Immunogenetics (2019) 71:347–351
https://doi.org/10.1007/s00251-018-1096-9

SHORT COMMUNICATION

Allelic nomenclature for the duplicated MHC class II *DQ* genes in sheep

Keith T. Ballingall[1] · Kara Dicks[2] · Panoraia Kyriazopoulou[1,2] · Lynne Herrmann-Hoesing[3]

Received: 16 October 2018 / Accepted: 6 November 2018 / Published online: 10 November 2018

Abstract

The principal MHC class II molecules involved in the presentation of peptides to the antigen specific receptors on $CD4^+$ T cells genes in sheep are derived from *DR* and *DQ* genes. Allelic nomenclature systems for the *DRB1* and its partner *DRA* loci are available for Ovid's; however, no official nomenclature is available for the *DQ* genes which creates ambiguity within the research community. Ovine MHC haplotypes include at least two pairs of *DQA* and *DQB* genes, termed *DQA1*, *DQB1* and *DQA2*, *DQB2* and both sets are polymorphic and both seem to be functional. In a number of haplotypes, the *DQA1* locus appears to be absent (*DQA1*-null) and is replaced by a second locus termed *DQA2-like*. Here, we identify families of alleles based on sequence similarity and phylogenetic clustering which correspond to each of the *DQA* and *DQB* genes identified in previous genomic and transcript analyses of homozygous animals. Using such criteria to cluster sequences, we have named 82 full-length and partial cDNA transcripts derived from domestic sheep (*Ovis aries*) which correspond to alleles at the *Ovar-DQA1*, *DQA2*, *DQA2-like*, *DQB1*, *DQB2* and *DQB2-like* genes and provide associated sequence resources available to the research community through the IPD-MHC Database. This sets the basis for naming and annotation of *DQ* genes within the ovine MHC and may be used as a template for *DQ* genes in other ruminant species which will ultimately support research in livestock infectious disease.

Keywords Sheep · MHC class II · *DQ* · Nomenclature · IPD-MHC

The major histocompatibility complex (MHC) is one of the most immunologically significant and genetically diverse regions of the vertebrate genome (Parham and Ohta 1996; Trowsdale 2011). An understanding of the nature, function and maintenance of allelic diversity at MHC loci in the major livestock species is increasingly important for the development of new and improved vaccines for global food security. The comparative MHC Nomenclature Committee provides guidance in the development of standardised nomenclature systems for alleles at MHC loci in non-human species (Ballingall et al. 2018a). The Ovine Nomenclature Committee has facilitated the development of a nomenclature system for alleles at the highly polymorphic MHC class II *DRB1* locus (Ballingall et al. 2011) with over 100 alleles having received official names. An associated database of these sequences is maintained in the ovine section of the IPD-MHC Database (https://www.ebi.ac.uk/ipd/mhc/). However, progress in developing an official nomenclature system for the *DQ* loci in sheep has been slower due to complexities associated with gene duplication, haplotype variation and limited sequence availability. As such, no officially named alleles are currently available for the ovine *DQ* genes which creates ambiguity within the research community when describing alleles at each of these loci.

The *DQ* repertoire in sheep is complicated by a functional duplication of the *DQA* and *DQB* genes (Scott et al. 1987, 1991a, b). Haplotypes analysed to date have two *DQA* genes with the majority including *DQA1* and *DQA2* loci (Scott et al. 1991a; Hickford et al. 2007). However, in a number of haplotypes, the *DQA1* gene appears absent (*DQA1* null, Scott et al. 1991a, Fabb et al. 1993). In these haplotypes, the *DQA2* locus is found in combination with a second locus which phylogenetic analyses of the second exon suggests is more closely related to *DQA2*, (Hickford et al. 2004) than to *DQA1*. Hence, the name is *DQA2-like*. As in other mammals, genomic analysis has indicated that the *DQ* genes in sheep occur as closely linked *A/B* gene pairs with *DQA1* linked to *DQB1* and

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00251-018-1096-9) contains supplementary material, which is available to authorized users.

✉ Keith T. Ballingall
keith.ballingall@moredun.ac.uk

[1] Moredun Research Institute, Midlothian, UK

[2] The University of Edinburgh, Edinburgh, UK

[3] Washington State University, Pullman, USA

DQA2 linked with *DQB2* (Wright and Ballingall 1994; Herrmann-Hoesing et al. 2008). No similar genomic analysis of haplotypes including *DQA2-like* sequences has been reported; however, transcriptional analysis of MHC homozygous animals identified unique transcripts which co-express with *DQA2-like* genes. It is likely that these represent the partner of *DQA2-like* and have therefore been termed *DQB2-like* (Ballingall et al. 2018b).

The ISAG/IUIS-VIC OLA Nomenclature Committee now considers that sufficient sequence data is available to propose a nomenclature framework for alleles at each of the *DQ* genes in sheep. We propose a nomenclature system for alleles at the Ovar-*DQA1*, *DQA2*, *DQA2-like*, *DQB1*, *DQB2* and *DQB2-like* genes. This will be developed using 82 full-length and partial *DQA* and *DQB* transcripts and will take an approach similar to that used to assign cattle MHC class I sequences to individual loci (Hammond et al. 2012).

The majority of the full-length *DQA* and *DQB* transcripts used to develop the nomenclature have previously been described (Ballingall et al. 2015, 2018b). Additional full-length transcripts are derived from MHC haplotype analysis within the Soay (Kara Dicks, personal communication) and Argos breeds (Panoraia Kyriazopoulou, personal communication). An additional 31 previously unpublished *DQB* transcripts covering 575 bp between exon 2 and exon 4 were included in the development of the *DQB* nomenclature. These sequences were derived from cloned PCR products, amplified from cDNA prepared from RNA extracted from PBMC from a Rambouillet flock using the primers and amplification conditions described by Herrmann-Hoesing et al. (2008). Each *DQ* nucleotide sequence has been submitted to the IPD-MHC Database (https://www.ebi.ac.uk/ipd/mhc) to allow construction of the allelic databases. The origin and identity of each *DQA* and *DQB* sequence is described in Tables 1 and 2 respectively. The recently updated tools available on the IPD-MHC Database (Maccari et al. 2017) allow alignment of nucleic acid or amino acid sequences from individual or multiple loci within and across species. Alignments of all *DQA* and *DQB* nucleotide sequences are provided in supplementary Figures 1 and 2 respectively or may be generated using the tools available on the IPD-MHC Database.

In our proposed nomenclature, gene specificity is based on nucleotide similarity to full-length reference sequences derived from MHC homozygous animals which in turn are based on sequence similarity to the initial descriptions of linked *DQA1/ B1* and *DQA2/ B2* genes on cosmid and BAC clones (Wright and Ballingall 1994; Herrmann-Hoesing et al. 2008). Sequence similarity was determined by a BLAST search of an in-house sequence database. Specificity was confirmed by constructing maximum likelihood phylogenetic trees in IQ-TREE (Trifinopoulos et al. 2016) using multiple alignments generated using CLUSTAL Omega (Sievers et al. 2011). The model selection tool (Kalyaanamoorthy et al. 2017) within IQ-tree was used to select the optimum substitution models, prior to phylogenetic tree estimation. The optimum substitution models selected for the *DQA* and *DQB* sequences were the Kimura 2 parameter (K2P, Kimura 1980) +R3 and K2P +G4 respectively. Tree topology was tested with 10,000 bootstrap replicates using the ultrafast boost strap method of Minh et al. (2013). The *DQA* and *DQB* trees are shown in Figs. 1 and 2 respectively. The clustering of sequences in Figs. 1 and 2 is consistent with alleles at each of the *DQ* genes in sheep.

Following sequence comparison and phylogenetic analysis, the nomenclature system was developed for each cluster of sequences, in accordance with the guidance provided in the recent report from the Comparative MHC Nomenclature

Table 1 List of *DQA* sequences with the proposed allelic nomenclature

Official name	Breed	Accession number	Reference
*DQA1*01:01:01*	SBF	HG798783	1
*DQA1*02:01:01*	SBF	HG798784	1
*DQA1*02:01:02*	SBF	HG798785	1
*DQA1*03:01:01*	PA	HG798786	1
*DQA1*03:01:02*	RB	EU176819	2
*DQA1*03:02:01*	Soay	LR025209	5
*DQA1*04:01:01*	AM	M93430	3
*DQA1*04:02:01*	Soay	LR025208	5
*DQA1*05:01:01*	PA	HG798787	1
*DQA1*05:02:01*	CM	FJ985876	4
*DQA1*06:01:01*	PA	HG798788	1
*DQA1*07:01:01*	Argos	LS990820	6
*DQA2*01:01:01*	SBF/RB	HG798789/EU176819	1/2
*DQA2*01:02:01*	SBF	HG798790	1
*DQA2*01:03:01*	AM	M93431	3
*DQA2*02:01:01*	SBF	HG798791	1
*DQA2*03:01:01*	PA	HG798792	1
*DQA2*04:01:01*	AM	M93433	3
*DQA2*04:02:01*	Soay	LR025212	5
*DQA2*05:01:01*	SBF	HG798793	1
*DQA2*06:01:01*	PA	HG798794	1
*DQA2*07:01:01*	PA	HG798795	1
*DQA2*08:01:01*	WM	HG798796	1
*DQA2*09:01:01*	Argos	LS990821	6
*DQA2*09:01:02*	Soay	LR025211	5
*DQA2*10:01:01*	Soay	LR025213	5
*DQA2-like*01:01:01*	SBF	HG798797	1
*DQA2-like*02:01:01*	WM	HG798798	1
*DQA2-like*03:01:01*	Soay	LR025210	5

Ballingall et al. 2015; Herrmann-Hoesing et al. 2008; Fabb et al. 1993; Gao et al. 2010; Kara Dicks, personal communication; Panoraia Kyriazopoulou, personal communication. *RB* Rambouillet, *AM* Australian Merino, *CM* Chinese Merino, *SBF* Scottish Blackface, *PA* French Prealpe, *WM* Welsh Mountain

Table 2 List of *DQB* sequences with the proposed allelic nomenclature

Official name	Breed	Accession number	Reference
*DQB1*01:01:01*	SBF/RB	LT837701/HQ728688	1/here
*DQB1*02:01:01*	RB	EU176819	2
*DQB1*03:01:01*	SBF	LT837702	1
*DQB1*03:02:01*	RB	HQ728684	Here
*DQB1*04:01:01*	SBF/RB	LT837703/HQ728667	1/here
*DQB1*05:01:01*	CM	FJ985876	3
*DQB1*05:01:02*	SBF	LT837704	1
*DQB1*05:02:01*	SBF	LT837705	1
*DQB1*05:03:01*	RB	HQ728670	Here
*DQB1*05:04:01*	RB	HQ728687	Here
*DQB1*06:01:01*	SBF	LT837706	1
*DQB1*07:01:01*	SBF	LT837707	1
*DQB1*07:02:01*	SBF/RB	LT837708/HQ728681	1/here
*DQB1*07:03:01*	RB	HQ728685	Here
*DQB1*08:01:01*	M	XM_012173129	7
*DQB1*08:02:01*	RB	HQ728678	Here
*DQB1*09:01:01*	Argos/RB	LS990822/HQ728696	6/here
*DQB1*10:01:01*	Argos	LS990823	6
*DQB1*11:01:01*	RB	HQ728668	Here
*DQB1*12:01:01*	RB	HQ728683	Here
*DQB1*13:01:01*	RB	HQ728690	Here
*DQB1*14:01:01*	RB	HQ728692	Here
*DQB2*01:01:01*	SBF	LT837709	1
*DQB2*02:01:01*	SBF/RB	LT837710/HQ728672	1/here
*DQB2*03:01:01*	SBF	LT837711	1
*DQB2*04:01:01*	RB	EU176819	2
*DQB2*05:01:01*	SBF	LT837712	1
*DQB2*06:01:01*	SBF	LT837714	1
*DQB2*07:01:01*	AM	L08792	4
*DQB2*07:02:01*	RB	HQ728675	Here
*DQB2*08:01:01*	SBF	LT837715	1
*DQB2*08:02:01*	SBF	LT837716	1
*DQB2*08:03:01*	RB	HQ728686	Here
*DQB2*09:01:01*	Soay/RB	LR025203/HQ728694	5/here
*DQB2*10:01:01*	Soay	LR025204	5
*DQB2*11:01:01*	Soay	LR025206	5
*DQB2*12:01:01*	Soay/RB	LR025205/HQ728680	5/here
*DQB2*13:01:01*	Argos	LS990824	6
*DQB2*14:01:01*	Argos	LS990825	6
*DQB2*14:02:01*	RB	HQ728695	Here
*DQB2*15:01:01*	RB	HQ728669	Here
*DQB2*16:01:01*	RB	HQ728671	Here
*DQB2*17:01:01*	RB	HQ728676	Here
*DQB2*18:01:01*	RB	HQ728677	Here
*DQB2*19:01:01*	RB	HQ728679	Here
*DQB2*20:01:01*	RB	HQ728682	Here
*DQB2*21:01:01*	RB	HQ728689	Here
*DQB2*22:01:01*	RB	HQ728691	Here
*DQB2*23:01:01*	RB	HQ728693	Here
*DQB2-like*01:01:01*	SBF/RB	LT837717/HQ728673	1/here
*DQB2-like*02:01:01*	SBF	LT837718	1
*DQB2-like*02:02:01*	M	XM_012159546	7
*DQB2-like*03:01:01*	Soay/RB	LR025207/HQ728697	5/here

Ballingall et al. 2018b; Herrmann-Hoesing et al. 2008; Gao et al. 2010; Fabb et al. 1993; Kara Dicks, personal communication; Panoraia Kyriazopoulou, personal communication; unpublished sequences from the Mouflon genome project. *RB* Rambouillet, *AM* Australian Merino, *CM* Chinese Merino, *SBF* Scottish Blackface, *PA* French Prealpe, *WM* Welsh Mountain, *M* Mouflon

Committee (Ballingall et al. 2018a). A full-length transcript was selected as the reference sequences for each cluster. For consistency, a single well-defined haplotype, 501b, was selected to provide the reference sequences for *DQA1*, *DQA2* and *DQB1*, *DQB2* genes (Ballingall et al. 2018b). The 501a haplotype was selected to provide the reference sequences for *DQA2-like* and *DQB2-like* (Ballingall et al. 2018b).

To maintain consistency with HLA nomenclature (Marsh et al. 2010), and to follow the guidelines of the MHC Nomenclature Committee (Maccari et al. 2018), the following system is proposed for alleles at each locus. The first number or field following the species and locus designation (*Ovar-DQA1*, *Ovar-DQB1*) and separated by an asterisk (*) represents the allelic group (*Ovar-DQA1*01*, **02* etc.) where alleles within a group differ by no more than four amino acids within the alpha 1 or beta 1 domain (encoded by the second exon) and no more than four amino acids predicted throughout the remainder of the transcript. The next field separated by a colon (:) as a field separator indicates coding (non-synonymous) change within the allelic group (*Ovar-DQA1*01:01, Ovar-DQB1*01:01*) and is manually assigned by the ovine group curator in order of submission to the IPD-MHC Database. The following field, again separated by a colon (*Ovar-DQA1*01:01:02, Ovar-DQB1*01:01:02*), may be used to indicate silent or synonymous substitutions. The flexibility of the system allows for additional fields to reflect diversity within intronic and regulatory regions.

Requests for official names for *DQ* sequences will be delivered through the IPD-MHC Database provided that sequence quality guidelines are followed (Ballingall et al. 2018a). Sequences should be submitted to the IPD-MHC Database using the online tools. Ideally, full-length transcripts should be submitted as these simplify gene assignment. The immediate three-prime untranslated regions which appear to contain gene specific motifs (Ballingall et al. 2015, 2018b) are especially helpful in gene assignment (see Supplementary Figures 1 and 2). Gene assignment and the subsequent allelic nomenclature will be based on the closest match to other DQ sequences held in the database. In cases where it is not clear from which a gene sequence is derived, the curator may request additional data.

A number of additional gene duplications within the ovine *DQ* region have been recently been reported (Ali et al. 2017, Ballingall et al. 2018b, Kara Dicks, personal communication). It is not yet clear if these represent functional duplications or gene fragments that co-amplify with the primers used. For the purpose of nomenclature, such sequences will be named depending on their clustering with other *DQ* sequences.

In summary, here we provide nomenclature for 82 alleles at the duplicated *DQ* genes within the MHC class II region of sheep. Associated databases of alleles are available within the IPD-MHC Database for alignment and download. Such a resource will support research in ruminant immunology, vaccine development, comparative studies of MHC evolution and population-based analyses of MHC diversity and disease. This is the first allelic nomenclature system proposed for the

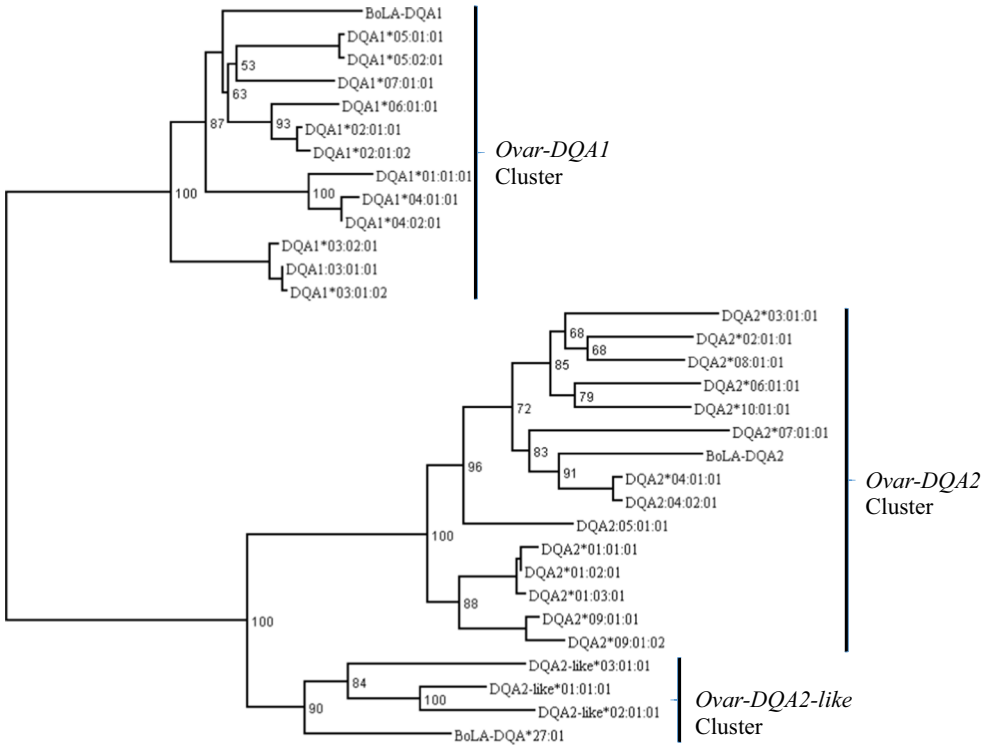

Fig. 1 Maximum likelihood tree estimating the relationships between ovine *DQA* sequences

DQ genes in a ruminant species which assigns sequences to individual loci based on sequence and phylogenetic analysis. It may also provide a framework for the *DQ* genes in other farmed ruminant species including cattle and goat.

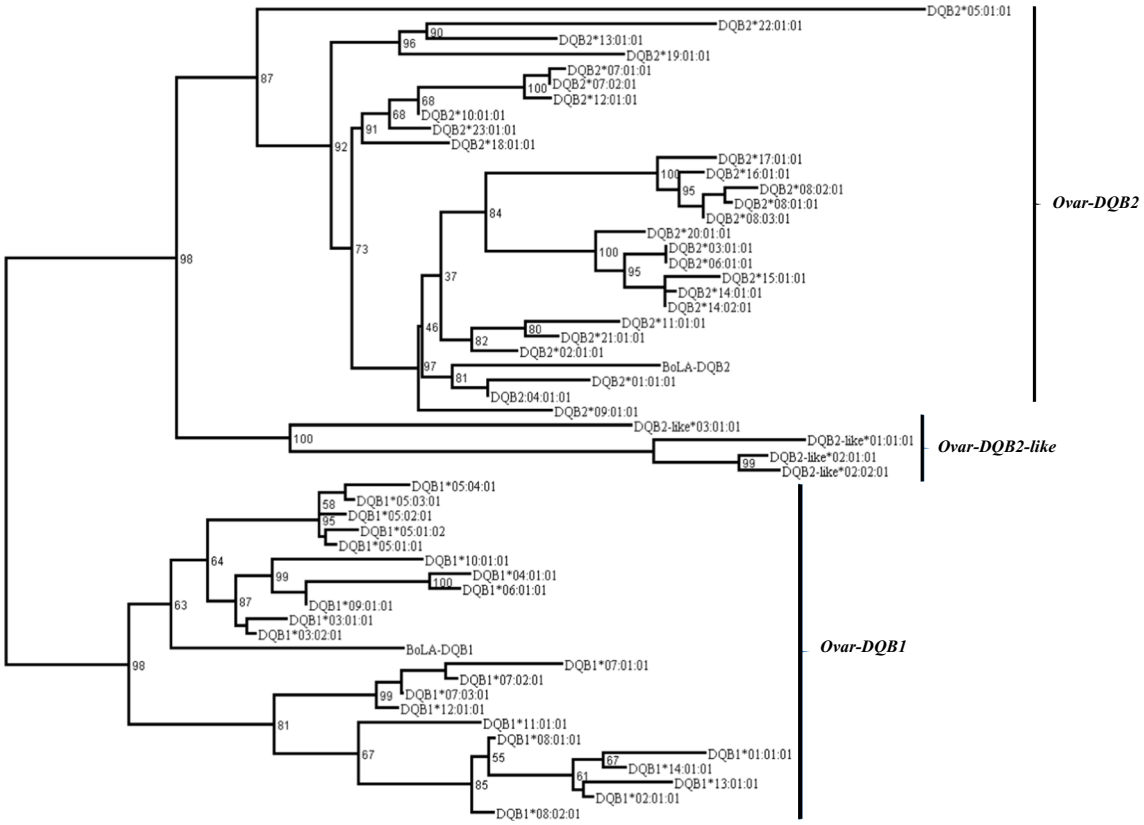

Fig. 2 Maximum likelihood tree estimating the relationships between ovine *DQB* sequences

Funding information The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7, 2007-2013), Research Infrastructures action, under the grant agreement No. FP7-228394 (NADIR) and European Union's Horizon 2020 research and innovation programme under grant agreement no. 731014 (VetBioNet). KB receives support from the Scottish Government's strategic research programme. The IPD-MHC project receives funding from the Biotechnology and Biological Sciences Research Council (BBSRC) under grant (BB/M011488/1). Thesupport provided by Giuseppe Maccari in processing the sequence data into the IPD-MHC Database is gratefully acknowledged. The work of the IPD-MHC Databases is supported by the European Molecular Biology Laboratories, the International Society for Animal Genetics. (ISAG) and the Veterinary Immunology Committee (VIC) of the International Union of Immunological societies (IUIS).

References

Ali AO, Stear A, Fairlie-Clarke K, Brujeni GN, Isa NM, Salisi MS, Donskow-Łysoniewska K, Groth D, Buitkamp J, Stear MJ (2017) The genetic architecture of the MHC class II region in British Texel sheep. Immunogenetics 69:157–163

Ballingall KT, Herrmann-Hoesing L, Robinson J, Marsh SGE, Stear MJ (2011) A single nomenclature and associated database for alleles at the MHC class II DRB1 locus of sheep. Tissue Antigens 77:546–553

Ballingall KT, Steele P, Lantier I, Cotelli M, Todd H, Lopez G, Martin E, Lantier F (2015) An ancient inter-locus recombination increases class II MHC *DQA* diversity in sheep and other bovidae. Anim Genet 46:333–336

Ballingall KT, Bontrop RE, Ellis SA, Grimholt U, Hammond JA, Ho C-S, Kaufman J, Kennedy LJ, Maccari G, Miller D, Robinson J, Marsh SGE (2018a) Comparative MHC nomenclature: report from the ISAG/IUIS-VIC committee 2018. Immunogenetics 70:625–632. https://doi.org/10.1007/s00251-018-1073-3

Ballingall KT, Lantier I, Todd H, Lantier F, Rocchi M (2018b) Structural and functional diversity arising from intra- and inter-haplotype combinations of duplicated DQA and B loci within the ovine MHC. Immunogenetics 70:257–269

Fabb SA, Maddox JF, Gogolin-Ewens KJ, Baker L, Wu MJ, Brandon MR (1993) Isolation, characterization and evolution of ovine major histocompatibility complex class II DRA and DQA genes. Anim Genet 24:249–255

Gao J, Liu K, Liu H, Blair HT, Li G, Chen C, Tan P, Ma RZ (2010) A complete DNA sequence map of the ovine major histocompatibility complex. BMC Genomics 11:466. https://doi.org/10.1186/1471-2164-11-466

Hammond JA, Marsh SG, Robinson J, Davies CJ, Stear MJ, Ellis SA (2012) Cattle MHC nomenclature: is it possible to assign sequences to discrete class I genes? Immunogenetics 64:475–480

Herrmann-Hoesing LM, White SN, Kappmeyer LS, Herndon DR Knowles DP (2008) Genomic analysis of Ovis aries (Ovar) MHC class IIa loci. Immunogenetics 60:167–176

Hickford JGH, Zhou H, Slow S, Fang Q (2004) Diversity of the ovine *DQA2* gene. J Anim Sci 82:1553–1563

Hickford JGH, Zhou H, Fang Q (2007) Haplotype analysis of the DQA genes in sheep: evidence supporting recombination between the loci. J Anim Sci 85:577–582

Kalyaanamoorthy S, Minh BQ, Wong TFK, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Maccari G, Robinson J, Ballingall K, Guethlein L, Grimholt U, Kaufmann J, Ho C-K, de Groot N, Flicek P, Bontrop R, Hammond J, Marsh S (2017) IPD-MHC 2.0: an improved inter-species database for the study of the major histocompatibility complex. Nucleic Acids Res 45(Database issue):D860–D864. https://doi.org/10.1093/nar/gkw1050

Maccari G, Robinson J, Bontrop RE, Otting N, de Groot NG, Ho C-S, Ballingall KT, Marsh SGE, Hammond JA (2018) IPD-MHC: nomenclature requirements for the non-human major histocompatibility complex in the next-generation sequencing era. Immunogenetics 70:619–623. https://doi.org/10.1007/s00251-018-1072-4

Marsh SGE, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA, Fernandez-Vina M, Geraghty DE, Holdsworth R, Hurley CK et al (2010) Nomenclature for factors of the HLA system, 2010. Tissue Antigens 75:291–455

Minh BQ, Nguyen MAT, von Haeseler A (2013) Ultrafast approximation for phylogenetic bootstrap. Mol Biol Evol 30:1188–1195

Parham P, Ohta T (1996) Population biology of antigen presentation by MHC class I molecules. Science 272:67–74

Scott PC, Choi C-L, Brandon MR (1987) Genetic organization of the ovine MHC class II region. Immunogenetics 25:116–122

Scott PC, Gogolin-Ewens KJ, Adams TE, Brandon MR (1991a) Nucleotide sequence, polymorphism and evolution of ovine MHC class II *DQA* genes. Immunogenetics 34:69–79

Scott PC, Maddox JF, Gogolin-Ewens KJ, Brandon MR (1991b) The nucleotide sequence and evolution of ovine MHC class II B genes: *DQB* and *DRB*. Immunogenetics 34:80–87

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, Mcwilliam H, Remmert M, Sö Ding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539

Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res 44:W232–W235

Trowsdale J (2011) The MHC, disease and selection. Immunol Lett 137: 1–8

Wright H, Ballingall KT (1994) Mapping and characterization of the *DQ* subregion of the ovine MHC. Anim Genet 25:243–249