

RESEARCH ARTICLE



## Isolation and characterization of the major histocompatibility complex *DQA1* and *DQA2* genes in gayal (*Bos frontalis*)

SAMEEULLAH MEMON, LIPING WANG, GUOZHI LI, XIANGYING LIU, WEIDONG DENG and DONGMEI XI\*

Yunnan Provincial Key Laboratory of Animal Nutrition and Feed, Faculty of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201, People's Republic of China

\*For correspondence. E-mail: 2522935343@qq.com.

Received 2 August 2016; revised 10 February 2017; accepted 20 June 2017; published online 28 February 2018

**Abstract.** The species origin of Yunnan gayal has been controversial since many years. However, few recent genetic studies have suggested that it has perhaps originated from the hybridization between male *Bos frontalis* and female *B. taurus* or *B. indicus*. Being an important semi-wild bovid species, this has also been listed under the red list of International Union of Conservation of Nature and Natural Resources. However, there is limited information available about the immunogenicity of this precarious species of *Bos*. Major histocompatibility complex (MHC) plays a pivotal role in immune response to infectious diseases in vertebrates. In the present study, we have investigated the structural and functional characteristics and possible duplication of the MHC-DQA genes in gayal (*B. frontalis*). Two full-length cDNA clones of the MHC-DQA genes were amplified and designated as Bofr-DQA1 (DQA\*0101) and Bofr-DQA2 (DQA\*2001) with GenBank accession numbers KT318732 and KT318733, respectively. A comparison between Bofr-DQA1, Bofr-DQA2 and to other MHC-DQA molecules from different animal species showed that nucleotide and encoded amino acid sequences of these two identified MHC-DQA genes have more similarity to alleles of specific DQA1 and DQA2 molecules from other *Ruminantia* species than to each other. The phylogenetic investigation also demonstrated a large genetic distance between these two genes than to homologous from the other species. The large genetic distance between Bofr-DQA1 and Bofr-DQA2, and the presence of different bovine DQA putative motifs clarify that these sequences are nonallelic type. These results could suggest that duplication of the *DQA* genes has also occurred in gayal. The findings of the present study have strengthened our understanding to MHC diversity in rare ruminants and mutation of immunological functions, selective and evolutionary forces that affect MHC variation within and between species.

**Keywords.** major histocompatibility complex; gayal; *DQA1* gene; *DQA2* gene; immunity; polymorphisms; *Bos frontalis*.

### Introduction

The potential of an organism for evolutionary interactions with pathogens or other species as well as the fitness are related to immunological functions (Lazzaro and Little 2009). The extent of genetic diversity is known to be associated with the capacity for adaptation and evolution to environmental changes (Reed and Frankham 2003). Diversity of genes was pivotal phenomenon for immune

functions which could be associated with the resistance or susceptibility to pathogens (Tibayrenc 2004; Trowsdale and Parham 2004). A cluster of associated genes named the MHC plays a key role in presenting antigenic peptides to T lymphocytes (Klein 1986). In the vertebrate genome, MHC has been known to be the most variable genes, which seem to be maintained by balancing selection, predating speciation events and reflecting the coevolution of hosts with their pathogens (Bernatchez and Landry 2003).

Electronic supplementary material: The online version of this article (<https://doi.org/10.1007/s12041-018-0882-3>) contains supplementary material, which is available to authorized users.

The class II genes of MHC encoded for  $\alpha$  and  $\beta$  chains of DR and DQ dimer molecule, which present antigenic peptides to the helper T cells (McKinney et al. 2013). In rat, mouse, rabbit and pig, there was a single gene of the DQ genes, whereas in dogs and human multiple DQ gene 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 varies in different species e.g., 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 genes are expressed (Russell et al. 1997). Hence, the polymorphisms as well as the duplication of DQ gene increase the differences at the cell surface by interhaplotype and intrahaplotype pairing of  $\alpha$  and  $\beta$  chains during dimerization. The formation of functional restriction elements is the result of interhaplotype combination of DQA and DQB molecules with duplicated DQA haplotypes (Glass et al. 2009).

Gayal or mithun (*Bos frontalis*) is a natural inhabitant of hilly forests, in India it is kept by ethnic groups living in the hills of Tripura, Mizoram, Assam, Arunachal Pradesh, Nagaland and Chittagong hill tracts. They are also found in the Trung and Salween river basins of northern Burma and Yunnan province of China (Simoons 1984). Gayal is an important source of meat in these areas than the other cattle and is considered to possess high percentage resistance to diseases (Rajkhowa et al. 2004). Gayal normally intakes local bamboo and other plant leaves and grasses but possesses high range (from cold to tropical regions) of adaptation to harsh environment (Zhao et al. 2003; Xi et al. 2007). However, its genetic composition has been controversial as many biologists regarded the gayal as the domestic gaur for morphological similarity between the gayal (*B. frontalis*) and the gaur (*B. gaurus*) (Walker et al. 1968; Lan et al. 1993; Nie et al. 1995). The findings of karyotyping, mt-DNA and Y-chromosome analyses have made the scenario little more complex. However, most studies have suggested that gaur has been one of the immediate species ancestor of gayal (Nie et al. 1995; Verkaar et al. 2004; Chi et al. 2005; Gou et al. 2010; Sun et al. 2014). A recent investigation of Yunnan gayal suggested that maternal lineages of both Yunnan gayal and cattle were the admixture of *B. indicus* and *B. taurus*, while the Y chromosomal phylogeny indicated that their parental lineages are almost *B. frontalis* and *B. indicus*, respectively (Gou et al. 2010).

In the present study, we have isolated and characterized two cDNAs of DQA1 and DQA2 from gayal and compared with other homologues MHC sequences from other animal species with regards to determine the disease resistance and susceptibility genetic factors. This work will possibly strengthen our understanding to the disease control in pet animals as well as in knowing MHC diversity in common ruminants. The study will assist to explore new horizons to investigate immunological functions, selective

and evolutionary forces that affect MHC variation within and between species.

## Materials and methods

Three healthy gayal (*B. frontalis*) liver samples were collected from the National Jiumudang Stud Gayal Farm, Gongshan, China. The RNA extraction was performed using the commercial kit (Tiangen Biotech, Beijing, China). The extracted RNA was incubated with DNaseI to cleave the DNA contamination. The cDNA was constructed using the commercial RevertAid First Strand cDNA synthesis kit (Fermentas, Ontario, Canada), following the manufacturers' protocol.

The *Bofr-DQA1* (784 bp) and *Bofr-DQA2* (801 bp) fragments were amplified from the template cDNA of gayal using three primer pairs, i.e. A1A2F, A1R and A2R, published previously for swamp buffalo (Niranjan et al. 2009). The forward primer (A1A2F: 5'-ACCTTGAGAAGAG GATGGTCCTG-3') was shared on the consensus region. The other two reverse primers (A1R: 5'-ATTGCACCTTC CTTCTGGAGTGT-3' and A2R: 5'-TCATAGATCGGC AGAACCACCTT-3') were different for both the *DQA1* and *DQA2*. By using the combined primers A1A2F, A1R and A2R, the two primers (A1A2F and A1R) amplified the *Bofr-DQA1* and the additional two primers (A1A2F and A2R) amplified the *Bofr-DQA2* fragments, respectively. Using Bioer Life Express Thermocycler, the polymerase chain reaction (PCR) was performed in a reaction volume of 25  $\mu$ L, containing 2.0  $\mu$ L template cDNA, 12.5  $\mu$ L PCR Power Mix, 1.0  $\mu$ L 10 pmoL  $\mu$ L<sup>-1</sup> of each primer, and 8.5  $\mu$ L double-distilled water. The PCR cycle was 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. Finally, the PCR products were sequenced bidirectionally using an ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, USA).

The cDNA sequences were translated to amino acid sequences using GenScan software (<http://genes.mit.edu/GENSCAN.html>) and compared with the orthologous sequences. The theoretical isoelectric point (pI) and molecular weight (Mw) of the two putative proteins of the gayal genes were also computed using the online pI/Mw tool ([http://www.expasy.org/tools/pi\\_tool.html](http://www.expasy.org/tools/pi_tool.html)). The sequence predictions were made using the open reading frame (ORF) Finder software (<http://www.ncbi.nlm.nih.gov/projects/gorf/>) and the neighbour-joining phylogenetic tree was constructed using MEGA software based on the coding regions of different orthologous DQA alleles from different species (Tamura et al. 2007). The nonsynonymous ( $d_n$ ) and synonymous substitution ( $d_s$ ) ratios between gayal and other livestock species in the genes *DQA1* and *DQA2* have been estimated

**Table 1.** Sequence comparison of the  $\alpha 1$ ,  $\alpha 2$ , CP/TM/CY motifs between the *Bofr-DQA1/DQA2* and *BoLA-DQA1/DQA2* genes.

	<i>Bofr-DQA*2001 (DQA2)</i>	<i>BoLA-DQA*0101 (DQA1)</i>	<i>BoLA-DQA*2201 (DQA2)</i>
<i>Bofr-DQA*0101 (DQA1)</i>			
$\alpha 1$	81.0	91.0	81.0
$\alpha 2$	90.0	87.0	90.0
CP/TM/CY	92.0	94.0	92.0
Entire gene (protein)	88.0 (80.0)	91.0 (85.0)	88.0 (80.0)
<i>Bofr-DQA*2001 (DQA2)</i>			
$\alpha 1$		80.0	100.0
$\alpha 2$		87.0	100.0
CP/TM/CY		91.0	100.0
Entire gene (protein)		85.0 (78.0)	100.0 (100.0)
<i>BoLA-DQA*0101 (DQA1)</i>			
$\alpha 1$			80.0
$\alpha 2$			87.0
CP/TM/CY			91.0
Entire gene (protein)			85.0(78.0)

The amino acid identity is shown in parentheses.

using the software PAML (Yang 2007) and the significant changes that has altered the amino acids among livestock with respect to gayal have been investigated by the web version of PAL2NAL (<http://www.bork.embl.de/pal2nal/>).

## Results and discussion

We searched the most homologous sequences for *Bofr-DQA1* and *Bofr-DQA2* genes using the BLAST tool of NCBI server (<http://www.ncbi.nlm.nih.gov/BLAST>). The sequence similarity search has revealed that the two genes were not similar to any of the known gayal genes but possesses high similarity to other ruminant genes. The nucleotide sequences of *DQA1* and *DQA2* were deposited to the NCBI GenBank database with accession number KT318732 and KT318733, respectively. Further, the sequences were also deposited to Immunopolymorphism database ([www.ebi.ac.uk/ipd/mhc/bola/nomenclature](http://www.ebi.ac.uk/ipd/mhc/bola/nomenclature)) with the assigned official names as *Bofr-DQA\*0101* (for *Bofr-DQA1*) and *DQA\*2001* (for *Bofr-DQA2*). The sequence prediction showed that the 784 bp and 801 bp cDNA sequences only represent two single genes with an ORF of 768 bp and both encoding a polypeptide of 255 amino acid residues. The computed pI of gayal *DQA1* and *DQA2* genes were 4.93 and 4.84, respectively. The computed Mw of the two putative proteins were 28298.34 and 27953.88 Da for *Bofr-DQA1* and *Bofr-DQA2*, respectively.

The nucleotide sequence comparison of *Bofr-DQA* with *BoLA-DQA* genes for homology showed that the *Bofr-DQA1* and *Bofr-DQA2* possess 91 and 100% sequence identities with that of *BoLA-DQA1* and *BoLA-DQA2*, respectively (table 1). However, the nucleotide sequence

identity between the *Bofr-DQA1* and *Bofr-DQA2* were 88% only. These findings corroborate to the study conducted by Niranjan *et al.* (2009) on water buffalo. However, these authors presented that the *Bubu-DQA* genes have different identity (93.9 and 97.7%) with that of cattle as compared to the sequence homology between the *DQA1* and *DQA2* genes (85.7%).

A considerable mutations of 49 amino acid polymorphisms were observed when *Bofr-DQA1* and *Bofr-DQA2* were compared to other alleles which resulted from 95 nucleotide polymorphisms within the coding regions (figure 1). A total of 29 amino acid replacements were found within the exon 2 motif ( $\alpha 1$ ), deduced from the 51 of the nucleotide mutations. The remaining amino acid differences including four in SP domain, 11 in the  $\alpha 2$  domain, two in the connecting peptide (CP), two in the transmembrane (TM) region and one in the cytoplasmic (CY) domain were observed. These results demonstrated that gayal has more amino acid substitutions than buffaloes with 45 amino acids variation (Niranjan *et al.* 2009).

Additionally, the peptide-binding sites (PBS, marked by green arrow sign), one N-glycosylation (NFT) within the  $\alpha 1$  domain and another (NIT) within the  $\alpha 2$  domain, one intrapeptide disulphide bond and the CD4+ binding site (marked by square) were identified, revealing the significance of maintaining their molecular conformation and function to against the invading pathogens (Rudd *et al.* 1999). There were 20 PBS (figure 1) which are specific functional motifs in contacting with the antigens (Brown *et al.* 1993; Kuduk *et al.* 2012). The highly conserved loci from different animal species were only eight residues at positions 11, 25, 29, 35, 57, 60, 63 and 70 between *DQA1* and *DQA2* homologues. The other 12 PBS sites had the different amino acids in both the polypeptide chains, demonstrating that it could have associated with gayal

SP domain

```

-23
MULNRALILGALALTTMTSLCGS   Bofr-DQA 0101 (DQA1)
.I.....V.....MGPS..     BoLA-DQA 0101
.....M.PS..                BoLA-DQA 12011
.I.....M.PS..                BoLA-DQA 12021
.....M.PS..                Bubu-DQA 0101
.I.....A.MNPS..            OLA-DQA1
.....M.PN.G                 Bofr-DQA 2001 (DQA2)
.....M.PN.G                 BoLA-DQA 2201
.....M.SS.G                 BoLA-DQA 22021
.....M.PS.G                 Bubu-DQA 2001
.....M.PS.C                 OLA-DQA2
.I.....UM.PS..             CLA-DQA

```

α1 domain

1	20	40	60	80		
EDIUADHIGTYGVSFVHSYGPSGYV	IIEFDGDEFF	YUDLEKRETU	WHLPUFSEFASFDP	QDGLRNITAKHTLEIM	IRRSNFTAVIN	Bofr-DQA 0101 (DQA1)
.....A.....I..NU...T.....	.....T.....	.....R.....K...T.....	.....GA.....N..UL.Q...S...AT	.....U..N...L.Q...S...AT	.....S...AT	BoLA-DQA 0101
.....A.....I..NU...T.....	.....T.....	.....R.....K...T.....	.....GA.....U..N...L.Q...S...AT	.....U..N...L.Q...S...AT	.....S...AT	BoLA-DQA 0102
.....A.....I..NI...T.....	.....T.....	.....N..L...K.RR...	.....GA.....N..V...Q...S...AT	.....U..N...L.Q...S...TAT	.....S...AT	BoLA-DQA 0204
.....A.....I..NU...T.....	.....T.....	.....N..L...K.RR...	.....GA.....N..UL.Q...S...TAT	.....U..N...L.Q...S...AT	.....S...AT	BoLA-DQA 0401
.....I.....I...T.....	.....T.....	.....R.....K...T.....	.....GA.....I..U..N...V...Q...S...AT	.....U..N...L.Q...S...AT	.....S...AT	BoLA-DQA 10011
.....I.....I...T.....	.....T.....	.....R.....K...T.....	.....GA.....T..N...V...Q...S...AT	.....U..N...L.Q...S...AT	.....S...AT	BoLA-DQA 12011
.....I.....I...T.....	.....T.....	.....R.....K...T.....	.....GA.....H..N...V...Q...S...AT	.....U..N...L.Q...S...AT	.....S...AT	BoLA-DQA 12021
.....A.....I..NU...T.....	.....T.....	.....R.....K...T.....	.....GA.....UG.R...V...S...AT	.....U..N...L.Q...S...AT	.....S...AT	BoLA-DQA 1401
.....A.....I..NU...T.....	.....T.....	.....K.....L...K...T.....	.....GA.....N..VN...QE...S...AT	.....U..N...L.Q...S...AT	.....S...AT	Bubu-DQA 0101
.....A.....I..NU...T.....	.....T.....	.....R.....K...T.....	.....GA.....UG.R...V...S...AT	.....U..N...L.Q...S...AT	.....S...AT	OLA-DQA1
.....U.S...TEI.Q.H...Q.TQ...	.....M...G.K...R.M.Q...G	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	Bofr-DQA 2001 (DQA2)
.....U.S...TEI.Q.H...Q.TQ...	.....M...G.K...R.M.Q...G	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	BoLA-DQA 2201
.....U.S...TEI.Q.H...Q.TQ...	.....M...G.K...R.M.Q...G	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	BoLA-DQA 22021
.....U.S...TEI.Q.H...Q.TQ...	.....L...G.K...R.M.GD.LT...	.....GA.SE...S..N..D.LT...P.A...	.....GA.SE...S..N..D.LT...P.A...	.....GA.SE...S..N..D.LT...P.A...	.....GA.SE...S..N..D.LT...P.A...	BoLA-DQA 2401
.....U...T.D...Q.H...Q...Q...	.....E..A...R..M.DKLR..H.GA...	.....I...N..DULTK.LY...P...	.....I...N..DULTK.LY...P...	.....I...N..DULTK.LY...P...	.....I...N..DULTK.LY...P...	BoLA-DQA 25012
.....U...T.D...Q.H...Q...Q...	.....E..A...R..M.DKLR..H.GA...	.....U...N..DULTK.LY...P...	.....U...N..DULTK.LY...P...	.....U...N..DULTK.LY...P...	.....U...N..DULTK.LY...P...	BoLA-DQA 2602
.....A.D...Q.H...Q...Q...L...	.....G.K...Q..M.G.LT...EA..A.NE.K...	.....DULTK...P...	.....DULTK...P...	.....DULTK...P...	.....DULTK...P...	BoLA-DQA 27012
.....U..I...I..I..Q...Q.TQ...	.....Q...K..A..Q..L.RML...LA...	.....IM.LH.UFLTKF...S...AT	.....IM.LH.UFLTKF...S...AT	.....IM.LH.UFLTKF...S...AT	.....IM.LH.UFLTKF...S...AT	BoLA-DQA 2801
.....U.S...TEI.Q.H...Q.TQ...	.....M...G.K...R.M.Q...G	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	.....AA.SE...S..N..DULTK...P...	Bubu-DQA 2001
.....F.S...TEI.Q.H...Q.TQ...	.....L...G.K...R.M.Q...G	.....GA.SE...QN.D.LTK...P.A...	.....GA.SE...QN.D.LTK...P.A...	.....GA.SE...QN.D.LTK...P.A...	.....GA.SE...QN.D.LTK...P.A...	OLA-DQA2
.....AA...I..NU...T.....	.....H...T.....	.....R.....K...U.G...GA...H..S.G...Q...QS...S...AT	.....R.....K...U.G...GA...H..S.G...Q...QS...S...AT	.....R.....K...U.G...GA...H..S.G...Q...QS...S...AT	.....R.....K...U.G...GA...H..S.G...Q...QS...S...AT	CLA-DQA

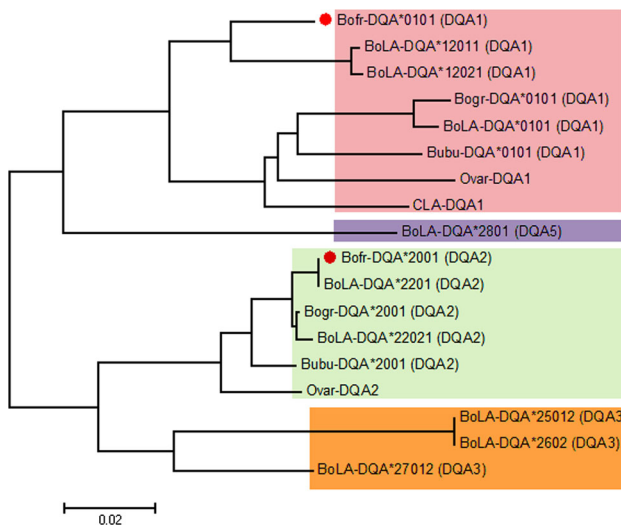
α2 domain

88	100	120	140	160	180	
KUPEUTQFSKSPUHLGQPNTLICHUDNI	FPPUINH	ILNRNGHSUTE	GESETSFLLKSDY	FLKINYLFLPSDDDUYDCKVEH	WGLDEPLLKHW	Bofr-DQA 0101 (DQA1)
.....M.....P.....	.....	.....L.I...I..D...S..D.H...A...S...	.....S..D.H...A...S...	.....S..D.H...A...S...	.....K.....	BoLA-DQA 0101
.....M.....P.....	.....	.....	.....I.....	.....I.....	.....I.....	BoLA-DQA 12011
.....M.....P.....	.....	.....S.....	.....I.....	.....S..D.H...S...S...	.....	BoLA-DQA 12021
.....M.....P.....	.....	.....S.....	.....I.....	.....S..D.H...S...S...	.....	Bubu-DQA 0101
.....M.....P.....	.....	.....I.....	.....A...S..D...S...S...	.....S..D.H...S...S...	.....	OLA-DQA1
.....E.....I..S.....	.....	.....K...A...P..D.H...G...N..I.....	.....P..D.H...G...N..I.....	.....P..D.H...G...N..I.....	.....	Bofr-DQA 2001 (DQA2)
.....E.....I..S.....	.....	.....K...A...P..D.H...G...N..I.....	.....P..D.H...G...N..I.....	.....P..D.H...G...N..I.....	.....	BoLA-DQA 2201
.....E.....I..S.....	.....	.....K...A...P..D.H...G...N..I.....	.....P..D.H...G...N..I.....	.....P..D.H...G...N..I.....	.....	BoLA-DQA 22021
.....E.....I..S.....	.....	.....K...A...P..D.H...G...N..I.....	.....P..D.H...G...N..I.....	.....P..D.H...G...N..I.....	.....	Bubu-DQA 2001
.....E.....I..S.....	.....	.....I...T...S..D.H...S...S...I.....E.....	.....S..D.H...S...S...	.....S..D.H...S...S...	.....	OLA-DQA2
.....E.....I..S.....	.....	.....I...T...S..D.H...S...S...I.....E.....	.....S..D.H...S...S...	.....S..D.H...S...S...	.....	CLA-DQA

CP domain TM domain CY domain

102	200	232		
EPDIPAPMSELTEI	TUVCALGLTVGLUGVIUGTULIT	QGLRSGGPRS	RHQGPL	Bofr-DQA 0101 (DQA1)
.....E.....	.....	.....R.....	.....	BoLA-DQA 0101
.....E.....	.....	.....M...I.....	.....	BoLA-DQA 12011
.....E.....	.....	.....M...I.....	.....	BoLA-DQA 12021
.....E.....	.....	.....R.....	.....	Bubu-DQA 0101
.....E.....	.....S.....M.....	.....R.....	.....	OLA-DQA1
.....EU.....	.....IF.....T.....	.....IF.....T.....	.....	Bofr-DQA 2001 (DQA2)
.....EU.....	.....IF.....T.....	.....IF.....T.....	.....	BoLA-DQA 2201
.....EU.....	.....IF.....A.....	.....IF.....A.....	.....	BoLA-DQA 22021
.....EU.....	.....IF.....A.....	.....IF.....A.....	.....	Bubu-DQA 2001
.....E.....	.....IF.....A.....	.....IF.....A.....	.....	OLA-DQA2
.....E.....	.....IF.....A.....	.....IF.....R.....	.....	CLA-DQA

**Figure 1.** An alignment between the amino acid sequences of Bofr-DQA and orthologous DQA sequences. The arrows indicate the amino acids positions consulting part of PBS. The putative N-linked glycosylation sites are underlined (—). The square (■) indicates the position of residues associated with binding of CD4+ molecules. A point (·) indicates 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). 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).



**Figure 2.** Phylogenetic tree based on DQA nucleotide sequences of gayal (neighbour-joining method).

adaptation to specific environment. Moreover, Indian buffaloes have extra three rare polymorphisms at positions 57 (hydrophilic > hydrophobic) and 36, 94 (hydrophobic > hydrophilic) resulting into the opposite water affinity (Nirajan *et al.* 2009). This may be from the animal germplasm because buffalo can well adapt to the tropical areas (Perera 2011). However, the replacements within the  $\alpha 1$  domain have impacted on the antigen-binding groove and could reveal differential binding ability to wide profiles of pathogens in different environments during the evolution process for livestock (Germain 1995; Williams *et al.* 2002). We conclude that the *Bofr-DQA1* and *-DQA2* genes are more identical with the corresponding sequences of their counterpart cattle. Similarly, the low-nucleotide sequence homology between *Bofr-DQA1* and *-DQA2* as well as the high proportion of nucleotide and amino acid substitutions clearly reveal inconsistency as allelic form. Our results also support the findings of Ballingall *et al.* (1998), that the bovine DQA3\*01 and DQA3\*02 sequences as nonallelic types have 92% nucleotide homology and larger genetic distance within two genes cluster.

From the phylogenetic tree exploration based on the nucleotide sequences, it appears that the split of the *DQA1* and *DQA2* sequences from the gayal and other ruminants into two major clades and further indicates their independently evolutionary relationship of the gayal DQA sequences (figure 2). We speculate from the results that gayal is genetically closer to cattle, which is in accordance with the previous studies (He *et al.* 2014; Sun *et al.* 2014). Further, a large distance between the two clades indicated that the *Bofr-DQA1* and *Bofr-DQA2* belong to two separate loci. It has been previously described that in case of cattle and buffalo, the DOA genes are present in duplicated form and both can be expressed (Russell *et al.* 1997; Nirajan *et al.* 2009) that seems to have similarity with gayal.

These duplicated genes with different mutations could be useful to promote immunological response as well as environmental adaptation for gayal.

There have been several nonsynonymous changes in the livestock species with respect to gayal that has altered the amino acid sequences in both *DQA1* and *DQA2* genes. A detailed results ( $d_n/d_s$  ratio) and synonymous/nonsynonymous substitution have been presented in electronic supplementary table (for  $d_n/d_s$  ratio see tables 1 & 2 in electronic supplementary material at <http://www.ias.ac.in/jgenet/> and for synonymous/nonsynonymous substitutions see tables 3 & 4 in electronic supplementary material).

In conclusion, the *Bofr-DQA1* and *Bofr-DQA2* genes have been characterized with extending our understanding to the MHC-DQA in rare ruminants. Like other animals *DQA* genes, the *Bofr-DQA* and *Bofr-DQA2* were also highly variable, especially in the  $\alpha 1$  domain as in most ruminants. It would be more interesting to clarify the effect of mutations from *Bofr-DQA1* and *Bofr-DQA2* on the pathogen's resistance for gayal adaption in further studies.

#### Acknowledgements

Authors acknowledge the support and cooperation received from the Dean, Faculty of Animal Science and Technology, Yunnan Agricultural University, China. We also thank the villagers and locals who helped us in sampling the gayal and the colleagues for synchronizing their work and help in wet lab experiments. Funding was provided by the National Nature Science Foundation of China under project nos. 31460583 and 31101640. Partial funding was also provided by the Yunnan Provincial Key Laboratory of Animal Nutrition and Feed with grant no. DYCX2015004.

#### References

- Andersson L. and Rask L. 1988 Characterization of the MHC class II region in cattle: the number of DQ genes varies between haplotypes. *Immunogenetics* **27**, 110–120.
- Ballingall K. T., Marasa B. S., Luyai A. and McKeever G. J. 1998 Identification of diverse BoLA DQA3 genes consistent with non-allelic sequences. *Anim. Genet.* **29**, 123–129.
- Bernatchez L. and Landry C. 2003 MHC studies in non model vertebrates: what have we learned about natural selection in 15 years? *J. Evol. Biol.* **16**, 363–377.
- Brown J. H., Jardetzky T. S., Gorge J. C., Stern L. J., Urban R. G., Strominger J. L. *et al.* 1993 Three dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**, 33–39.
- Chi J., Fu B., Nie W., Wang J., Graphodatsky A. S. and Yang F. 2005 New insights into the karyotypic relationships of Chinese muntjac (*Muntiacus reevesi*), forest muskdeer (*Moschus berezovskii*) and gayal (*Bos frontalis*). *Cytogenet. Genome Res.* **108**, 310–316.
- Germain R. N. 1995 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.

- Glass E. J., Oliver R. A. and Russell G. C. 2009 Duplicated DQ haplotypes increases the complexity of restriction element usage in cattle. *J. Immunol.* **165**, 134–138.
- Gou X., Wang Y. Q., Yang S. L., Deng W. D. and Mao H. M. 2010 Genetic diversity and origin of gayal and cattle in Yunnan revealed by mtDNA D-loop and SRY gene sequence variation. *J. Anim. Breed. Genet.* **127**, 154–160.
- He Y., Xi D., Leng J., Qian T., Jin D., Chen T. et al. 2014 Genetic variability of MHC class II DQB exon 2 alleles in yak (*Bos grunniens*). *Mol. Biol. Rep.* **41**, 2199–2206.
- Kappes D. and Strominger J. L. 1988 Human class II major histocompatibility complex genes and proteins. *Annu. Rev. Biochem.* **57**, 991–1028.
- Klein J. 1986 Seeds of time: fifty years ago Peter A. Gorer discovered the H-2 complex. *Immunogenetics* **24**, 331–338.
- Kuduk K., Babik W., Bojarska K., Sliwinska E. B., Kingberg J., Taberlet P. et al. 2012 Evolution of major histocompatibility complex class I and class II genes in the brown bear. *BMC Evol. Biol.* **12**, 197.
- Lan H., Xiong X., Lin S., Liu A. and Shi L. 1993 Mitochondrial DNA polymorphism of cattle (*Bos taurus*) and mithun (*Bos frontalis*) in Yunnan Province. *Acta Gen. Sin.* **20**, 419–425.
- Lazzaro B. P. and Little T. J. 2009 Immunity in a variable world. *Philos. Trans. R. Soc. London, Ser. B* **364**, 15–26.
- McKinney D. M., Southwood S., Hinz D., Osseroff C., Arlehamn C. S., Schulten V. et al. 2013 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.
- Nie L., Shi L., He X. D., Zhao Y. L., Mu W. G. and Zhang J. L. 1995 The gayal's genetic diversity and its genetic structure's enzyme analysis. *Acta Genet. Sin.* **22**, 185–191.
- Niranjan S. K., Deb S. M., Sharma A. and Kumar S. 2009 Isolation of two cDNAs encoding MHC-DQA1 and -DQA2 from the water buffalo, *Bubalus bubalis*. *Vet. Immunol. Immunopathol.* **130**, 268–271.
- Perera B. M. 2011 Reproductive cycles of buffalo. *Anim. Reprod. Sci.* **124**, 194–199.
- Rajkhowa S., Rajkhowa C., Rahman H. and Bujarbaruah K. M. 2004 Seroprevalence of infectious bovine rhinotracheitis in mithun (*Bos frontalis*) in India. *Rev. Sci. Tech.* **23**, 821–829.
- Reed D. H. and Frankham R. 2003 Correlation between fitness and genetic diversity. *Conserv. Biol.* **17**, 230–237.
- Rudd P. M., Wormald M. R., Stanfield R. L., Huang M., Mattsson N., Speir J. A. et al. 1999 Roles for glycosylation of cell surface receptors involved in cellular immune recognition. *J. Mol. Biol.* **293**, 351–366.
- Russell B. C., Gallagher A., Craigmile S. and Glass E. J. 1997 Characterization of cattle cDNA sequences from two DQA loci. *Immunogenetics* **45**, 455–458.
- Sena L., Schneider M. P., Brenig B. B., Honeycutt R. L., Honeycutt D. A., Womack J. E. et al. 2011 Polymorphism and gene organization of water buffalo MHC-DQB genes show homology to the BoLA DQB region. *Anim. Genet.* **42**, 378–385.
- Sigurdardottir S., Borsch C., Gustafsson K. and Andersson L. 1992 Gene duplications and sequence polymorphism of bovine class II DQB genes. *Immunogenetics* **35**, 205–213.
- Simoons F. J. 1984 Gayal or mithan. In *Evolution of domesticated animals* (ed. I. L. Mason), pp. 34–38. Longman, London.
- Sun Y., Xi D., Li G., Hao T., Chen Y. and Yang Y. 2014 Genetic characterization of MHC class II DQB exon 2 variants in gayal (*Bos frontalis*). *Biotechnol. Biotech. Eq.* **28**, 827–833.
- Tamura K., Dudley J., Nei M. and Kumar S. 2007 MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596–1599.
- Tibayrenc M. 2004 A molecular biology approach to tuberculosis. *Proc. Natl. Acad. Sci. USA* **101**, 4721–4722.
- Trowsdale J. 2001 Genetic and functional relationships between MHC and NK receptor genes. *Immunity* **15**, 363–374.
- Trowsdale J. and Parham P. 2004 Defense strategies and immunity-related genes. *Eur. J. Immunol.* **34**, 7–17.
- Verkaar E. L. C., Nijman I. J., Beeke M., Hanekamp E. and Lenstra J. A. 2004 Maternal and paternal lineages in cross-breeding bovine species. Has Wisent a hybrid origin? *Mol. Biol. Evol.* **21**, 1165–1170.
- Walker E., Warnick F. and Hamlet S. 1968 **Mammals of the world**. The Johns Hopkins Press, Baltimore.
- Williams A., Peh C. A. and Elliott T. 2002 The cell biology of MHC class I antigen presentation. *Tissue Antigens* **59**, 3–17.
- Xi D. M., Wanapat M., Deng W. D., He T. B., Yang Z. F. and Mao H. M. 2007 Comparison of gayal (*Bos frontalis*) and Yunnan Yellow Cattle (*Bos taurus*): in vitro dry matter digestibility and gas production for a range of forages. *Asian-Aust. J. Anim. Sci.* **20**, 1208–1214.
- Yang Z. 2007 PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591.
- Zhao K. D., Ou C. H., Huang Y. L. and He T. B. 2003 Rare animal germplasm resources in Yunnan Province: present situation and counter measures of preservation and research on Dulong cattle (*Bos frontalis*). *J. Yellow Cattle Sci.* **29**, 71–74 (in Chinese with English abstract).

Corresponding editor: INDRAJITH NANDA