TÜBİTAK

Turkish Journal of Veterinary and Animal Sciences

http://journals.tubitak.gov.tr/veterinary/

Short Communication

Turk J Vet Anim Sci
(2017) 41: 136-141
© TÜBİTAK
doi:10.3906/vet-1510-86

Molecular cloning and sequence characterization of two genes, DQA1 and DQA2, from the Chinese yakow (*Bos grunniens* × *Bos taurus*)

Dongmei XI[1,2], Sameeullah MEMON[1,2], Guozhi LI[1,2], Xiangying LIU[1,2], Chao SU[1,2], Jing LENG[1,2], Liping WANG[1,2], Sehroon KHAN[3,4] , Weidong DENG[1,2,*]

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

Received: 30.10.2015 • **Accepted/Published Online:** 02.06.2016 • **Final Version:** 21.02.2017

Abstract: The major histocompatibility complex (MHC) plays a crucial role in the immune response to infectious diseases. The aim of the present study was to investigate the structural and functional characteristics and possible duplication of the *MHC-DQA* genes in the Chinese yakow, which is well adapted to the Qinghai–Tibetan Plateau. Two full-length cDNA sequences and partial 5'- and 3'-untranslated regions (UTRs) were retrieved and denoted as Bogr×BoLA-DQA*0101 (DQA1) and Bogr×BoLA-DQA*2001 (DQA2), respectively. The percentage of nucleotide and amino acid sequences homology between Bogr×BoLA-DQA1 and Bogr×BoLA-DQA2 showed that these two sequences have more identity to alleles of the corresponding *DQA1* and *DQA2* genes from other counterparts than to each other. Moreover, the *Bogr×BoLA-DQA1* and *-DQA2* exhibited considerable variation with 99 nucleotide polymorphisms along the length of the coding regions, leading to 52 amino acid polymorphisms. The phylogenetic tree investigation also demonstrated that there is a larger genetic gap within these two identified genes than within orthologous genes from different animals, suggesting that these sequences belong to nonallelic and duplicated types in the Chinese yakow. This is the first report to isolate and identify the *MHC-DQA1* and *-DQA2* genes. This result will deepen our understanding of MHC diversity among different ruminant species based on comparative population genetics.

Key words: Major histocompatibility complex (MHC), Chinese yakow (*Bos grunniens* × *Bos taurus*), *DQA1* and *DQA2* genes, molecular structure and function

In both humans and animals, the major histocompatibility complex (MHC) molecules cluster plays a vital role against antigens, and thus is helpful for knowing the hosts' resistance or susceptibility to infectious parasites and pathogens (1–3). The mammalian MHC gene is divided into functional regions including class I and class II (4). The MHC class II DQ and DR molecules are known to be associated with many diseases (5–8). In some vertebrates, there is a single copy of the DQ gene for the rat, pig, mouse, and rabbit, whereas in others like humans and dogs, multiple DQ genes have been observed but expression is limited to one copy only (9).

Variability in the number of DQ loci is reported in ruminants. For cattle, most haplotypes carry duplicated DQ genes that are mapped to the autosome 23q13-23 (10–12). Evidence suggests that in these cases both DQ molecules are expressed (13). Recognition of a wider range of nonself-bodies is thought to have a combined effect of polymorphism along with the duplication of the MHC genes at the surface of the cell.

The Chinese yakow is a remarkable hybrid of yak and Yellow cattle. They are often used for transportation, ploughing, milk, meat, and even cash income for local herdsman in the harsh environment of the Qinghai–Tibetan Plateau (14). In farming the pastoral region, local farmers have favored to feed the Chinese yakow over the yak for adapting agricultural activities due to their power and tolerance (15).

To date, the *MHC-DQA1* and *DQA2* genes in this species remain unexplored. In the present study, we aimed to isolate and characterize the *MHC-DQA1* and *DQA2* genes and compare the amino acid sequence with the consensus sequence of its counterpart for finding the specific characterization from the Chinese yakow,

* Correspondence: 1692425306@qq.com

which could suffer due to long-term selection during close agricultural activities with local humans. This work will be helpful in understanding MHC diversity, as well as duplication in the immunological defense gene of the Chinese yakow, and the high genetic variation in the *DQA* gene that might be generated to recognize species as well as region-specific antigens.

In October 2013, five apparently healthy mature Chinese yakow (*Bos grunniens* × *Bos taurus*, 3♀ and 2♂) were sacrificed in the official slaughterhouse of Shangri La City, Diqing District, Yunnan Province, China. Liver samples were collected quickly and stored at –80 °C. The total RNA was extracted using a commercial kit (Beijing Tiangen Biotech Co., Ltd, Beijing, China). Furthermore, the DNase I was added to the RNA, and the cDNA was constructed following the manufacturer's instructions, using RevertAidTM First Strand cDNA Synthesis Kits (Fermentas Inc., Ontario, Canada).

The three primers previously used to retrieve the buffalo DQA genes (16) were used to clone the complete cDNA fragments of the Chinese yakow *Bogr×BoLA-DQA1* and *Bogr×BoLA-DQA2* genes. The forward primer (A1A2F: 5'-ACCTTGAGAAGAGGATGGTCCTG-3') was shared. The other two reverse primers (A1R: 5'-ATTGCACCTTCCTTCTGGAGTGT-3' and A2R: 5'-TCATAGATCGGCAGAACCACCTT-3') were different. Therefore, the combined primers A1A2F and A1R, and A1A2F and A2R were used to amplify the Bogr×BoLA-DQA1 and Bogr×BoLA-DQA2 fragments, respectively. Using a Bioer Life Express Thermocycler, the PCR was carried out in a reaction volume of 25 µL, containing 2.0 µL template cDNA (about 100 ng/µL), 12.5 µL PCR Power Mix, 1.0 µL 10 pmoL/µL of each primer, and 8.5 µL of double-distilled water. The PCR amplification program was as follows: denaturation at 94 °C for 3 min, followed by 35 cycles; at 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. Amplicons were visualized on agarose gel stained with ethidium bromide. Finally, the PCR products were sequenced bi-directionally using an ABI 3730 DNA Analyzer (Applied Biosystems Inc.) with the above primers at the Sun Biotechnology Company (Beijing, China). All samples were sequenced twice.

The gene analysis for the cDNA sequence was conducted using GenScan software (http://genes.mit.edu/GENSCAN.html). The sequences prediction was conducted using the ORF Finder software (http://www.ncbi.nlm.nih.gov/projects/gorf/). The theoretical isoelectric point (pI) and molecular weight (Mw) of proteins were computed using the online Compute pI/Mw Tool (http://www.expasy.org/tools/pi_tool.html). The complete cDNA and putative amino acid sequences were compared with the orthologous sequences. Phylogenetic analysis was done using MEGA software version 4 (17) by neighbor-joining method for coding regions of different DQA orthologous fragments from different species.

Two nucleotide sequences of 783 bp (Bogr×BoLA-DQA1) and 815 bp (Bogr×BoLA-DQA2) were amplified using the template cDNA from Chinese yakow. The cDNA sequence identification showed that the two genes are not homologous to any of the known Chinese yakow genes and it was then deposited into the GenBank database with accession numbers JQ904621 (for DQA1) and JQ904622 (for DQA2). Moreover, these sequences were submitted to the Immunopolymorphism database (www.ebi.ac.uk/ipd/mhc/bola/nomenclature) with the official names Bogr×BoLA-DQA*0101 (DQA1) and Bogr×BoLA-DQA*2001 (DQA2), respectively, based on BoLA-DQA sequence similarity.

The sequence predictions were conducted and the results showed that the 783 bp (including 7 bp and 8 bp 5'- and 3'-untranslated regions) and 815 bp cDNA fragments (including 6 bp and 41 bp 5'- and 3'-untranslated regions) denote two single genes containing a complete open reading frame (ORF) of 768 nucleotides, with both encoding a polypeptide of 255 amino acids. The pI of Chinese yakow DQA1 and DQA2 proteins was 5.52 and 4.84, respectively. The Mw of the two assigned molecules was 28,202.35 and 27,904.82, respectively.

The *Bogr×BoLA-DQA* genes were compared with the reference *BoLA-DQA* gene sequences with GenBank accession numbers Y07898 and Y07820. The Bogr×BoLA-DQA1 and -DQA2 showed the highest similarities (96% and 99%) at the nucleotide sequence level, with that of BoLA-DQA1 and -DQA2, respectively (Table). However, the sequence homology percentage between the Bogr×BoLA-DQA1 and -DQA2 was 86% less than that of cattle (BoLA-DQA) (Table). These findings were similar to the results from water buffalo (16). However, they illustrated that the *Bubu-DQA* genes have less similarity (93.9% and 97.7%) with that of cattle as compared to the fragment homology between the DQA genes (85.7%).

In fact, the Bogr×BoLA-DQA1 and -DQA2 exhibited considerable variation with 99 nucleotide polymorphisms along the length of the coding regions (Figure 1), leading to 52 amino acid polymorphisms including 2 in the signal peptide (SP), 32 in the α1, 12 in the α2, 2 in the connecting peptide (CP), 2 in the transmembrane (TM), and 2 in the cytoplasmic (CY) domains (Figure 2). Nevertheless, the Chinese yakow has more amino acid substitutions than buffaloes with 45 amino acids differences (16).

In the present study, 20 peptide binding sites (PBSs) were retrieved (Figure 2). Within them, only 7 residues corresponding to the positions 11, 29, 35, 57, 60, 63, and 70 were fixed between DQA1 and DQA2 molecules from the investigated different animal species. The remaining

Table. Sequence identity comparisons from the α1, α2, and CP/TM/CY domains between *Bogr×BoLA-DQA1/DQA2* and *BoLA-DQA1/DQA2* genes (the amino acid identity is shown in parentheses).

	Bogr×BoLA-DQA*2001 (DQA2)	BoLA-DQA*0101 (DQA1)	BoLA-DQA*2201 (DQA2)
Bogr×BoLA-DQA*0101 (DQA1)			
α1	79.0	96.0	79.0
α2	88.0	95.0	88.0
CP/TM/CY	91.0	98.0	91.0
Entire gene	86.0 (79.0)	96.0 (92.0)	86.0 (78.0)
Bogr×BoLA-DQA*2001 (DQA2)			
α1		80.0	100.0
α2		88.0	98.0
CP/TM/CY		91.0	100.0
Entire gene		86.0 (79.0)	99.0 (98.0)
BoLA-DQA*0101 (DQA1)			
α1			80.0
α2			87.0
CP/TM/CY			91.0
Entire gene			85.0 (78.0)

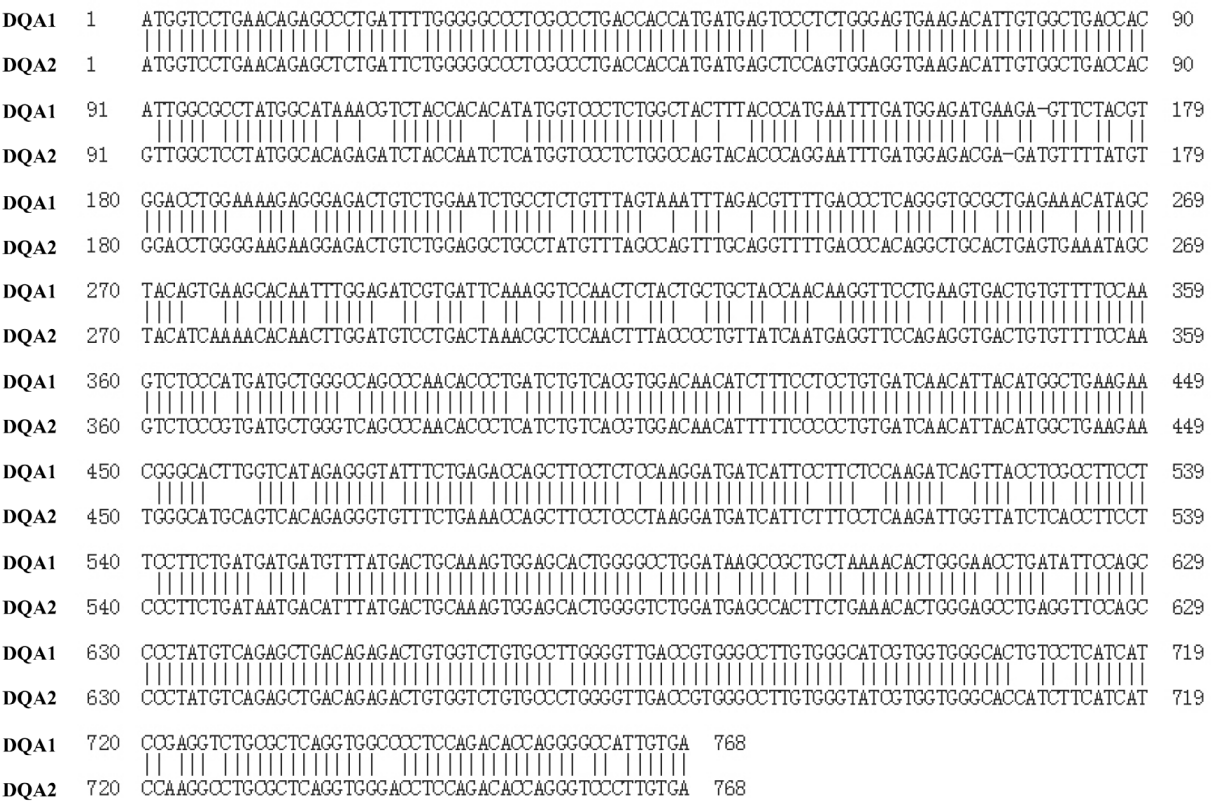

Figure 1. Nucleotide polymorphism comparison between the *Bogr×BoLA-DQA1* and *Bogr×BoLA-DQA2*.

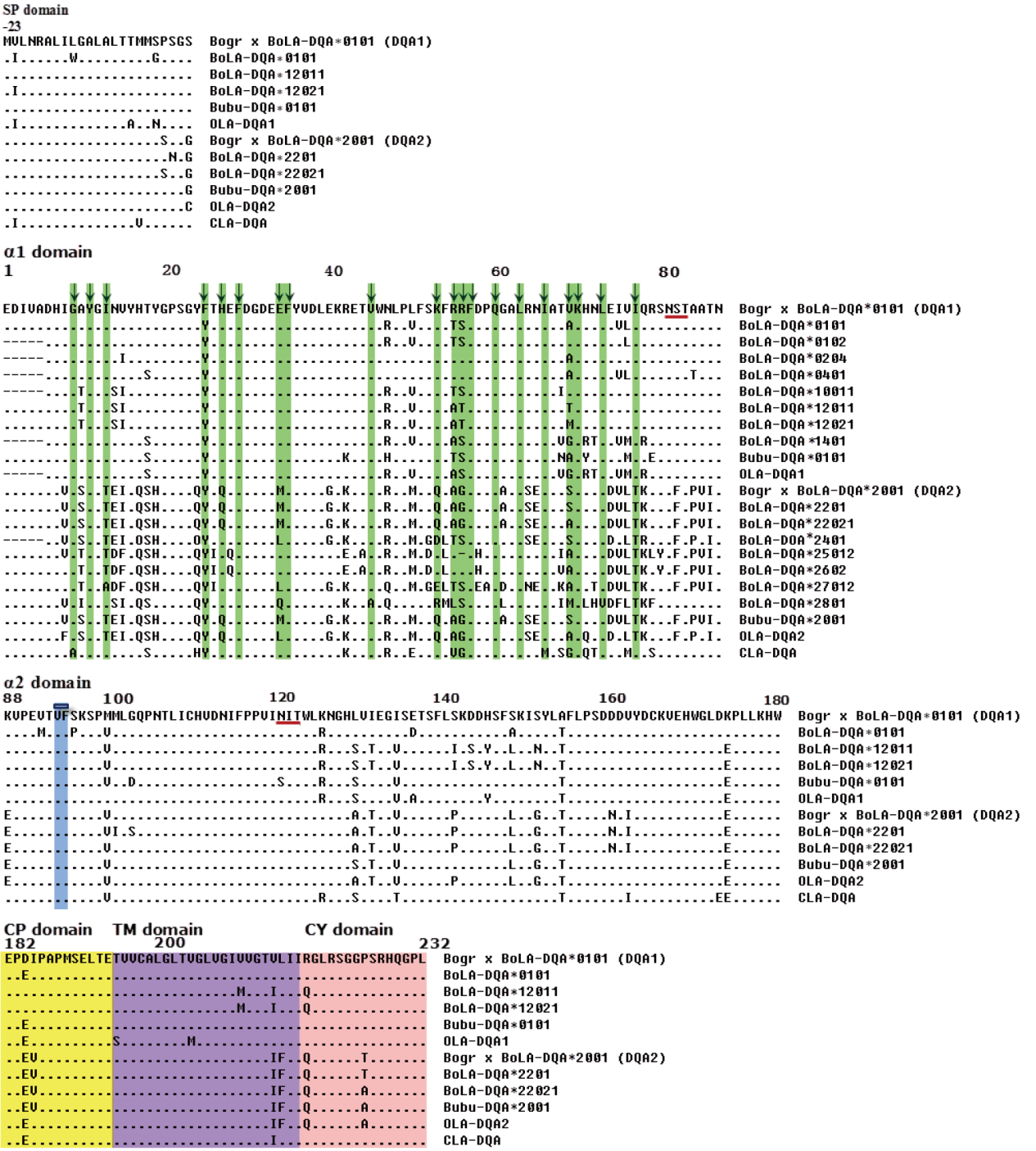

Figure 2. Amino acid sequences alignment between the Bogr×BoLA-DQA and orthologous DQA molecules [the green arrows denote the peptide binding sites (PBS). The deduced N-linked glycosylation sites are underlined (▬). The square (■) denotes the position of residues associated with binding of CD^{4+} molecules. A point (.) denotes amino acid identity and hyphen (-) indicates gap inserted to maximize]. 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), respectively. 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), respectively.

13 PBS functional sites had different amino acid residues in the two polypeptide sequences. In addition, the Bogr×BoLA-DQA1 molecule possessed one rare residue at position 25 within the PBS motif that is uncommon in other investigated animal species (Figure 2). Obviously, with the advantage of such mutations, the nonsynonymous replacements at the PBS of the exon 2 region might have affected the antigen binding groove and could demonstrate differential ability binding to a wide spectrum of pathogens for adapting to cold and hypoxia environments during the long evolutionary history (18,19).

These results revealed that the *Bogr×BoLA-DQA1* and *-DQA2* genes are more homologous with the corresponding sequences from their counterpart. Within coding regions, the replacements from 99 nucleotides and their corresponding 52 nonsynonymous mutations between Bogr×BoLA-DQA1 and -DQA2 molecules clearly indicated inconsistency as an allelic form. Moreover, the TM and CY domains of the bovine MHC class II genes often displayed locus-specific nucleotide sequence motifs (20). Different bovine DQA locus-specific motifs have also been identified in cattle and buffalo DQA sequences (16). Those motifs differ between the *Bogr×BoLA-DQA1* and *-DQA2* genes, providing further evidence that these sequences might be nonallelic type.

From the phylogenetic analysis based on the investigated nucleotide sequences, the DQA1 and DQA2 sequences from the Chinese yakow and other animals are separated into two major groups and this further indicates their independent evolutionary history (Figure 3). It was apparent that the Chinese yakow is closest to cattle as shown by previous results (14,21). Moreover, higher deviation between the two clusters shows that the Bogr×BoLA-DQA1 and -DQA2 fall into two different loci.

For buffalo and cattle, there is some evidence to demonstrate that the DQA molecules belong to duplicated type and can be expressed together (13,16), which also looks to be similar for the Chinese yakow. Gene duplication is a frequent case in eukaryotic organisms including yeast, plants, and animals. When the cellular hardware creates two copies of a gene, gene duplication occurs. Recent studies have demonstrated that the proportion of duplicated genes in mammalian genomes is correlated with environmental variability within a habitat. Moreover, the species under low habitat variability have a higher proportion of lost duplicated genes, particularly small-scale duplication genes, than those under high habitat variability. These events are an essential source of genetic originality leading to evolutionary novelty. One copy, freed from selection, could adapt to a new function, or be turned off or even serve as a "spare" if the original gene is damaged. Alternatively, the original function can be dissected, giving different roles to each copy of the gene. These results reveal that the species that inhabit variable environments may maintain more SSD genes in their genomes and hint that SSD genes are important for adapting to newly environments and surviving after environmental changes (22,23). Therefore, these duplicated genes (DQA1 and DQA2) with more mutations could improve the immunological ability to adapt to harsh environments for the Chinese yakow.

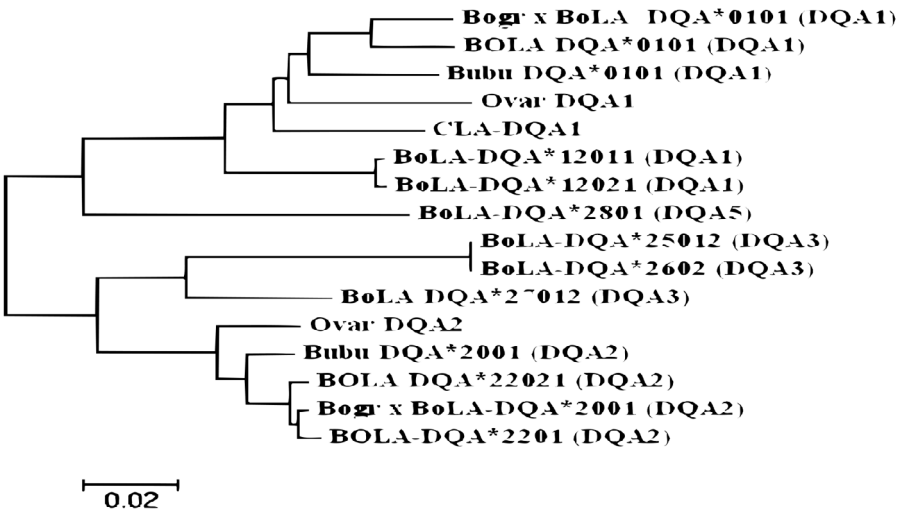

Figure 3. Phylogenetic tree based on the DQA nucleotide sequences of Chinese yakow (neighbor-joining method). The reference GenBank accession numbers for phylogenetic tree are Y07898 (BoLA-DQA*0101), DQ440647 (Bubu-DQA*0101), M93430 (OLA-DQA1), AY464652 (CLA-DQA), Y07819 (BoLA-DQA*12011), D50454 (BoLA-DQA*12021), AF037314 (BoLA-DQA*2801), Y14020 (BoLA-DQA*25012), Y14021 (BoLA-DQA*2602), Y14022 (BoLA-DQA*27012), M93433 (OLA-DQA2), DQ440648 (Bubu-DQA*2001), D50045 (BoLA-DQA*22021), and Y07820 (BoLA-DQA*2201), respectively.

In summary, the two cDNAs encoded by the *Bogr×BoLA-DQA1* and *-DQA2* genes have been amplified and characterized for the first time, therefore expanding our knowledge of the MHC-DQA for ruminants. The *Bogr×BoLA-DQA* and *-DQA2* genes are highly variable, especially in the α1 domain as in most ruminants. It would be more interesting to decrypt the effect of variability from Bogr×BoLA-DQA1 and -DQA2 on Chinese yakow resistance to adapt to extremely low temperature and hypoxia conditions in future.

Acknowledgment
Funding was provided by the National Nature Science Foundation of China under project numbers 31460583 and 31101640.

References

1. Park YH, Joo YS, Park JY, Moon JS, Kim SH, Kwon NH, Ahn JS, Davis WC, Davies CJ. Characterization of lymphocyte subpopulations and major histocompatibility complex haplotypes of mastitis-resistant and susceptible cows. J Vet Sci 2004; 5: 29-39.
2. Castillo S, Srithayakumar V, Meunier V, Kyle CJ. Characterization of major histocompatibility complex (MHC) DRB exon 2 and DRA exon 3 fragments in a primary terrestrial rabies vector (*Procyon lotor*). PLoS ONE 2010; 5: e12066.
3. Ellis SA, Codner G. The impact of MHC diversity on cattle T cell responses. Vet Immunol Immunopathol 2012; 148: 74-77.
4. Sin YW, Dugdale HL, Newman C, Macdonald DW, Burke T. MHC class II genes in the European badger (*Meles meles*): characterization, patterns of variation, and transcription analysis. Immunogenetics 2012; 64: 313-327.
5. Hughes AL, Yeager M. Natural selection at major histocompatibility complex loci of vertebrates. Annu Rev Genet 1998; 32: 415-435.
6. Gelhaus A, Förster B, Wippern C, Horstmann RD. Evidence for an additional cattle DQA locus, BoLA-DQA5. Immunogenetics 1999; 49: 321-327.
7. Hameed KGA, Sender G, Maynt M. Major histocompatibility complex polymorphism and mastitis resistance – a review. Anim Sci Pap Rep 2006; 24: 11-25.
8. Miyasaka T, Takeshima SN, Matsumoto Y, Kobayashi N, Matsuhashi T, Miyazaki Y, Tanabe Y, Sentsui H, Aida Y. The diversity of bovine MHC class II DRB3 and DQA1 alleles in different herds of Japanese Black and Holstein cattle in Japan. Gene 2011; 472: 42-49.
9. Kappes D, Strominger JL. Human class II major histocompatibility complex genes and proteins. Annu Rev Biochem 1988; 57: 991-1028.
10. Andersson L, Rask L. Characterization of the MHC class II region in cattle: the number of DQ genes varies between haplotypes. Immunogenetics 1988; 27: 110-120.
11. Sigurdardottir S, Borsch C, Gustafsson K, Andersson L. Gene duplications and sequence polymorphism of bovine class II DQB genes. Immunogenetics 1992; 35: 205-213.
12. Fries R, Hediger R, Stranzinger G. Tentative chromosomal localization of the bovine major histocompatibility complex by *in situ* hybridization. Anim Genet 1986; 17: 287-294.
13. Russell BC, Gallagher A, Craigmile S, Glass EJ. Characterization of cattle cDNA sequences from two DQA loci. Immunogenetics 1997; 45: 455-458.
14. Xi D, Wu M, Fan Y, Huo Y, Leng J, Gou X, Mao H, Deng W (2012). Isolation and characteristics of the melanocortin 1 receptor gene (*MC1R*) in the Chinese yakow (*Bos grunniens* × *Bos taurus*). Gene 498: 259-263.
15. Xi D, Hao T, He Y, Leng J, Sun Y, Yang Y, Mao H, Deng W. Nucleotide sequence and polymorphism of MHC class II DQB exon 2 alleles in Chinese yakow (*Bos grunniens* × *Bos taurus*). Int J Immunogenet 2014; 41: 269-275.
16. Niranjan SK, Deb SM, Sharma A, Kumar S, Mitra A. Isolation of two cDNAs encoding MHC-DQA1 and DQA2 from the water buffalo *Bubalus bubalis*. Vet Immunol Immunopathol 2009; 130: 268-271.
17. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 2007; 24: 1596-1599.
18. Germain RN, Margulies DH. The biochemistry and cell biology of antigen processing and presentation. Annu Rev Immunol 1993; 11: 403-450.
19. Williams MA, Trout R, Spector SA. HIV-1 gp120 modulates the immunological function and expression of accessory and co-stimulatory molecules of monocyte-derived dendritic cells. J Hematother Stem Cell Res 2002; 11: 829-847.
20. Ballingall KT, Marasa BS, Luyai A, McKeever DJ. Identification of diverse BoLA DQA3 genes consistent with non-allelic sequences. Anim Genet 1998; 29: 123-129.
21. Sun Y, Zheng H, Xi D, Zhang X, Du M, Pu L, Lin M, Yang Y. Molecular characteristics of the MHC-DRA genes from yak (*Bos grunniens*) and Chinese yakow (*Bos grunniens* × *Bos taurus*). Int J Immunogenet 2014; 41: 69-73.
22. Tamate SC, Kawata M, Makino T. Contribution of nonohnologous duplicated genes to high habitat variability in mammals. Mol Biol Evol 2014; 31: 1779-1786.
23. Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, Fan S, Simakov O, Ng AY, Lim ZW, Bezault E et al. The genomic substrate for adaptive radiation in African cichlid fish. Nature 2014; 513: 375-381.