.34±17	3.07±0.13	0.19±0.03	4.62±0.05	3.07±0.03	2.47±0.11	2889±116
	Surezi	123.86±31	2.56±0.33	0.25±0.07	4.65±0.06	3.16±0.05	2.45±0.28	1879±174
	Khazana	186.45±36	3.17±0.31	0.41±0.25	5.16±0.06	3.37±0.06	2.5±0.25	6196±378
	<i>P</i> value	<0.01	<0.01	<0.03	<0.01	<0.01	0.5	<0.01
Parity	1	195.68	3.12	0.23	4.79	3.2	2.34	3649.02
	2	270.68	3.21	0.24	4.63	3.1	2.57	3412.21
	3 and above	204.78	3.32	0.42	4.5	3.02	2.75	3112.32
	<i>P</i> value	0.55	0.33	0.22	0.9	0.64	0.08	0.9

*Significant (P value<0.05); AMY= Annual Milk Yield, SCC= Somatic Cells Count, SCS= Somatic Cells Score, FOM= Frequency of Mastitis, FP= Fats Percentage, LP= Lactose Percentage, PP= Proteins Percentage, FP= Fat Percentage

Discussion

Recent researchers have disclosed the important role of *CD4* in immunity, but its association with production traits and mastitis related traits requires more explanation. SNP (C868T) in *CD4* gene of human was discovered to be having potential link with enhanced vulnerability to HIV-1 disease in African populations (Oyugi *et al.* 2009; Hennig *et al.* 2011). Significant link of SNP in *CD4* gene with rheumatoid arthritis was confirmed in Egyptian female patients (Hussein *et al.* 2012). *CD4* gene has potential function in the defence reaction against pathogen that induces

mastitis in cattle (Gao *et al.* 2012). The experimental work revealed that *CD4* gene is directly linked to the immune resistance offered to mastitis (Wang *et al.* 2013). Keeping in view the importance of *CD4* gene in various inflammatory conditions, the present study was planned to explore SNPs in the exons 1, exon 2 and exon 4 of *CD4* gene, and to examine their association with milk production traits and mastitis indicator traits.

He *et al.* (2011) reported a significant association of SNPs in *CD4* gene with production traits and mastitis indicator traits (SCS) in Chinese Holstein cattle. Usman

et al. (2016) recorded a significant association of SNPs in *CD4* gene fat percentage but a non-significant association with SCC and SCS. In their study, homozygous wild type genotype cows showed 0.91% higher fat percentage than the heterozygous genotype. In addition, the study discovered that the wild type genotypes were linked with higher protein and lactose percentage compared to the other genotypes, although these differences were not statistically significant ($P>0.05$). In another study, Usman *et al.* (2017) found that polymorphism in bovine *CD4* and *LAG-3* genes of Holstein cattle have significant association with the SCC. In the present study, all the SNPs in *CD4* gene were significantly associated with the frequency of clinical mastitis (FOM) but none was associated with the SCC and SCS. The possible reasons for discrepancies may be differences in the breeds of cattle, the environment conditions, and population structure.

Schutz *et al.* 1990 reported that fat and protein yield increase with parity i.e. from parity one to parity two; whereas, there was a decrease in the fat percentage in 3rd and later parities. Furthermore, the authors reported that SCC increase with parity. Similarly, different breed of cattle were found to have different fat, lactose and protein percentage. Nash *et al* (2000) recorded that SCS increases with the increase in parity number i.e. from 2.73 to 2.85 in the first and second-lactation cows, respectively. Poso and Maantyasari (1996) in their study on Finnish Ayrshrie cattle reported that decrease in SCS occur for first, second and third lactation (4.42, 4.28 and 4.27, respectively). Results of the current study are in line with the findings of Schutz *et al.* (1990) and Nash *et al* (2000). The differences between the two studies might be due to the variation in the breeds, management and environmental conditions at the respective farms.

Strandberg and Shook (1989) recommended that if importance of selection for milk yield is privileged over selection for mastitis, it might be estimated to increase the frequency of CM as a result of the genetic correlation between more milk yield and clinical mastitis. Present study is in complete agreement with the statement of Strandberg and Shook (1989), as we found higher rate of mastitis in high milk yielding breed i.e. the Holstein-Friesian compared to Jersey and the indigenous breed (Achai).

In conclusion, present study suggests that the *CD4* gene should be considered as potential candidate gene for studies of mastitis resistance in dairy cattle. The four SNPs in *CD4* gene that were found to have significant association with the mastitis indicator traits could be strong genetic markers against mastitis resistance. In addition, suitable dairy breed in early parity should be selected providing adequate management condition for maximum production and to lower the incidence of mastitis occurrence.

Conflict of Interest: The authors declare that they have no conflict of interest.

Acknowledgements

This research was financially supported by the projects (SRGP-536, HEC Pakistan), Modern Agro-industry Technology Research System (CARS-36), and

the Program for Changjiang Scholar and Innovation Research Team in University (IRT-15R62). Beijing Sanyuan Breeding Technology Ltd Co. funded project (SYZY-20190005).

References

- Fabbri M, Smart C and Pardi R 2003. T lymphocytes. *Int J Biochem Cell Biol*, 35: 1004-1008.
- Gao J, Ferreri M, Yu F, Liu X, Chen L, Su J and Han B 2012. Molecular types and antibiotic resistance of *Staphylococcus aureus* isolates from bovine mastitis in a single herd in China. *Vet J*, 192(3): 550-552.
- He Y, Chu Q, Ma P, Wang Y, Zhang Q, Sun D, Zhang Y, Yu Y and Zhang Y 2011. Association of bovine *CD4* and *STAT5b* single nucleotide polymorphisms with somatic cell scores and milk production traits in Chinese Holsteins. *J dairy Res*, 78(2): 242-249.
- Hennig BJ, Velez-Edwards DR, Schim van der LMF and Bisseye C 2011. *CD4* intragenic SNPs associate with HIV-2 plasma viral load and *CD4* count in a community-based study from Guinea-Bissau, West Africa. *J Acquir Immune Defic Syndr*, 56: 1-8.
- Hussein YM, El-Tarhouny SA and Mohamed RH 2012. Association of *CD4* enhancer gene polymorphisms with rheumatoid arthritis in Egyptian female patients. *Rheum Int*, 32(8): 23-25.
- Leyva-Baca I, Schenkel F, Martin J and Karrow NA 2008. Polymorphisms in the 5'upstream region of the CXCR1 chemokine receptor genes, and their association with somatic cell score in Holstein cattle in Canada. *J Dairy Sci*, 91: 407-417.
- Nash DL, Rogers GW, Cooper JB, Hargrove GL, Keown JF and Hansen LB 2000. Heritability of clinical mastitis incidence and relationships with sire transmitting abilities for somatic cell score, udder type traits, productive life, and protein yield. *J Dairy Sci*, 83: 2350-2360.
- Oyugi JO, Vouriot FC, Alimonti J and Wayne S 2009. A common *CD4* gene variant is associated with an increased risk of HIV-1 infection in Kenyan female commercial sex workers. *J Infect Dis*, 199: 1327-1334.
- Perez-Diez A, Joncker NT, Choi K, Chan WF, Anderson CC, Lantz O and Matzinger P 2007. *CD4* cells can be more efficient at tumor rejection than *CD8* cells. *Blood J*, 109: 5346-354.
- Posô I and Mantysaari EA 1996. Relationships between clinical mastitis, somatic cell score, and production for the first three lactations of Finnish Ayrshire. *J Dairy Sci*, 79: 1284-1291.
- Rainard P and Riollot C 2006. Innate immunity of the bovine mammary gland. *Vet Res*, 37(3): 369-400.
- Schutz MM, Hansen LB and Steuernagel GR 1990. Variation of Milk, Fat, Protein, and Somatic cells for Dairy Cattle. *J Dairy Sci*, 73: 484-493.
- Shelley S and Hill AVS 2003. Genetic susceptibility to infectious disease. *Trends Microbiol*, 11: 445-448.
- Sordillo LM and Shafer-Weaver K 1997. Immunobiology of the mammary gland. *J Dairy Sci*, 80(8): 1851-1865.
- Strandberg E and Shook GE 1989. Genetic and economic responses to breeding programs that consider mastitis. *J Dairy Sci*, 72: 21-36.

- Summerfield A, Rziha HJ and Saalmüller A 1996. Functional characterization of porcine *CD4*⁺ *CD8*⁺ extrathymic T lymphocytes. *Cell Immunol*, 168: 291-296.
- Usman T, Wang Y, Song M, Wang X, Dong Y, Liu C, Wang S, Zhang Y, Xiao W and Ying Yu 2017. Novel polymorphisms in bovine *CD4* and *LAG-3* genes associated with somatic cell counts of clinical mastitis cows. *Genet Mol Res*, (4). doi: 10.4238/gmr16039859.
- Usman T, Yu Y, Zhai L, Liu C, Wang X and Wang Y 2016. Association of *CD4* SNPs with fat percentage of Holstein cattle. *Genet Mol Res*, 15(3). doi: 10.4238/gmr.15038697.
- Zhao X and Lacasse P 2008. Mammary tissue damage during bovine mastitis: causes and control. *J Anim Sci*, 86(13): 57-65.