



Yak TNF α gene: Sequence analysis reveals high amino acid identity with cattle but variations in potential binding sites for transcription factors

SHUBHAM GOYAL¹, P KATHIRAVAN², R S KATARIA³, S K NIRANJAN⁴ and S JAYAKUMAR⁵

National Bureau of Animal Genetic Resources, Karnal, Haryana 132001 India

Received: 24 March 2012; Accepted: 6 July 2012

ABSTRACT

The present study reports the sequence characterization of TNF α gene in Indian yak (*Bos grunniens*). A total of 2786 nucleotides long sequence data were generated by primer walking method covering complete TNF α gene and 5' upstream basal promoter region. Predicted coding DNA sequence (CDS) of yak TNF α gene revealed only 4 nucleotide changes compared to that of cattle however, without altering amino acid sequence. Similarly, predicted CDS of yak had a very high sequence similarity with other ruminants like buffalo (99%), goat (96.2%) and sheep (95.9%). Phylogenetic analysis of conceptualized amino acid sequences of TNF α gene of various species revealed the clustering of yak sequence with other ruminants under a major group distinct from the non-ruminants species. A total of 14 putative transcription factor binding sites were predicted within 818 nucleotides long 5' upstream region of yak TNF α gene. Interestingly, 2 potential binding sites for transcription factors Elk-1 and v-Myb were present in yak and indicine cattle but absent in taurine cattle due to an A-G transition within promoter at 526 nucleotides upstream to start codon. Variations within 5' upstream sequence need to be explored further for their functional and regulatory roles in altering expression level of TNF α gene in these species.

Key words: *Bos grunniens*, Phylogeny, TNF α gene, Transcription factor binding sites

Tumor necrosis factor alpha (TNF α), one of the most important inflammatory cytokines, plays a significant role in tissue destruction process through induction of apoptosis in tumour cells and regulates cell proliferation and differentiation. TNF α gene is well studied in many mammalian species and is one of the important candidate genes in bovine breeding research on genetic resistance/susceptibility to diseases. TNF α gene was mapped to chromosome 23 in cattle and polymorphism within coding region, and 5'UTR of bovine TNF α gene was found associated with somatic cell score in Chinese Holstein cows (Xu *et al.* 2010). Also, SNPs within the promoter and 5' upstream region of TNF α gene were reported to be associated with progression of bovine leukemia virus infection (Konnai *et al.* 2006) and poly cystic ovary syndrome (Yun *et al.* 2011) in cattle. High nucleotide and amino acid sequence similarities were reported between TNF α genes of river buffalo, swamp buffalo and cattle (Gupta *et al.* 2004, Mingla

et al. 2007). However, no information is available on yak (*Bos grunniens*). Mastitis and many commonly occurring bacterial, viral and parasitic diseases were reported earlier, but yak is said to be relatively more resistant to some of them including mastitis as compared to cattle (Joshi *et al.* 1997, Weggi, 1983, Wiener *et al.* 2003). In this paper, we report the nucleotide sequence of yak TNF α gene, its upstream basal promoter region and phylogenetic relationship with other ruminant and farm animal species. We also predicted the putative transcription factor binding sites within TNF α promoter to understand the variations within the regulatory regions of yak, taurine (*Bos taurus*) and indicine (*Bos indicus*) cattle.

MATERIALS AND METHODS

DNA sample from a yak in Nagaland state of North-East hilly region of India was utilized for the present study. Blood was collected from jugular vein into vacutainer tubes containing EDTA and DNA was extracted from whole blood following standard phenol-chloroform extraction method (Sambrook and Russell 2001). Eight sets of oligonucleotide primers were designed based on cattle genome sequence data available at ENSEMBL (Chromosome: BTAU4.0:23) using PRIMERSELECT program of LASERGENE software. Primers were designed to amplify the complete TNF α gene

Present address:^{1,4}(shubhamgoyal28@gmail.com, sakethniranjan@gmail.com, jeyvet@gmail.com). ²Animal Production and Health Laboratory, Joint FAO-IAEA Division of Nuclear Techniques in Food and Agriculture, International Atomic Energy Agency, Vienna (kathirvet@yahoo.co.in). ³Technical Officer (Livestock Genetics and Breeding), Animal Production and Health Section, IAEA, Vienna (kathirvet@yahoo.co.in).

Table 1. Details of oligonucleotide primer sequence, annealing temperature and product size for TNF α gene

Region	Primer sequence	Annealing temperature	Product size
TNF α -I	For: 5' GCCCTGCTGTGCTGGAGTTCGT 3' Rev: 5' GGGGCGGGGTTTGGAAAAGT 3'	60.7°C	470
TNF α -II	For: 5' CGAGCGGGCGGAGTGTAGGA 3' Rev: 5' CCGGATCATGCTTTTGGTGCTCA 3'	61.0°C	452
TNF α -III	For: 5' CGCCCTCTTCTCACCCACAG 3' Rev: 5' CTCCCCATCTTTCCCCATCAACAC 3'	61.0°C	542
TNF α -IV	For: 5' CCGGCCTTGGCTCATTCTC 3' Rev: 5' GGGCCACCTGGGGACTGC 3'	56.8°C	581
TNF α -VA	For: 5' TTGGACGAATGAGTAAGCAGAGAT 3' Rev: 5' AAAAGAGGAAAGGAAAAGGTAAGC 3'	54.7°C	575
TNF α -VB	For: 5' GTGGGTGTGGGTGGGGTTTAGTC 3' Rev: 5' TGAGGGCATTGGCATAACGAGTCC 3'	61.2°C	560
TNF α -VI	For: 5' GGGGCGGGCAGAACACGAG 3' Rev: 5' GGCGATGATCCCAAAGTAGACCTG 3'	61.3°C	507
TNF α -VII	For: 5' CATCAGCCGCATTGCAGTCTCCTA 3' Rev: 5' GGTGGTTGGTGGCCCGTTGTC 3'	60.1°C	434

including intronic regions and 5' upstream sequence along with TNF α promoter. Sequence of oligonucleotide primers along with their annealing temperature and expected product size of amplicons are listed in Table 1. Polymerase chain reaction was performed in a final reaction volume of 25 μ L, containing ~100ng of genomic DNA, 5 pmole of each primer, 200 μ M of each dNTP, 2.5 μ L of 10X buffer with 1.5 mM MgCl₂ and one unit of Taq DNA polymerase. Amplification was performed using MASTERCYCLER EP with an initial denaturation at 95°C for 2.5 min followed by 35 cycles of 94°C for 30s, respective annealing temperature of primer for 30s and 72°C for 1 min, with a final extension for 5 min at 72°C. The PCR products were visualized following electrophoresis through a 2% ethidium bromide stained agarose gel.

Amplified PCR product from each region was sequenced from both ends with forward and reverse primers using big dye terminator cycle sequencing kit on an automated genetic analyzer. Raw sequence data were edited using Chromas (Ver. 1.45, <http://www.technelysium.com.au/chromas.html>). Homology among the coding DNA sequences of TNF α gene in different species was estimated using MegAlign program of LASERGENE software. Multiple sequence alignments and phylogenetic analysis were performed with CLC Main Workbench version 5.0.2. Various binding sites for putative transcription factors were identified using MATCH (Kel *et al.* 2003). The structure of yak TNF α gene including its exon-intron junctions was predicted using GENSCAN web server (Burge and Karlin 1997) available at <http://genes.mit.edu/GENSCAN.html>.

RESULTS AND DISCUSSION

A total of 2786 bp long sequence data were generated after assembling 8 overlapping sequence fragments of yak

TNF α gene. The assembled sequence was submitted to NCBI GenBank (JQ043175). Comparison of yak TNF α gene with *Bos taurus* (ENSEMBL:BTau 4.0:23) and *Bos indicus* (GenBank accession AF 011927) cattle revealed 23 nucleotide variations. This included 13 transitions, 6 transversions and 4 deletions. Among the deletions, 3 were observed in 5' upstream sequence of *Bos indicus* cattle while the others were found within intron 1 of *Bos taurus* cattle. Yak TNF α gene sequence was found to have more nucleotide differences with indicine cattle (20 nucleotide changes) than taurine cattle (10 nucleotide changes). The predicted coding DNA region of yak TNF α gene was compared with mRNA sequences of *Bos taurus* (NM_173966) and *Bos indicus* (AY221122). There were 4 transitions among bovines with 1 found in exon 2 and 3 in exon 4 (Fig. 1). However, all the 4 transitions within the coding DNA region were synonymous revealing the highly conserved nature of TNF α protein among the 3 bovine species.

Splicing variations in TNF α transcripts were reported with variable length of ORF across different species. In cattle, the ORF was of 705 bp spanning across 4 exonic regions while it was of 702 bp in buffaloes (Gupta *et al.* 2004). This was due to difference in the splice site between intron1 and exon 2 resulting in deletion of 3 nucleotides at 5' end of exon 2. Normally the consensus splice site of an intron will be having GT residues at 5' end while AG residues at 3' end. The splice site at 3' end of intron 2 in TNF α gene was found to beCAGCAG-3' with 2 AG residues resulting in splice variants with or without 3 nucleotides 'CAG' at 5' end of exon 2. This deletion was present in TNF α genes of certain species like dog (NM_001003244; EU249361) and cat (M92061; NM_001009835), and absent in goats (AY304502; D86587) and horses (NM_001081819; EU438779; AB035735). In sheep TNF α gene both types of splice variants

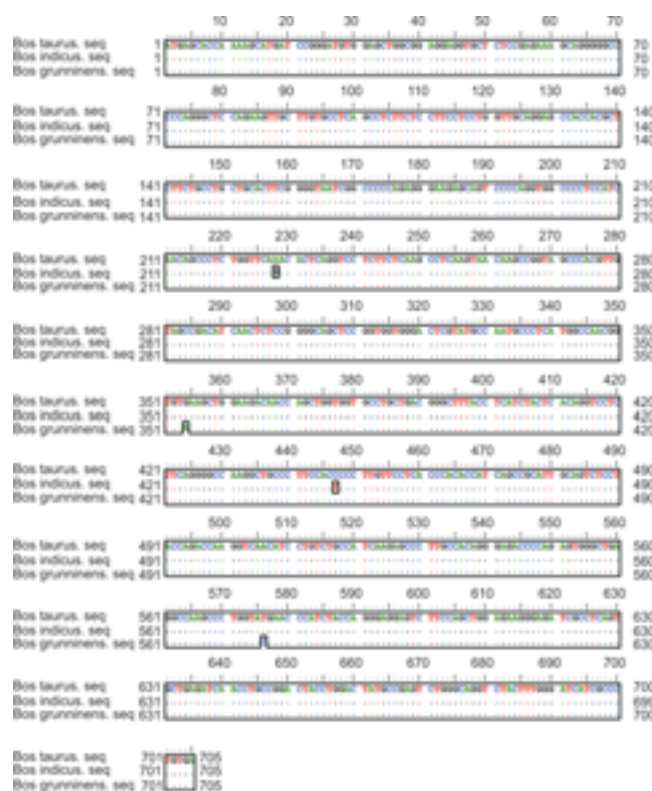


Fig. 1. Multiple alignment of predicted coding DNA sequences of TNF alpha genes of different ruminant species.

were found (705bp-NM_001024860; X56756; X55152 and 702 bp-X55966). The predicted coding DNA sequence of yak-TNF α gene was similar to cattle with an ORF of 705bp. However, this needs to be confirmed by sequencing the mRNA of yak-TNF α gene as the sequence of present study was conceptually derived from genomic DNA.

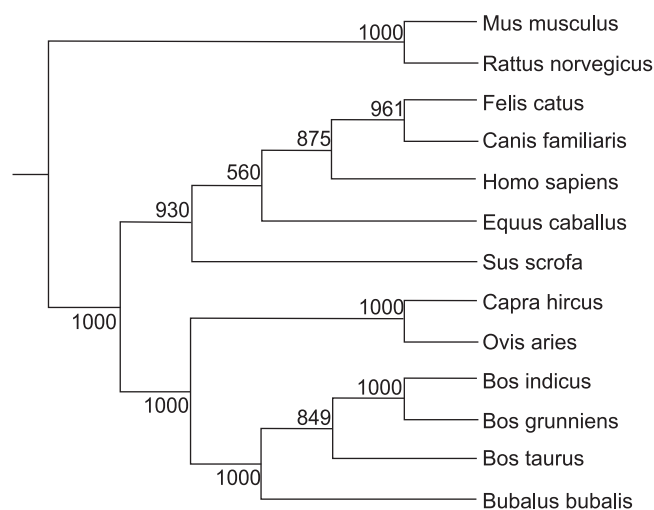


Fig. 2. Phylogenetic analysis of conceptualized amino acid sequences of TNF alpha genes of different species following UPGMA algorithm (values at nodes indicate bootstrap values out of 1000 re-sampled subsets).

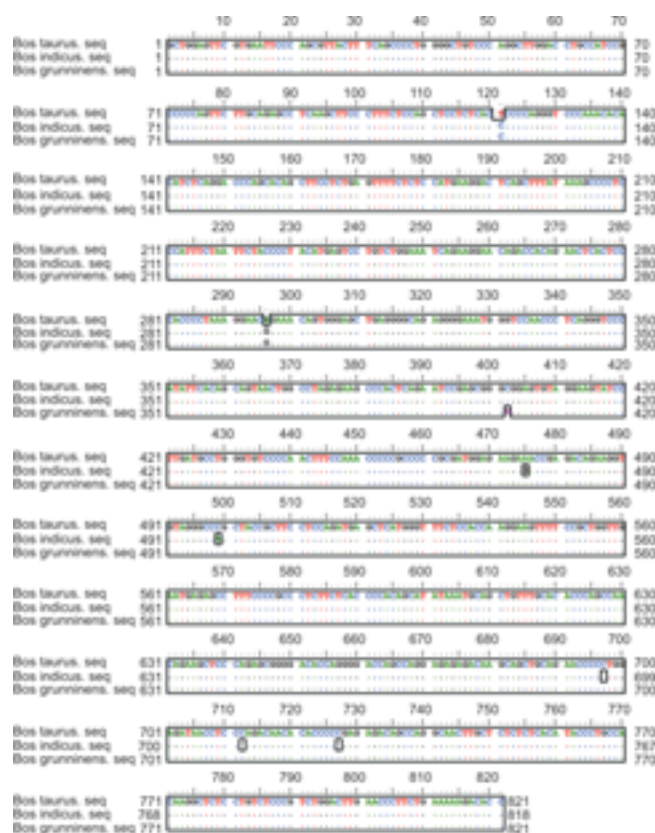


Fig. 3. Multiple alignment of 5' flanking sequences of TNF alpha gene of bovines.

Conceptualized CDS of yak TNF α was further compared with that of other species to establish the sequence homology. The river and swamp type buffalo had a sequence similarity of 99% each with yak while goat and sheep had sequence similarity of 96.2 and 95.9% respectively. With respect to non-ruminants, sequence homology of CDS of yak TNF α ranged from 72.1% (rat) to 84.3% (pig). The trend was similar in sequence homology of yak TNF α amino acid sequence with other species (Table 2). The conceptualized amino acid sequences were subjected to phylogenetic analysis following UPGMA algorithm (Fig. 2) with a bootstrap re-sampling of 1000 overlapping sequence subsets to understand the evolutionary relationship of TNF α among mammalian species. Three major clusters were observed with ruminants, mono-gastric animals along with humans and murines clustering distinctly. Among the ruminants, bovines (*Bos indicus*, *Bos taurus* and *Bos grunniens*) clustered together with internal branches while sheep and goats clustered distinctly. Within the cluster of mono-gastric species, dogs and cats clustered together while human, horse and pig joined the node successively. Interestingly, human TNF α amino acid sequence was closer with dog and cat rather than mouse and rat which together clustered distinctly in a separate node. This is understandable as the rat and mouse were having higher amino acid and coding DNA sequence divergence

Table 2. Sequence homology of predicted coding DNA and amino acid sequences of yak TNF alpha genes with other species

Species	Accession number	CDS	Amino acid
Cattle (taurus)	NM_173966	99.7	100.0
Cattle (indicus)	AY221122	99.4	100.0
River buffalo	JQ043173	99.0	99.6
Swamp buffalo	JQ043174	99.0	99.6
Goat	AY304502	96.2	93.2
Sheep	NM_001024860	95.9	92.7
Pig	NM_214022	84.3	81.9
Dog	NM_001003244	80.9	78.5
Horse	NM_001081819	81.6	78.6
Cat	M92061	81.9	80.7
Human	NM_000594	81.8	79.0
Mouse	NM_013693	73.0	71.8
Rat	NM_012675	72.1	70.9

(about 30%) than the mono gastric species (about 20%). These findings are in agreement with our earlier report on S100A8 gene in buffalo and yak (Kathiravan *et al.* 2011).

The promoter region of TNF α gene were found significant phylogenetically in serving as markers of ancestry and disease association in humans (Baena *et al.* 2002). Several extended haplotypes that contain SNPs at positions-856, -862 and -307 of human TNF α gene were found to be associated with differential susceptibility to a variety of autoimmune diseases (Alper *et al.* 1992, Wilson *et al.* 1993, Ugliarolo *et al.* 1998). The 818 nucleotides long upstream sequence of yak TNF α gene generated in the study was subjected to MATCH analysis to predict putative binding sites for regulatory transcription factors. A total of 14 putative transcription factor binding sites were predicted within this region of which 7 were found with sense strand and remaining with anti-sense strand (Table 3). All the putative transcription factor binding sites were having 100% match of core

sequence and more than 90% matrix match. There were 3 binding sites for Hand1/E47 factor at positions -23, -100 and -571 bases upstream to start codon. Two sites each for Elk-1 and GCR1 were found at positions -264, -519 and -309, -655 respectively. Other predicted binding sites included RFX1, Elf1, Zeste, v-Myb, NF-E2, dl and Ik-1. The set of binding sites predicted that in 5' upstream sequence of yak TNF α gene was different from that of reported earlier in humans (Takashiba *et al.* 1993). The most notable difference was the absence of potential binding sites for NF κ B, a commonly reported LPS inducible transcription factor in humans (Faisst and Meyer 1992, Takashiba *et al.* 1993). Another interesting feature within 5' upstream sequence of TNF α gene was the variation among the 3 bovine species; 2 nucleotide differences between yak and taurine cattle and 5 nucleotide variations between yak and indicine cattle. Among these, a A-G transition found at 526 bases upstream (Fig. 3) to the start codon in yak and indicine cattle resulted in 2 additional potential binding sites, viz. Elk-1 and v-Myb at -519 to -532 and -523 and -532 positions respectively. Both Elk-1 and v-Myb were absent in taurine cattle due to alteration in the core sequence of binding sites of these factors. Elk-1 (E 26 (ETS)-like transcription factor 1) functions as a transcriptional activator and is characterized by a common protein domain that regulates DNA binding to target sequences. v-Myb is transcription factor for proto-oncogenic component of avian myeloblastoma family and is involved in transcriptional repression. In animals v-Myb has the cellular counterpart known as c-Myb (Lüscher and Eisenman 1990). Animal Myb genes were reported to be involved in the regulation of cell proliferation and differentiation (Thompson and Ramsay 1995). Leung *et al.* (2000) reported that the phylogenetic footprint of TNF α promoter revealed sequences containing functional activator binding sites that conserved in the primate lineage while

Table 3. Predicted transcription factor binding sites within 5' flanking sequence of cattle and yak TNF α gene

Factor	<i>Bos taurus</i>		<i>Bos indicus</i>		<i>Bos grunniens</i>		Strand	Core match	Matrix match	Sequence
	From	To	From	To	From	To				
Ik-1	-810	-798	-807	-795	-810	-798	-	1	0.951	tgaattccccagcg
GCR1	-663	-655	-660	-652	-663	-655	+	1	0.968	agcttccctc
dl	-653	-643	-650	-640	-653	-643	-	1	0.962	gagttttctct
NF-E2	-636	-626	-633	-623	-636	-626	-	1	0.900	aggacTCAGCt
Hand1/E47	-586	-571	-583	-568	-586	-571	+	1	0.949	agtcctgTCTGGaaat
Elk-1	ND	ND	-529	-516	-532	-519	+	1	0.940	aggaaCGGAaAcag
v-Myb	ND	ND	-529	-520	-532	-523	+	1	0.944	aggAACGGaa
Zeste	-444	-429	-441	-426	-444	-429	-	1	0.955	aagccCACTCagaatc
Elf-1	-379	-364	-376	-361	-379	-364	-	1	0.914	tttccAAACCcccgc
GCR1	-317	-309	-314	-306	-317	-309	+	1	0.936	cgcttccctc
Elk-1	-277	-264	-274	-261	-277	-264	-	1	0.929	agttTTCCGctgtgt
Hand1/E47	-115	-100	-114	9-9	-115	-100	-	1	0.932	cctcCCAGAcacaca
RFX1	-91	-75	-91	-75	-91	-75	+	1	0.924	agacagccagGCAACtt
Hand1/E47	-38	-23	-38	-23	-38	-23	+	1	0.932	tctcccgTCTGGact

regions that are not essential for gene regulation display a high density of variable sites. However, the present study revealed variations of potential binding sites within TNF α promoter of 3 bovine species. These variations identified *in silico* needs to be confirmed experimentally and explored further for their functional significance with respect to regulation of TNF α gene expression in these bovine species. In conclusion, the present study reported the characterization of complete TNF α gene and its 5' upstream sequence in yak. The predicted amino acid sequence was found to be highly conserved with 100% identity between yak and cattle species although variations existed within binding sites of 5' regulatory region. Our results will also help in identification of candidate gene markers within yak TNF α gene for association with disease resistance traits in future.

ACKNOWLEDGEMENT

The authors are thankful to the Director, National Bureau of Animal Genetic Resources, Karnal, for providing necessary facilities to the study. The study was part of the Institute Research Project funded by ICAR, which is duly acknowledged. Technical assistance provided by Naresh Kumar is also gratefully acknowledged.

REFERENCES

- Alper C A, Awdeh Z, Yunis E J. 1992. Conserved, extended MHC haplotypes. *Experimental Clinical and Immunogenetics* **9**: 58–71.
- Burge C and Karlin S. 1997. Prediction of complete gene structures in human genomic DNA. *Journal of Molecular Biology* **268**: 78–94.
- Faisst S and Meyer S. 1992. Compilation of vertebrate encoded transcription factors. *Nucleic Acids Research* **20**: 3–26.
- Gupta P K, Bind R B, Walunj S S and Saini M. 2004. High nucleotide and amino acid sequence similarities in tumor necrosis factor α amongst Indian buffalo (*Bubalus bubalis*), Indian cattle (*Bos indicus*) and other ruminants. *European Journal of Immunogenetics* **31**: 189–93.
- Joshi D D, Lensch J, Sasaki M and Hentsch G. 1997. Epidemiological aspects of yak diseases in Nepal. *Proceedings of the second international congress on yak*. Xining, China, 1–6 September 1997. Xining, China, Qinghai People's Publishing House, pp. 229–233.
- Kathiravan P, Goyal S, Kataria R S, Mishra B P, Jayakumar S and Joshi B K. 2011. Sequence characterization of S100A8 gene reveals structural differences of protein and transcriptional factor binding sites in water buffalo and yak. *Animal Biotechnology* **22**: 124–32.
- Kel A E, Gobling E, Reuter I, Cheremushkin E, Kel-Margoulis O V and Wingender E. 2003. MATCHTM: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Research* **31**: 3576–79.
- Konnai S, Usui T, Ikeda M, Kohara J, Hirata T I, Okada K, Ohashi K and Onuma M. 2006. Tumor necrosis factor-alpha genetic polymorphism may contribute to progression of bovine leukemia virus infection. *Microbes and Infection* **8**: 2163–71.
- Leung J Y, McKenzie F E, Ugialoro A M, Flores-Villanueva P O, Sorkin B C, Yunis E J, Hartl D L, Goldfeld A E. 2000. Identification of phylogenetic footprints in primate tumor necrosis factor-alpha promoters. *Proceedings of National Academy of Sciences USA* **97**: 6614–18.
- Lüscher B and Eisenman R N. 1990. New light on Myc and Myb. 2. Myb. *Genes and Development* **4**: 2235–41.
- Mingla C N, Odbileg R, Konnai S, Ohashi K and Onuma M. 2007. Molecular cloning, sequencing and phylogenetic analysis of inflammatory cytokines of swamp type buffalo contrasting with other bubaline breeds. *Comparative Immunology, Microbiology and Infectious Diseases* **30**: 119–31.
- Sambrook J and Russell D W. 2001. *Molecular Cloning: A Laboratory Manual III*. Cold Spring Harbour, Cold Spring Laboratory Press, NY.
- Takashiba S, Shapira L, Amar S and Van Dyke T E. 1993. Cloning and characterization of human TNF α promoter region. *Gene* **131**: 307–08.
- Thompson M A and Ramsay R G. 1995. Myb: an old oncoprotein with new roles. *Bioessays* **17** (4): 341–50.
- Ugialoro A M, Turbay D, Pesavento P A, Delgado J C, McKenzie F E, Gribben J G, Hartl D, Yunis E J and Goldfeld A E. 1998. Identification of three new single nucleotide polymorphisms in the human tumor necrosis factor-alpha gene promoter. *Tissue Antigens* **52**: 359–67.
- Weggi S. 1983. Observation on streptococcal mastitis I yaks. *Chinese Journal of Veterinary Medicine* **9**: 12–14.
- Wiener G, Jianlin H and Ruijun L. 2003. *The Yak*. 2nd edn. RAP Publication, Rome.
- Wilson A G, de Vries N, Pociot F, di Giovine F S, van der Putte L B A and Duff G W. 1993. An allelic polymorphism within the human tumor necrosis factor alpha promoter region is strongly associated with HLA A1, B8 and DR3 alleles. *Journal of Experimental Medicine* **177**: 557–60.
- Xu A J, Liu X L, Guo J Z and Xia Z. 2010. Polymorphism of bovine TNF α gene and its association with mastitis in Chinese Holstein cows. *Yi Chuan* **32**: 929–34.
- Yun J H, Choi J W, Lee K J, Shin J S and Baek K H. 2011. The promoter -1031 (T/C) polymorphism in tumor necrosis factor-alpha associated with polycystic ovary syndrome. *Reproductive Biology and Endocrinology* **9**: 131–36.