ONLINE RESOURCES

Genetic variation in the β-lactoglobulin of Chinese yak (*Bos grunniens*)

YANHUA CUI[1], YU CAO[1], YING MA[1]*, XIAOJUN QU[2] and AIJUN DONG[1]

[1]*School of Food Science and Engineering, Harbin Institute of Technology, Harbin 150090, People's Republic of China*
[2]*Institute of Applied Microbiology, Heilongjiang Science Academy, Harbin 150010, People's Republic of China*

[Cui Y., Cao Y., Ma Y., Qu X. and Dong A. 2012 Genetic variation in the β-lactoglobulin of Chinese yak (*Bos grunniens*). *J. Genet.* **91**, e44–e48. Online only: http://www.ias.ac.in/jgenet/OnlineResources/91/e44.pdf]

Introduction

Yak *(Bos grunniens)* is distributed in the area of Central Asian highlands, it thrives in conditions of extreme harshness with severely cold winters, short growing seasons for herbage and no absolutely frost-free periods (Wiener *et al.* 2003). The total population of yak is estimated to be 14 million, about 90% of the domestic yaks are found on the plateau of western China, with the largest population in the Sichuan province (Wiener *et al.* 2003; Zhang *et al.* 2008). In contrast to the milk of dairy cattle and goat, yak's milk contains a higher percentage of protein, organic calcium, and conjugated linoleic acid (Wiener *et al.* 2003; Li *et al.* 2010).

β-lactoglobulin (β-LG) is an important protein that constitutes approximately up to 65% of the total whey protein of bovine milk. β-LG is coded by *LGB* gene, located on chromosome 11 of the bovine genome (Hayes and Petit 1993). It is a single-stranded protein of 18 kDa comprising of 162 amino acid residues, and is stabilized by five disulphide bonds (Apenten *et al.* 2002).

β-LG is a member of the lipocalin protein superfamily with the function of transporter of hydrophobic molecules (Flower *et al.* 2000). The ability of β-LG to bind hydrophobic and amphiphilic molecules has been confirmed, ranging from hexane to palmitic acid to retinol to vitamin D (Perez and Calvo 1995; Narayan and Berliner 1997; Kontopidis *et al.* 2004). Other biological functions of β-LG have been reported, including enzyme regulation, the neonatal acquisition of passive immunity, source of bioactive peptides and antimicrobial activity against mastitis-causing bacteria (Pérez *et al.* 1992; Kontopidis *et al.* 2004; Pellegrini *et al.* 2004; Hernández-Ledesma *et al.* 2008; Chaneton *et al.* 2011).

Milk proteins show genetic polymorphism due to nucleotide sequence substitution or deletion, various degrees of glycosylation and phosphorylation (Farrell *et al.* 2004). Polymorphism of the *LGB* gene was discovered in 1957 (Aschaffenburg and Drewry 1957). To date, 14 β-LG variants in *Bos* genus (*B. taurus, B. javanicus* and *B. grunniens*) have been identified at the protein and DNA levels, including *A, B, C, D, E, F, G, H, I, J, W* and three nomenclature nonunified variants (*X14712*, *EU883598* and *M19088*) (Caroli *et al.* 2009). The *A* and *B* variants occur at high frequency in most breeds of cow. The occurrence of these variants is based on nucleotide exchanges located in exons II (*C*, *D*, *F*, *W*), III (*A*, *H*, *X14712*, *EU883598* and *M19088*), IV (*A*, *G*, *H*, *I*, *X14712*, *EU883598* and *M19088*), V (*F* and *J*), and VI (*E*, *F* and *G*) of *LGB* gene. These variants have been associated with differences in the protein yield, milk composition (fat, protein, casein and total solid content), technological properties of milk, and antimicrobial activity (Lunden *et al.* 1997; Qin *et al.* 1999; Heck *et al.* 2009; Chaneton *et al.* 2011).

At present, the yak β-LG genetic polymorphism has been less extensively investigated at DNA level. Little information is available about yak DNA encoding β-LG. The purpose of this study was to identify and characterize the genetic variation in *LGB* gene of Chinese yak at the DNA sequence level and to establish a phylogenetic relationship of yak β-LG with that of closely related species.

Materials and methods

Collection of milk samples and extraction of DNA samples

Milk samples were collected from 62 Maiwa yak to study the genetic polymorphism at *LGB* locus, in Sichuan province, China, for genomic DNA extraction. After addition of sodium azide (0.4 g/L^{-1}), the milk samples were immediately frozen at −20°C. Genomic DNA was extracted from milk according to the modified procedure by d'Angelo *et al.* (2007) and Cui *et al.* (2011).

*For correspondence. E-mail: maying@hit.edu.cn.

Keywords. yak; β-lactoglobulin; polymorphism; PCR-SSCP; variant.

Table 1. Primers used in this study.

Primer	Sequences (5′–3′)	Location	T_m value (°C)	Product length (bp)	PCR annealing temperature (°C)
Lg-E2F	AGTCCTGTGTTGAGGTTCTG	2671–2689	55.4	598	54
Lg-E2R	TGGATAAGCAGCCTTGGGT	3251–3268	55.2		
Lg-E3F	TTCTCCCTGGCTCCATCTGA	3757–3775	57.4	406	57.5
Lg-E3R	GCAGTGTCTTCATCACCACCA	4143–4162	57.6		
Lg-E4F	TCACTTTCCTCCCGTCTTG	5137–5154	55.2	478	52.2
Lg-E4R	GCTTTGTGCCTGAGTTCTTA	5596–5614	53.4		

Primers and PCR

A 598-bp fragment containing exon II of the yak *LGB* gene was amplified using primers Lg-E2F and Lg-E2R designed according to the sequence of cattle *LGB* gene exon II region at GenBank accession number *X14710*. The fragment comprised positions from 2671 to 3268 of the *LGB* gene. The exons III and IV regions of the yak *LGB* gene were amplified using primers Lg-E3F/Lg-E3R and Lg-E4F/Lg-E4R, respectively (table 1).

The PCR reaction was performed in a 50 μL final volume containing 1 U of KOD Plus DNA polymerase (Toyobo, Shanghai, China), 1× PCR buffer (Toyobo), 1 mM $MgCl_2$, 200 μM each dNTP, 0.4 μM each primer, and approximately 50 ng of yak genomic DNA. Thermal cycling conditions were 94°C for 2 min, followed by 30 cycles of 94°C for 15 s, annealing for 30 s, 68°C for 30 s, and a final extension step of 68°C for 5 min.

PCR-SSCP Analysis

The DCode, universal mutation detection system (Bio-Rad Laboratories, Hercules, USA) was used for single strand conformation polymorphism (SSCP) analysis. The SSCP analysis was carried out as follows, 10 μL of PCR product was added to 10 μL of 2× SSCP Gel Loading Dye (DCode™, Hercules, USA) (0.05% xylene–cyanol, 0.05% bromophenol blue, 20 mM EDTA, 95% formamide). After heat denaturizing at 95°C for 10 min, the samples were cooled rapidly in ice and then loaded on 20 cm × 20 cm, 8% acrylamide, bisacrylamide gels (37.5 : 1). Electrophoresis was performed at 150 V for 15 h at 4°C in 1× TBE buffer. Gels were stained in a solution containing 0.5 μg/mL ethidium bromide for 5–15 min.

Cloning and sequencing

The PCR products showing different patterns on SSCP gels were randomly selected for cloning and sequencing. PCR products were cloned into pTA2 vector using the Target clone™ plus kit following the instructions provided by the manufacturer (Toyobo, Shanghai, China). Positive clones were selected by colony PCR. Plasmid DNA was prepared and sequenced at least twice in both directions, by using an ABI PRISM 377 DNA sequencer (Perkin-Elmer Cetus Instruments, Norwalk, USA). The sequences were submitted to GenBank and assigned accession numbers JN811549, JN811550, JN811551 and JN811552.

Sequence and phylogenetic analyses

LGB-B (*X14710*), *B. taurus* (*X14712*, *EU883598* and *M19088*) were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/). DNA sequences for *LGB-A, C, D, E, F, G, H, I, J* and *W* were retrieved based on the previous studies (Caroli *et al.* 2009).

Only partial sequences corresponding to the coding region of exons II, III and IV were used for phylogenetic analysis because of different lengths of the sequences. Sequences were aligned using the CLUSTAL X software (Thompson *et al.* 1997). The evolutionary distances were calculated using the software package TREECON (van der Peer and de Wachter 1994). The phylogenetic tree was generated by the neighbour-joining (NJ) method using the software package TREECON (van der Peer and de Wachter 1994).

Results and discussion

PCR-SSCP analysis of yak LGB gene

The PCR-SSCP analysis of yak *LGB* gene exon II region presented two distinct migration patterns (figure 1a). The more

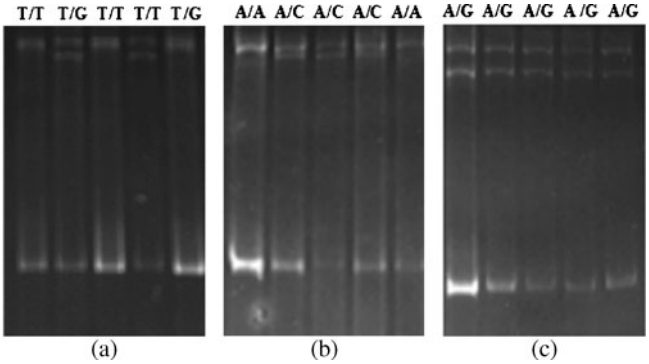

Figure 1. PCR-SSCP analysis of yak *LGB* exons II (a), III (b) and IV (c).

Table 2. Differences in nucleotide and corresponding AA sequences of β-lactoglobulin gene in the *Bos* genus.

Variants	Nucleotide position of β-lactoglobulin gene and corresponding AA position of mature β-lactoglobulin																		Resource
	25 3007	45 3065	50 3080	56 3098	59 3109	63 3982	64 3984	70 4003	78 4025	78 4027	88 5174	100 5210	105 5223	108 5233	118 5263	126 5962	129 5970	143 6017	158 6280	
B *(X14710)*	Gla GC**G**	Glu **G**AG	Pro **C**CT	Ile **A**TC	Gln CA**G**	Asn AA**C**	Gly G**G**T	Lys AA**G**	Ile **A**TC	Ile AT**C**	Asn AA**T**	Lys AAA	Phe **T**TC	Glu G**A**G	Ala G**C**C	Pro C**C**G	Asp **G**AC	Pro CC**C**	Glu G**A**G	*Bos taurus, B. indicus,* *B. grunniens*
A						Asn AA*T*	Asp G**A**T				Asn AA*C*				Val G**T**C					*B. taurus, B. indicus,* *B. grunniens*
C					His CA**T**						NA									*B. taurus*
D		Gln **C**AG									NA									*B. taurus*
E						NA					NA								Gly G**G**G	*B. grunniens*, *B. javanicus*
F			Ser **T**CT			NA											Tyr\ **T**AC		Gly G**G**G	*B. javanicus*
G						NA				Met AT**G**	NA								Gly G**G**G	*B. javanicus*
H						NA	Asp G**A**T	Asn AAC			NA				Val G**T**C					*B. taurus*
I						NA					NA			Gly G**G**G						*B. taurus*
J						NA					NA					Leu C**T**G				*B. taurus*
W				Leu **C**TC		NA					NA									*B. taurus*
X14712						Asn AA*T*	Asp G**A**T				Asn AA*C*		Val **G**TC		Val G**T**C					*B. taurus*
EU883598						Asn AA*T*	Asp G**A**T				Asn AA*C*				Val G**T**C					*B. taurus*
M19088						Asn AA*T*	Asp G**A**T				Asn AA*C*				Val G**T**C			Pro CC**A**		*B. taurus*
JN811549	Gla GC**T**										Asn AA*C*					NA	NA	NA	NA	*B. grunniens*
JN811550	Gla GC**T**										Asn AA*C*	Lys AA**G**				NA	NA	NA	NA	*B. grunniens*
JN811551									Leu **C**TC		Asn AA*C*					NA	NA	NA	NA	*B. grunniens*
JN811552									Leu **C**TC		Asn AA*C*	Lys AA**G**				NA	NA	NA	NA	*B. grunniens*

The reference sequence is GenBank no. *X14710*, corresponding to the *LGB-B* allele.
Amino acids are provided in the upper lane, and nucleotide in the lower lane of each sequence.
NA indicates the nucleotide is not available from the sequence. Nonsynonymous and synonymous mutations are bold and italics, respectively.
?, information not available (Caroli *et al.* 2009).

common pattern is heterozygous, and the other pattern is homozygous. The PCR products showing different patterns by PCR-SSCP were cloned into pTA2 vector and sequenced. Most samples were heterozygous, and included two kinds of sequences, i.e. two variants. One variant corresponded to the sequence of *X14710*. The second variant did not correspond to any of the known cattle *LGB* variants, and entered into GenBank with accession numbers JN811549 and JN811550 (two sequences are identical in exon II). The sequencing of the second pattern revealed one polymorphic site at position 3007 (G/T) of the referring sequence *X14710*, resulting in two alleles named *G3007* and *T3007*, respectively, on the basis of nucleotide at position 3007. This single mutation is synonymous mutation (codon 25 Gla/Gla).

The PCR-SSCP analysis of yak *LGB* gene exon III region presented two distinct migration patterns (figure 1b). The more common pattern is heterozygous, including two kinds of sequences, i.e. two variants. One variant corresponded to the sequence of *X14710*. The second variant did not correspond to any of the known cattle *LGB* variants, and entered into GenBank with accession numbers JN811551 and JN811552 (two sequences are identical in exon III). The sequencing of variants revealed one polymorphic site at position 4025 (A/C) of the referring sequence *X14710*, resulting in two alleles named *A4015* and *C4025*, respectively, on the basis of nucleotide at position 4025. This single mutation was a nonsynonymous mutation (codon 78 Ile/Leu).

The PCR-SSCP analysis of yak *LGB* gene exon III region presented one migration pattern (figure 1c). All samples were heterozygous, and including two kinds of sequences, i.e. two variants. The variants did not correspond to any of the known cattle *LGB* variants, and entered into GenBank with accession numbers JN811549 and JN811551 that are identical in exon IV, JN8115450 and JN811552 which are identical in exon IV. Comparing with *X14710*, JN811549 and JN811551 showed 1 nucleotide mutation (5174 T/C), and JN8115450 and JN811552 nucleotide mutation (5174 T/C, 5210 A/G). The sequencing of variants revealed one polymorphic site at position 5210 (A/G) of the referring sequence *X14710*, resulting in two alleles named *A5210* and *G5210*, respectively, on the basis of nucleotide at position 5210. This single mutation was synonymous mutation (codon 100, Lys/Lys).

Phylogenetic analysis of LGB genes of Bos genus

The differences of nucleotide and corresponding amino acid in the genus *Bos*, including the new sequences from the present work, are summarized in table 2. Because most changes of variants were in exons II, III and IV, we compared the differences in sequences of these regions. The NJ consensus tree was constructed based on the *LGB* gene sequences in the genus *Bos* in order to reveal and understand the evolution of protein polymorphism (figure 2). The tree shows three main branches, one with the out-grouped goat sequence and two main sequence groups. The nucleotide and AA sequence differences are summarized in table 2.

The first group is subdivided into two subgroups. The subgroup I-1 comprises *LGB-B* (*X14710*), *C*, *E*, *F*, *G*, *J*, and *W*. The nucleotide sequences of *LGB-C*, *E*, *F*, *G*, *J*, *W* each differ from *LGB-B* in a single mutation, and thus probably directly evolved from LGB-B.

The subgroup I-2 comprises *LGB-A*, *H*, *M19088*, *EU0883598* and *X14712*, all being identical in codons 64 (Asp/GAT) and 118 (Val/GTC). Besides the variant H was not deterimined, the other variants were identical in codons 63 (Asn/AAT) and 88 (Asn/AAC). The *EU0883598* was identical to variant A. The nucleotide sequences of *LGB-H*, *M19088*, and *X14712* each differ from LGB-A in a single mutation, and thus probably directly evolved from *LGB-A*.

The second main group comprises JN811549, JN811550, JN811551 and JN811552 from *B. grunniens.* JN811549 and JN811550 both differ from all other sequences by an

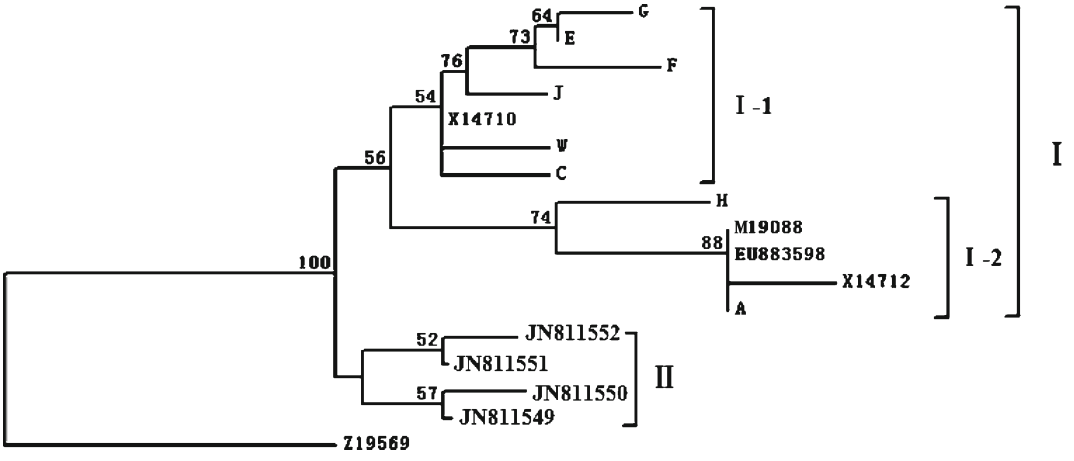

Figure 2. Nucleotide sequence-based neighbor joining consensus tree (random input order, 100-bootstrap datasets of *LGB* sequences in the genus *Bos* (*B. taurus*, *B. javanicus*, and *B. grunniens*) drawn as ladderized rectangular cladogram.

unique nucleotide difference in position 3007. JN811551 and JN811552 have an unique nucleotide difference in position 4025, which result in a nonsynonymous mutation (codon 78 Ile/Leu). The phylogenetic analysis clearly presented β-LG evolution and phylogenetic inter relationship among genus *Bos*.

In conlusion, four genetic variants of yak β-LG were detected by PCR-SSCP analysis and show clearly different migration patterns. To our knowledge, this is the first report of polymorphisms in the yak *LGB* gene at DNA level and can be considered in future researches on the genetic variation within and between yak populations, and on the influence of *LGB* polymorphisms on technological or nutritional parameters of yak milk and association with production traits.

Acknowledgements

This work was supported by National Nature Science Foundation of China (grant nos. 30871953 and 31071571).

References

Apenten R. K. O., Khokhar S. and Galani D. 2002 Stability parameters for β-lactoglobulin thermal dissociation and unfolding in phosphate buffer at pH 7. *Food Hydrocolloids* **16**, 95–103.

Aschaffenburg R. and Drewy J. 1957 Genetics of the β-lactoglobulins of cow's milk. *Nature* **180**, 376–378.

Caroli A. M., Chessa S. and Erhardt G. J. 2009 Invited review, milk protein polymorphisms in cattle, effect on animal breeding and human nutrition. *J. Dairy Sci.* **92**, 5335–5352.

Chaneton L., Pérez Sáez J. M. and Bussmann L. E. 2011 Antimicrobial activity of bovine β-lactoglobulin against mastitis-causing bacteria. *J. Dairy Sci.* **94**, 138–145.

Cui Y. H., Ma Y., Qu X. J., Li H. M. and He S. H. 2011 Establishment of the method of extraction of DNA from yak milk. *J. Harbin Inst. Technol.* **43**, 1892–1895.

d'Angelo F., Santillo A., Sevi A. and Albenzio M. 2007 A simple salting-out method for DNA extraction from milk somatic cells, investigation into the goat *CSN1S1* gene. *J. Dairy Sci.* **90**, 3550–3552.

Farrell Jr H. M., Jimenez-Flores R., Bleck G. T., Brown E. M., Butler J. E., Creamer L. K. *et al.* 2004 Nomenclature of the proteins of cows' milk-sixth revision. *J. Dairy Sci.* **87**, 1641–1674.

Flower D. R., North A. C. T. and Sansom C. E. 2000 The lipocalin protein family-structural and sequence overview. *Biochim. Biophys. Acta* **1482**, 9–24.

Hayes H. and Petit E. 1993 Mapping of the β-lactoglobulin gene and of an immunoglobulin M heavy chain-like sequence in homologous cattle, sheep, and goat chromosomes. *Mamm. Genome* **4**, 207–210.

Heck J. M., Schennink A., van Valenberg H. J., Bovenhuis H., Visker M. H., van Arendonk J. A. *et al.* 2009 Effects of milk protein variants on the protein composition of bovine milk. *J. Dairy Sci.* **92**, 1192–1202.

Hernández-Ledesma B., Recio I. and Amigo L. 2008 β-Lactoglobulin as source of bioactive peptides. *Amino Acids* **35**, 257–265.

Kontopidis G., Holt C. and Sawyer L. 2004 Invited review, β-lactoglobulin, binding properties, structure, and function. *J. Dairy Sci.* **87**, 785–796.

Li H. M., Ma Y., Dong A. J., Wang J. Q., Li Q. M., He S. H. *et al.* 2010 Protein composition of yak milk. *Dairy Sci. Technol.* **90**, 111–117.

Lunden A., Nisson M. and Janson L. 1997 Marked effect of β-lactoglobulin polymorphism on the ratio of casein to total protein in milk. *J. Dairy Sci.* **80**, 2996–3005.

Narayan M. and Berliner L. J. 1997 Fatty acids and retinoids bind independently and simultaneously to β-lactoglobulin. *Biochemistry* **36**, 1906–1911.

Pellegrini A., Hülsmeier A. J., Hunziker P. and Thomas U. 2004 Proteolytic fragments of ovalbumin display antimicrobial activity. *Biochim. Biophys. Acta* **1672**, 76–85.

Pérez M. D. and Calvo M. 1995 Interaction of beta-lactoglobulin with retinol and fatty acids and its role as a possible biological function for this protein, A review. *J. Dairy Sci.* **78**, 978–988.

Pérez M. D., Sanchez L., Aranda P., Ena J. M., Oria R. and Calvo M. 1992 Effect of beta-lactoglobulin on the activity of pregastriclipase. A possible role for this protein in ruminant milk. *Biochim. Biophys. Acta* **1123**, 151–155.

Qin B. Y., Bewley M. C., Creamer L. K., Baker E. N. and Jameson G. B. 1999 Functional implications of structural differences between variants A and B of bovine beta-lactoglobulin. *Prot. Sci.* **8**, 75–83.

Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F. and Higgins D. G. 1997 The clustalx windows interface, flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876–4882.

van der Peer Y. and de Wachter R. 1994 TREECON for windows, a software package for construction and drawing of evolutionary trees for the microsoft windows environment. *Comput. Appl. Biosci.* **10**, 569–570.

Wiener G., Jianlin H. and Ruijun L. 2003 The yak. Bangkok, Thailand, FAO regional office for Asia and the Pacific.

Zhang G., Chen W., Xue M., Wang Z., Chang H., Han X. *et al.* 2008 Analysis of genetic diversity and population structure of Chinese yak breeds *Bos grunniens* using microsatellite markers. *J. Genet. Genomics* **35**, 233–238.

Received 7 October 2011, accepted 8 January 2012
Published on the Web: 28 March 2012