Arch. Anim. Breed., 59, 395–400, 2016
www.arch-anim-breed.net/59/395/2016/
doi:10.5194/aab-59-395-2016

Open Access
Archives Animal Breeding

Cloning and characterization of MHC-DQA1 and MHC-DQA2 molecules from yak (*Bos grunniens*)

Fei Ge[1], Sameeullah Memon[1], Dongmei Xi[1], Shijun Li[1], Xiangying Liu[1], Guozhi Li[1], Liping Wang[1], Jing Leng[1], Sehroon Khan[2,3], and Weidong Deng[1]

[1]Yunnan Provincial Key Laboratory of Animal Nutrition and Feed, Faculty of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201, China
[2]Center for Mountain Ecosystem Studies, Kunming Institute of Botany, Chinese Academy of Science, Kunming 650201, China
[3]The World Agroforestry Center, East and Central Asia, Kunming 650201, China

Correspondence to: Weidong Deng (1692425306@qq.com)

Received: 23 May 2016 – Revised: 26 August 2016 – Accepted: 5 September 2016 – Published: 14 September 2016

Abstract. The major histocompatibility complex (MHC) plays a crucial role in the processing and presentation of antigens and in discrimination between self and non-self. The aim of this investigation was to scrutinize the structural diversity and possible duplication of the *MHC-DQA* genes in yak (*Bos grunniens*). Two cDNA sequences were amplified and designated as *Bogr-DQA1* (DQA*0101) and Bogr-DQA2 (DQA*2001) with GenBank accession numbers JQ864314 and JQ864315, respectively. The nucleotide and amino acid sequence alignment between Bogr-DQA1 and Bogr-DQA2 molecules showed that these two identified *MHC-DQA* gene sequences had more similarity to alleles of specific *DQA1* and *DQA2* genes from other Ruminantia species than to each other. The result from phylogenic investigation also revealed that there was a larger genetic distance between these two genes than between homologous genes from different species. The presence of different bovine DQA putative motifs and the large genetic distance between Bogr-DQA1 and Bogr-DQA2 suggest that these sequences are non-allelic. Further, these results indicate that *DQA* gene duplication occurs in ruminants. This study will be helpful in knowing MHC diversity in common ruminants and will deepen our understanding of the variation of immunological functions, evolutionary constraints, and selective forces that affect MHC variation within and between species.

1 Introduction

The extent of genetic diversity is known to be linked to the ability for adapting to environmental changes and with the capacity to evolve (Reed and Frankham, 2003). The potential of an organism for evolutionary interactions with either pathogens or other species as well as its Darwinian fitness is related to immunological functions (Lazzaro and Little, 2009). Diversity of genes important for immune functions may be associated with resistance or susceptibility to pathogens (Trowsdale and Parham, 2004; Tibayrenc, 2007). A cluster of associated genes called the major histocompatibility complex (MHC) play a key role in presenting antigenic peptides to T lymphocytes (Klein, 1986). In the vertebrate genome, genes of MHC are known to be the most polymorphic genes which are maintained by balancing selection, predating speciation events, and reflecting the co-evolution of hosts with their pathogens (Bernatchez and Landry, 2003).

The class II genes of MHC encode for the α and β chains of DR and DQ dimer molecules, which present antigenic peptides to helper T cells (McKinney et al., 2013). In rats, mice, rabbits, and pigs, there is a single gene of the *DQ* genes, whereas in dogs and humans multiple *DQ* genes copies have been identified but only one of them appears to be expressed (Kappes and Strominger, 1988; Trowsdale, 2001). The number of *DQ* loci in ruminants has been reported to be different in different species; for example, the haplotypes in cattle and buffalo contain two copies of

DQ genes (Andersson and Rask, 1988; Sigurdardottir et al., 1992; Sena et al., 2011) and both of these *DQ* genes are expressed (Russell et al., 1997). Hence, the polymorphism as well as the duplication of *DQ* gene increases the differences at the cell surface by inter- and intra-haplotype pairing of the α and β chains during dimerization. The formation of functional restriction elements is the result of inter-haplotype combination of *DQA* and *DQB* molecules with duplicated *DQA* haplotypes (Glass et al., 2000).

The yak (*Bos grunniens*) is a native ruminant of the Qinghai–Tibet Plateau and is well adapted to that hypoxic and low-temperature environment (An et al., 2005; Xi et al., 2014). The yak plays an important role in the Himalaya region, contributing in particular in terms of high-quality foods, fuel, conveyance, and cash income for the local herdsmen, and is believed to have more resistance than cattle to some endemic diseases under the harsh environmental conditions (Wiener et al., 2003; Qiu et al., 2012).

In the present investigation, the *DQA* genes of the yak MHC were cloned and characterized to identify structural changes and possible duplication which remains in other ruminants. This study will be helpful in characterizing MHC diversity in common ruminants and will deepen our understanding of the variation of immunological functions, evolutionary constraints, and selective forces that affect MHC variation within and between species.

2 Material and methods

Liver samples from six mature (8.2 ± 2.8 years) Zhongdian yak (*B. grunniens*) were randomly collected from the slaughterhouse in Shangri-La City, Diqing District, Yunnan Province, China. The samples were snap-frozen, kept in liquid nitrogen during transportation, and then stored at $-80\,^{\circ}$C. The total RNA was extracted using a total RNA extraction kit (Beijing Tiangen Biotech Co., Ltd., Beijing, China). Before the first-strand cDNA synthesis, DNase I treatment of the total RNA was carried out before the cDNA was constructed using RevertAid™ first-strand cDNA synthesis kits (Fermentas Inc., Ontario, Canada) following the manufacturer's protocol.

The three primers, which have been used successfully to clone the buffalo *DQA* genes (Niranjan et al., 2009), were used to amplify the complete cDNA of the yak *DQA1* and *DQA2* sequences. The forward primer (A1A2F: 5′-ACCTTGAGAAGAGGATGGTCCTG-3′) was matched onto the consensus region. For amplifying unique sequence independently, two different reverse primers (A1R: 5′-ATTGCACCTTCCTTCTGGAGTGT-3′; A2R: 5′-TCATAGATCGGCAGAACCACCTT-3′) were matched onto the bovine regions specific to either *DQA1* or *DQA2*. Thus, the primers A1A2F and A1R as well as A1A2F and A2R were combined to amplify the *Bogr-DQA1* and *Bogr-DQA2* genes, respectively.

Polymerase chain reaction (PCR) was performed using a Bioer Life Express thermal cycler using 2.0 µL of DNA template (about 50 ng), 12.5 µL of PCR Power Mix, 1.0 µL of 10 pmoL µL^{-1} of each primer and 8.5 µL of double-distilled water in a total reaction volume of 25 µL. The PCR conditions started with an initial denaturation of 94 °C for 3 min, followed by 35 cycles of the following steps: 94 °C for 1 min, 59 °C for 45 s, and 72 °C for 45 s, with a final extension of 10 min at 72 °C. Amplified products were visualized on agarose gel stained with ethidium bromide. The PCR products were sequenced bidirectionally using an ABI 3730 DNA analyzer (Applied Biosystems Inc.) at the Sun Biotechnology Company (Beijing, China).

The nucleotide sequence analysis was conducted using the BLAST software from the NCBI server (http://www.ncbi.nlm.nih.gov/BlAST). The sequence predictions were carried out using the ORF Finder software (http://www.ncbi.nlm.nih.gov/projects/gorf/). The theoretical isoelectric point (pI) and molecular weight (Mw) of the two putative proteins of the yak genes were also computed using the compute pI / Mw tool (http://www.expasy.org/tools/pi_tool.html). The complete cDNA and deduced amino acid sequences were compared with the orthologous sequences by using the online Clustal W (http://www.genome.jp/tools/clustalw/). A phylogenetic tree for coding region of different orthologous *DQA* alleles from different species was constructed by means of the neighbor-joining method using MEGA software version 4 (Tamura et al., 2007).

3 Results and discussion

In the present study, two fragments of 799 bp (*Bogr-DQA1*) and 810 bp (*Bogr-DQA2*) were amplified from the template cDNA. The cDNA sequence analysis showed that the two genes are not homologous to any of the known yak genes. Both the sequences were deposited into the NCBI GenBank database with accession numbers JQ864314 (for *DQA1*) and JQ864315 (for *DQA2*) as well as the Immuno Polymorphism Database (www.ebi.ac.uk/ipd/mhc/bola/nomenclature) with assigned official names Bogr-DQA*0101 and DQA*2001, respectively, based on identity to existing *BoLA-DQA* sequences. Furthermore, the 799 and 810 bp cDNA sequences represent two single genes containing a complete open reading frame (ORF) of 768 nucleotides and both encoding a polypeptide of 255 amino acids. The pI of yak DQA1 and DQA2 was 5.52 and 4.84, respectively. The Mw of the two putative proteins was 28 293.43 and 27 874.79.

The *Bogr-DQA* genes were compared for homology with bovine leukocyte antigen (BoLA) *DQA* genes (*BoLA-DQA*). The *Bogr-DQA1* and *Bogr-DQA2* showed the highest identities (97–99 and 99–100 %) at the nucleotide sequence level, with that of *BoLA-DQA1* and *BoLA-DQA2*, respectively (Table 1). However, the sequence identity between the *Bogr-DQA1* and *Bogr-DQA2* was found to be only 85 %, similar

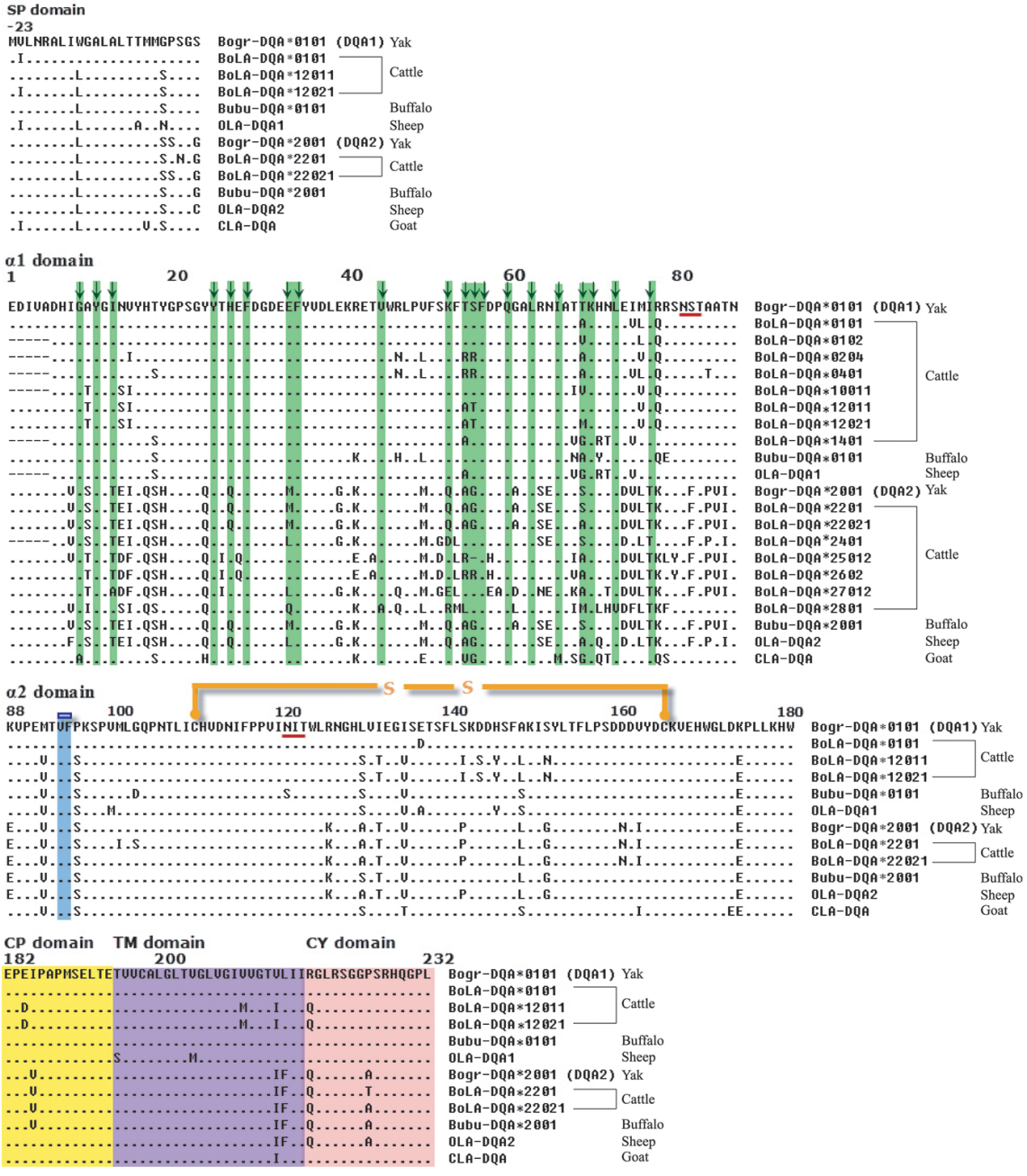

Figure 1. An alignment comparison between the amino acid sequences of Bogr-DQA and orthologous DQA sequences. (The arrows indicate the amino acid positions constituting part of the peptide binding sites. The putative N-linked glycosylation sites are underlined (–). The blue rectangle indicates the position of residues associated with binding of CD^{4+} molecules. A point (·) indicates amino acid identity and a hyphen (-) indicates a gap inserted to maximize alignment.) The reference GenBank accession numbers for DQA1 alignment are Y07898 (BoLA-DQA*0101), U80884 (BoLA-DQA*0102), U80872 (BoLA-DQA*0204), U80871 (BoLA-DQA*0401), AB257109 (BoLA-DQA*10011), Y07819 (BoLA-DQA*12011), D50454 (BoLA-DQA*12021), U80869 (BoLA-DQA*1401), DQ440647 (Bubu-DQA*0101), and M93430 (OLA-DQA1). The reference GenBank accession numbers for DQA2 alignment are Y07820 (BoLA-DQA*2201), D50045 (BoLA-DQA*22021), U80868 (BoLA-DQA*2401), Y14020 (BoLA-DQA*25012), Y14021 (BoLA-DQA*2602), Y14022 (BoLA-DQA*27012), AF037314 (BoLA-DQA*2801), DQ440648 (Bubu-DQA*2001), M93433 (OLA-DQA2), and AY464652 (CLA-DQA).

Table 1. Nucleotide similarity comparisons from the $\alpha 1$, $\alpha 2$, and CP/TM/CY regions between *Bogr-DQA1/DQA2* and *BoLA-DQA1/DQA2* genes (the amino acid similarity is shown in parentheses).

	Bogr-DQA*2001 (DQA2)	BoLA-DQA*0101 (DQA1)	BoLA-DQA*(2201)
Bogr-DQA*0101 (DQA1)			
$\alpha 1$	79.0	97.0	79.0
$\alpha 2$	89.0	99.0	88.0
CP/TM/CY	90.0	99.0	90.0
Entire gene (protein)	85.0 (79.0)	98.0 (97.0)	85.0 (78.0)
Bogr-DQA*2001 (DQA2)			
$\alpha 1$		80.0	100.0
$\alpha 2$		88.0	99.0
CP/TM/CY		91.0	99.0
Entire gene (protein)		86.0 (79.0)	99.0 (98.0)
BoLA-DQA*0101 (DQA1)			
$\alpha 1$			80.0
$\alpha 2$			87.0
CP/TM/CY			91.0
Entire gene (protein)			85.0 (78.0)

to that of cattle (83.5 %). These results were highly similar to the finding of Niranjan et al. (2009) for water buffalo. However, they demonstrated that *Bubu-DQA* genes have less identity (93.9 and 97.7 %) with that of cattle as compared to the sequence identity between *DQA* genes (85.7 %).

The *Bogr-DQA1* and *Bogr-DQA2* revealed considerable divergence with 115 nucleotide substitutions within the coding regions, resulting in 52 amino acid polymorphisms (Fig. 1). Among these, 54 of the nucleotide mutations occurred within the $\alpha 1$ motif (exon 2), resulting in 30 amino acid substitutions. Furthermore, other amino acid differences were observed, including 4 in the signal peptide (SP), 13 in the $\alpha 2$ domain, 1 in the connecting peptide (CP), 2 in the transmembrane (TM) region, and 2 in the cytoplasmic (CY) domain. Obviously, yak has more amino acid replacements than buffaloes with 45 amino acid polymorphisms (Niranjan et al., 2009).

In addition, the peptide binding sites (PBSs; marked by green arrow sign), the CD^{4+} binding site (marked by a blue rectangle), one intramolecular disulfide bond, one N-glycosylation within the $\alpha 2$ domain and another within the $\alpha 1$ domain were retrieved, indicating the importance of these motifs for maintaining molecular structure and function to protect against invading pathogens (Rudd et al., 1999). The PBS is a specific functional motif which is contacted with the antigens (Brown et al., 1993; Kuduk et al., 2012). In the present study, 20 PBSs were observed (Fig. 1). Among them, only 8 residues were highly conserved at positions 11, 25, 29, 35, 57, 60, 63, and 70 between DQA1 and DQA2 protein sequences from different Ruminantia species. Importantly, the remaining 12 PBS sites have the varied amino acids in both of the polypeptide chains, indicating that it could have association with adaptation of yak to the harsh environment. However, Indian buffaloes have an additional three specific substitutions locating at the positions 36, 94 (hydrophobic > hydrophilic), and 57 (hydrophilic > hydrophobic), leading to the opposite water affinity (Niranjan et al., 2009), which could be due to the breed characteristics because buffalo can adapt well to the tropical areas (Perera, 2011). Furthermore, the mutations, especially within the $\alpha 1$ domain, have an impact on the antigen binding groove and could present differential binding ability to wide profiles of antigens in different environments during the long evolutionary history for domesticated animals (Germain, 1995; Williams et al., 2002).

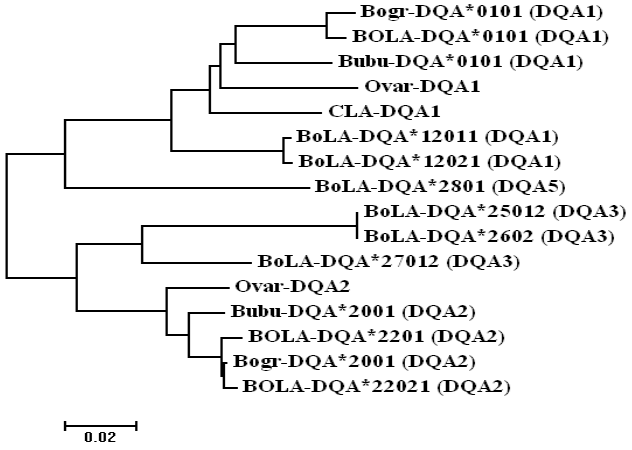

Figure 2. Phylogenetic tree based on DQA nucleotide sequences of yak (neighbor-joining method).

These results showed that the *Bogr-DQA1* and *Bogr-DQA2* genes are more closely related to their cattle counterparts. Similarly, the low nucleotide sequence similarity between Bogr-DQA1 and Bogr-DQA2 as well as the high proportion of nucleotide and amino acid replacements clearly present inconsistencyin allelic form. Moreover, the bovine DQA3*01 and DQA3*02 sequences as non-allelic types have 92 % nucleotide similarity and larger genetic distance within two gene families (Ballingall et al., 1998).

From the phylogenetic tree analysis based on the nucleotide sequences, it appears that the split of the DQA1 and DQA2 sequences from the yak and other ruminants into two major clusters and further reveals the independently evolutionary relationship of the yak DQA sequences (Fig. 2). It was apparent that yak is closest to cattle, as previous results have shown (He et al., 2014; Sun et al., 2014). Moreover, higher divergence between the two branches reveals that the *Bogr-DQA1* and *Bogr-DQA2* belong to two different loci. For cattle and buffalo, there is some evidence to show that the DQA molecules are in the form of duplicated types and can both be expressed (Russell et al., 1997; Niranjan et al., 2009). These duplicated genes with more polymorphisms could be helpful to consolidate the immunological defenses of yak under harsh environmental conditions.

4 Conclusions

The *Bogr-DQA1* and *Bogr-DQA2* genes have been cloned and characterized for the first time, thereby extending the knowledge of the MHC-DQA in ruminants. The genes of *Bogr-DQA* and *Bogr-DQA2* are highly polymorphic, especially in the α1 domain, as in most ruminants. It would be more interesting to study the effect of polymorphisms from *Bogr-DQA1* and *Bogr-DQA2* on the pathogen resistance for yak adaptation in the Himalayan Plateau, with its extremely low temperature and varying oxygen conditions, in subsequent studies.

5 Data availability

All data have been deposited into the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/) under accession numbers JQ864314 and JQ864315.

Acknowledgements. Funding for the study was provided by the Training Programs of Innovation for Undergraduates from Yunnan Agricultural University (Fei Ge) and the National Nature Science Foundation of China under project no. 31460583 and 31101640.

Edited by: S. Maak
Reviewed by: three anonymous referees

References

An, D., Dong, X., and Dong, Z.: Prokaryote diversity in the rumen of yak (*Bos grunniens*) and Jinnan cattle (*Bos taurus*) estimated by 16S rDNA homology analyses, Anaerobe, 11, 207–215, 2005.

Andersson, L. and Rask, L.: Characterization of the MHC class II region in cattle: the number of DQ genes varies between haplotypes, Immunogenetics, 27, 110–120, 1988.

Ballingall, K. T., Marasa, B. S., Luyai, A., and McKeever, D. J.: Identification of diverse BoLA DQA3 genes consistent with non-allelic sequences, Anim. Genet., 29, 123–129, 1998.

Bernatchez, L. and Landry, C.: MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years?, J. Evol. Biol., 16, 363–377, 2003.

Brown, J. H., Jardetzky, T. S., Gorge, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., and Wiley, D. C.: Three dimensional structure of the human class II histocompatibility antigen HLA-DR1, Nature, 364, 33–39, 1993.

Germain, R. N.: The biochemistry and cell biology of antigen presentation by MHC class I and class II molecules, Implications for development of combination vaccines, Ann. N. Y. Acad. Sci., 754, 114–125, 1995.

Glass, E. J., Oliver, R. A., and Russell, G. C.: Duplicated DQ haplotypes increases the complexity of restriction element usage in cattle, J. Immunol., 165, 134–138, 2000.

He, Y., Xi, D., Leng, J., Qian, T., Jin, D., Chen, T., Yang, C., Hao, T., Yang, Z., and Deng, W.: Genetic variability of MHC class II DQB exon 2 alleles in yak (*Bos grunniens*), Mol. Biol. Rep., 41, 2199–2206, 2014.

Kappes, D. and Strominger, J. L.: Human class II major histocompatibility complex genes and proteins, Annu. Rev. Biochem., 57, 991–1028, 1988.

Klein, J.: Seeds of time: Fifty years ago Peter A. Gorer discovered the H-2 complex, Immunogenetics, 24, 331–338, 1986.

Kuduk, K., Babik, W., Bojarska, K., Sliwińska, E. B., Kindberg, J., Taberlet, P., Swenson, J. E., and Radwan, J.: Evolution of major histocompatibility complex class I and class II genes in the brown bear, BMC Evol. Biol., 12, 197, doi:10.1186/1471-2148-12-197, 2012.

Lazzaro, B. P. and Little, T. J.: Immunity in a variable world, Philos. T. Roy. Soc. B, 364, 15–26, 2009.

McKinney, D. M., Southwood, S., Hinz, D., Oseroff, C., Arlehamn, C. S., Schulten, V., Taplitz, R., Broide, D., Hanekom, W. A., Scriba, T. J., Wood, R., Alam, R., Peters, B., Sidney, J., and Sette, A.: A strategy to determine HLA class II restriction broadly covering the DR, DP, and DQ allelic variants most commonly expressed in the general population, Immunogenetics, 65, 357–370, 2013.

Niranjan, S. K., Deb, S. M., Sharma, A., Mitra, A., and Kumar, S.: Isolation of two cDNAs encoding MHC-DQA1 and -DQA2 from the water buffalo, *Bubalus bubalis*, Vet. Immunol. Immunopathol., 130, 268–271, 2009.

Perera, B. M.: Reproductive cycles of buffalo, Anim. Reprod. Sci., 124, 194–199, 2011.

Qiu, Q., Zhang, G., Ma, T., Qian, W., Wang, J., Ye, Z., Cao, C., Hu, Q., Kim, J., Larkin, D. M., Auvil, L., Capitanu, B., Ma, J., Lewin, H. A., Qian, X., Lang, Y., Zhou, R., Wang, L., Wang, K., Xia, J., Liao, S., Pan, S., Lu, X., Hou, H., Wang, Y., Zang, X., Yin, Y., Ma, H., Zhang, J., Wang, Z., Zhang, Y., Zhang, D., Yonezawa, T.,

Hasegawa, M., Zhong, Y., Liu, W., Zhang, Y., Huang, Z., Zhang, S., Long, R., Yang, H., Wang, J., Lenstra, J. A., Cooper, D. N., Wu, Y., Wang, J., Shi, P., Wang, J., and Liu, J.: The yak genome and adaptation to life at high altitude, Nat. Genet., 44, 946–949, 2012.

Reed, D. H. and Frankham, R.: Correlation between fitness and genetic diversity, Conserv. Biol., 17, 230–237, 2003.

Rudd, P. M., Wormald, M. R., Stanfield, R. L., Huang, M., Mattsson, N., Speir, J. A., DiGennaro, J. A., Fetrow, J. S., Dwek, R. A., and Wilson, I. A.: Roles for glycosylation of cell surface receptors involved in cellular immune recognition, J. Mol. Biol., 293, 351–366, 1999.

Russell, B. C., Gallagher, A., Craigmile, S., and Glass, E. J.: Characterization of cattle cDNA sequences from two DQA loci, Immunogenetics, 45, 455–458, 1997.

Sena, L., Schneider, M. P., Brenig, B. B., Honeycutt, R. L., Honeycutt, D. A., Womack, J. E., and Skow, L. C.: Polymorphism and gene organization of water buffalo MHC-DQB genes show homology to the BoLA DQB region, Anim. Genet., 42, 378–385, 2011.

Sigurdardottir, S., Borsch, C., Gustafsson, K., and Andersson, L.: Gene duplications and sequence polymorphism of bovine class II DQB genes, Immunogenetics, 35, 205–213, 1992.

Sun, Y., Zheng, H., Xi, D., Zhang, X., Du, M., Pu, L., Lin, M., and Yang, Y.: Molecular characteristics of the MHC-DRA genes from yak (*Bos grunniens*) and Chinese yakow (*Bos grunniens* × *Bos taurus*), Int. J. Immunogenet., 41, 69–73, 2014.

Tamura, K., Dudley, J., Nei, M., and Kumar, S.: MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0, Mol. Biol. Evol., 24, 1596–1599, 2007.

Tibayrenc, M.: Human genetic diversity and the epidemiology of parasitic and other transmissible diseases, Adv. Parasitol., 64, 377–422, 2007.

Trowsdale, J.: Genetic and functional relationships between MHC and NK receptor genes, Immunity, 15, 363–374, 2001.

Trowsdale, J. and Parham, P.: Mini-review: defense strategies and immunity-related genes, Eur. J. Immunol., 34, 7–17, 2004.

Wiener, G., Han, J. L., and Long, R. J.: Origins, domestication and distribution of yak, and production characteristics of yak, in: The yak, 2nd Edn., Regional Office for Asia and the Pacific of the Food and Agriculture Organization of the United Nations, Bangkok, 136 pp., 2003.

Williams, A., Peh, C. A., and Elliott, T.: The cell biology of MHC class I antigen presentation, Tissue Antigens, 59, 3–17, 2002.

Xi, D., Hao, T., He, Y., Leng, J., Sun, Y., Yang, Y., Mao, H., and Deng W.: Nucleotide sequence and polymorphism of MHC class II DQB exon 2 alleles in Chinese yakow (*Bos grunniens* × *Bos taurus*), Int. J. Immunogenet., 41, 269–275, 2014.