Evolution of the ovine MHC *DQA* region

J G H Hickford, H J Ridgway, A P Escayg

Summary

Southern hybridisation was used to define an apparent gene duplication event at the ovine *DQA2* locus. Approximately 500 sheep from five different breeds were genotyped at their *DQA1* and *DQA2* loci. A subset of these were selected for further characterisation. Southern hybridisation of *Taq*I digested DNA revealed no *DQA1* region in some sheep. It was also noted in these *DQA1* null animals the *DQA2* specific probe hybridised to two bands. An *Eco*RV-RFLP designed to distinguish copy number confirmed this duplication of the *DQA2* region. The results showed that the duplication was exclusively associated with the *DQA1* null haplotype and occured only in alleles *DQA2-F, -G, -I* and *-J*. Comparison with bovine MHC genes revealed that they also contained a *DQA1* null haplotype and that this haplotype was associated with a putative *DQA3* gene. The potential for an ovine *DQA3* locus is discussed.

Keywords: Mhc*Ovar*, *DQA1*, *DQA2*, null allele, gene duplication

Introduction

The polymorphism of MHC class I and II molecules has been a long-standing puzzle, raising questions as to the significance of this diversity, its genetic and structural basis, the mechanism by which it is produced, and the rate at which it evolves. Numerous mechanisms for the generation of this polymorphism have been suggested and no consensus has been reached in more than two decades.

Extensive MHC class II polymorphism and the existence of gene families in humans has led to the postulate that gene duplication (Rask *et al.* 1985; Tokunaga *et al.* 1988; Rask *et al.* 1990; Kuhner & Peterson 1992; Horne *et al.* 1993), has played an important role during the evolution of the various *HLA* class II loci. Given that all vertebrates have several MHC loci and that there are clusters of genes which share

homology, it is likely that gene duplication is a common driving force in MHC evolution.

There has been considerable interest in the genetic organisation of the MHC region of production animals because of its association with resistance to a wide range of diseases (Weigel *et al.* 1990; Xu *et al.* 1993; Buitkamp *et al.* 1996; Escayg *et al.* 1997; Stear & Wakelin 1998). In sheep, the *DQA* genes are found within a 130-kb region on chromosome 20 and both *DQA1* and *DQA2* are transcribed (Scott *et al.* 1991; Wright & Ballingall 1994). A common ancestor gene probably duplicated to give rise to the *DQA1* and *DQA2* genes (Scott *et al.* 1991).

We describe an investigation of the *DQA1* and *DQA2* loci of the ovine MHC (Mhc*Ovar*) class II region and present evidence of a gene duplication event that appears to precede divergence of the artiodactyl lineages.

Materials and methods

Animals

The animals used in this study have been detailed previously (Escayg *et al.* 1996). Briefly, they comprise nearly 500 animals derived from the Coopworth, Perendale, Romney, Corriedale and Merino breeds, most being members of full or half-sib families. Following *DQA1* and *DQA2* genotyping, a representative group of the sheep was selected for polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) analysis and further Southern hybridisation.

DNA preparation

DNA was extracted from either blood (Montgomery & Sise 1990) or skin tissue ($\approx$ 35 mg), by digestion overnight at 37 °C with 0·7 ml digestion buffer (150 mM NaCl, 20 mM Tris–HCl (pH 8·0), 40 mM EDTA (pH 8·0), 1% SDS, 0·5 mg/ml proteinase K). Samples were extracted with 0·8 ml phenol/chloroform/isoamyl alcohol (25 : 24 : 1), the DNA was precipitated, washed twice with ethanol, pelleted, dried briefly and dissolved in 0·2 ml TE (pH 8·0) containing pancreatic RNase (5·0 μg/ml). DNA isolated from the same individual by both procedures generated identical banding patterns following Southern hybridisation (data not shown).

J G H Hickford
H J Ridgway
Animal and Food Sciences Division, PO Box 84, Lincoln University, Canterbury, New Zealand

A P Escayg
Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, United States of America

Correspondence: J G H Hickford.

Southern hybridisation analysis and naming of alleles

Sheep DNA (5–10 μg) was digested for analysis with various restriction endonucleases using the manufacturers' recommended procedures (New England Biolabs, Beverly, MA; Boehringer Mannheim, Mannheim, Germany). Digested DNA was electrophoresed for 14–18 h in 0·8% agarose gels and after depurination transferred to Hybond™-N + membrane (Amersham Life Sciences, Little Chalfont, Buckinghamshire, UK) by the alkaline capillary method (Sambrook *et al.* 1989). Prehybridisation was performed at 65 °C for 2 h in buffer comprising 4× SSPE, 5× Denhardt's solution and 0·5% sodium dodecylsulphate (SDS). Hybridisation was performed at 65 °C for 14–18 h following the addition of herring sperm DNA (0·5 mg) and radioactively labelled probe. DNA probes were radioactively labelled with α-^{32}P-dCTP to a specific activity of 1×10^9 cpm/μg using the Megaprime DNA Labelling System (Amersham). After hybridisation, the membranes were rinsed in 2× SSC, 0·1% SDS, then washed successively for 15 and 30 min in 2× SSC, 0·1% SDS with a final wash at 65 °C comprising 0·3× SSC, 0·5% SDS for 10 min. Membranes were exposed to Kodak X-Omat AR film with intensifying screens at − 80 °C for 3–14 days. Autoradiographs were analysed and alleles assigned based on the calculated molecular weight of bands, as described previously (Escayg *et al.* 1996).

Membrane stripping

Bound radioactivity was removed by pouring boiling 0·25% SDS on the membranes and agitating for 20 min. The solution was discarded and the process repeated with 0·5% SDS. Once stripped, membranes were stored at 4 °C.

Probes

The *DQA1* and *DQA2* exon 2 probes used for Southern hybridisation are as described by Escayg *et al.* (1996) and are detailed in Fig. 1a,b. A longer 1230 bp *DQA1* probe covering exon 2, intron 2, exon 3, intron 3 and a small
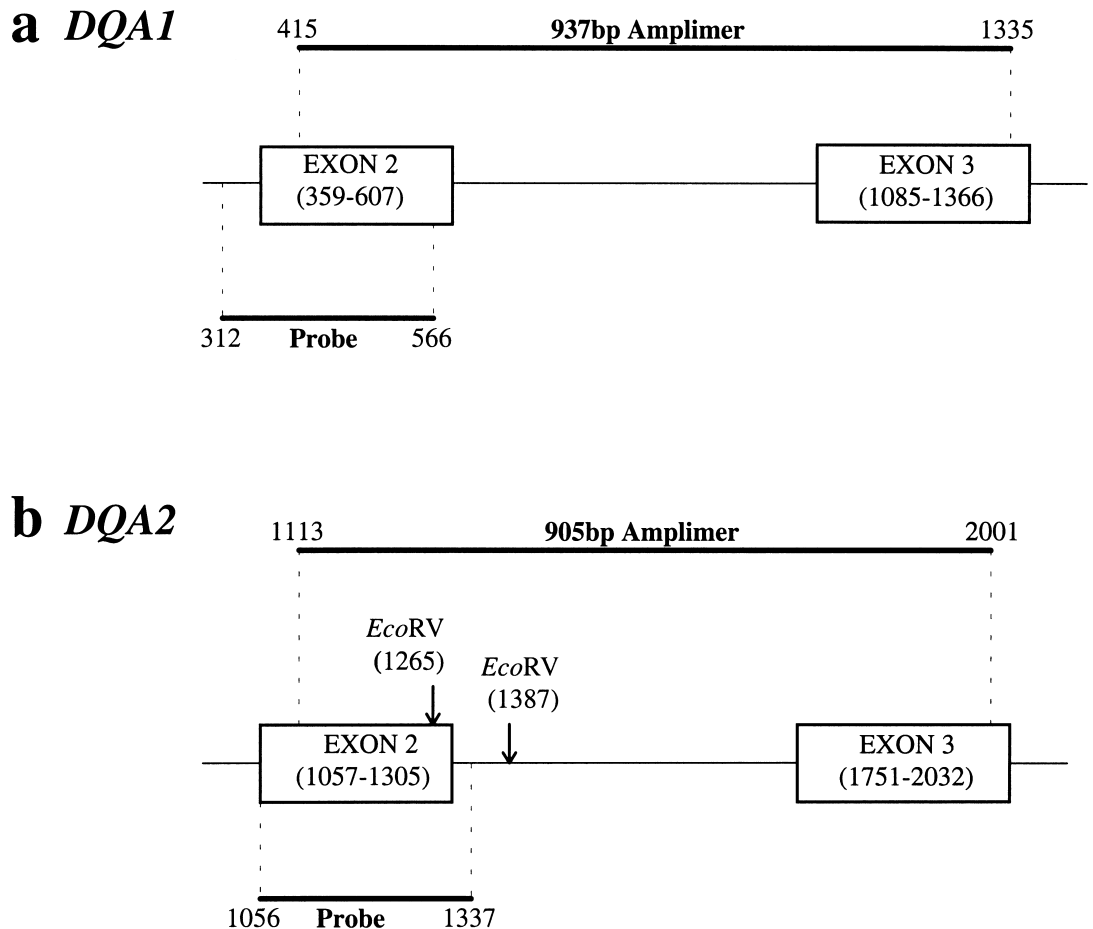

Fig. 1. Diagrammatic representation of exons 2–3 of the *DQA1* and *DQA2* genes. Coordinates are given in bp relative to the published sequences M33304 and M33305 (Scott *et al.* 1991). (a) The exon 2 specific *DQA1* probe and 937 bp amplimer. (b) The exon 2 specific *DQA2* probe, *Eco*RV cleavage sites and 905 bp amplimer.

portion of exon 4 was generated from the plasmid pC13–4BS (Escayg *et al.* 1996), from which the 255 bp *DQA1* exon 2 probe was generated.

PCR primers

Primers used to amplify part of the *DQA2* and *DQA1* genes were gifted by N. Gemmell, University of Canterbury, New Zealand. They were 5′ CCGGATCCCAGTACACCCATGAATTTGATGG 3′ (upstream) and 5′ CCGGATCCCCAGTGCTCCACCTTGCAGTC 3′ (downstream), corresponding to nucleotides 415–437 and 1315–1335 of the *DQA1* gene (Scott *et al.* 1991; GenBank M33304) and nucleotides 1113–1135 and 1981–2001 of the *DQA2* gene (Scott *et al.* 1991; GenBank M33305). The 5′ end of both primers contained a *Bam*HI site and two terminal cytosines and they amplify a 937-bp and 905-bp product from the *DQA1* (Fig. 1a) and *DQA2* (Fig. 1b) genes, respectively. Use of higher annealing temperatures (> 64 °C) favours amplification of only the *DQA2* product.

PCR amplification and analysis

Reactions were performed in a 25-μl volume containing 50 ng genomic template DNA, 200 nM each primer, 100 μM dNTPs (Boehringer Mannheim), 1·5 U *Taq* polymerase (Boehringer Mannheim), and the supplied reaction buffer (containing 1·5 mM MgCl$_2$). All reactions included a negative control. Cycling conditions were 94 °C; 1 min, 65 °C; 1 min, 72 °C; 2 min, for 30 cycles followed by a final extension step at 72 °C for 5 min. One microlitre of the PCR products was digested with 5 U of the appropriate restriction endonuclease (New England Biolabs, Boehringer Mannheim) at 37 °C for 5 h in a total volume of 15 μl. Samples were visualised in 1·8% agarose gels containing ethidium bromide (0·25 μg/ml) and sized with the 100 bp marker (Gibco BRL, Gaithersberg, MD).

Results

Southern analysis of the DQA1 and DQA2 genes

Comparison of the results generated with the *DQA1* and *DQA2* exon 2 probes and the restriction enzyme *Taq*I revealed that *DQA2* alleles defined by two hybridising fragments appeared to coincide with the segregation of the *DQA1*-null allele that had been previously described (Scott *et al.* 1991; Fabb *et al.* 1993; Escayg *et al.* 1996). The longer *DQA1* probe also failed to detect *DQA1* specific bands in animals typed as *DQA1*-null, but at lower stringency did detect the *DQA2* locus (results not shown). Analysis of segregation within families confirmed the association of *DQA2-F*, *-G*, *-I* and *-J* with the *DQA1*-null allele. Only one animal each possessed *DQA2-M*, *-O* and *-P*, the *DQA2* genotypes of these animals being *G/M*, *J/O* and *G/P*, respectively. As each animal was homozygous for *DQA1*-null, it is probable that *DQA2-M*, *-O* and *-P* were also associated with the *DQA1*-null allele. No animal was observed which generated a *DQA2* allele comprising two hybridising bands without the *DQA1*-null allele. Animals typing with other *DQA2* alleles (*DQA2-B*, *-C*, *-D*, *-E*, *-H*, *-K*, *-L*, and *-N*) had identifiable *DQA1* alleles.

Amplification of the DQA2 gene

The association of *DQA2* alleles comprising two hybridising bands with the *DQA1*-null allele suggested that the absence of a *DQA1* locus may coincide with a duplication at the *DQA2* locus.

However, the presence of two bands may have been explained by alleles *F*, *G*, *I* and *J* containing a *Taq*1 site within exon 2. Cleavage at such a site could have produced two fragments capable of hybridising to the probe sequence.

To investigate whether duplication of *DQA2* had occurred, an experiment was devised to ascertain how many sequences the *DQA2* exon 2 probe was detecting. The *DQA2* gene of animals homozygous at *DQA2* (both with and without detectable *DQA1* alleles) was amplified using PCR under conditions that favoured amplification of the *DQA2* amplimer. These amplimers were restriction mapped. Restriction enzymes were chosen with reference to the published *DQA2* gene sequence (Scott *et al.* 1991), and so as to cut once, close to the upstream or downstream boundaries of the region detected by the *DQA2* exon 2 probe. Ideally they would not cut within this region.

While no enzyme was found that fitted these criteria, sequence data (Scott *et al.* 1991) indicated there were two *Eco*RV restriction sites in the *DQA2* gene. One site was located in exon 2 at position 1265 while the other occurs in intron 2 at position 1387 (Fig. 1b). Both sites are within the region that was amplified, between nucleotides 1113 and 2001. In contrast, only the exon 2 restriction site occurs in the probe binding region, which corresponds to nucleotides 1056–1337 (Fig. 1b).

*Eco*RV digestion of the *DQA2* amplimer indicated that both *Eco*RV restriction sites
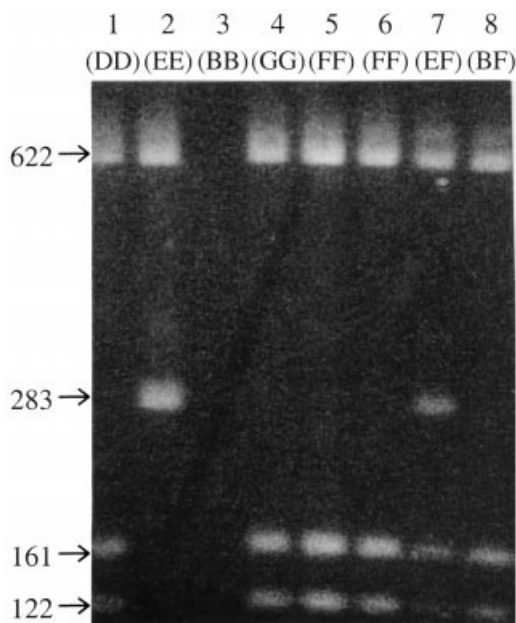
Fig. 2. *Eco*RV digestion of the 905 bp *DQA2* amplimer. Sheep homozygous for alleles *DQA2-F* (lanes 5 and 6), *-G* (lane 4) and *-D* (lane 1) have two *Eco*RV recognition sites in their amplimer producing bands of size 622, 161 and 122 bp, while *DQA2-E* (lane 2) has one site producing bands of 622 and 283 bp. No amplimer was obtained for the *BB* homozygous sheep.

Fig. 3. (a) Autoradiograph of DNA from 6 sheep (*DQA2* genotype marked) digested with *Eco*RV and probed with the *DQA2* exon 2 probe. Two hybridising bands were seen for alleles *G* and *F* while only one was seen for alleles *D* and *E*. (b) Membrane from (a) reprobed with *DQA1* exon 2 probe. No bands are seen for sheep homozygous for *DQA2* alleles *G* and *F* (marked *).

were polymorphic. Conservation of both restriction sites was associated with *DQA2-F*, *-G* and *-D*, whereas the exon 2 (position 1265) restriction site was absent in *DQA2-E* (Fig. 2). There was no amplimer visible with *DQA2-B*. This suggested that upon Southern hybridisation, the *DQA2* exon 2 probe would bind to one restriction fragment with allele *E* and a maximum of two fragments with alleles *F*, *G* and *D*. One of the fragments with alleles *F*, *G* and *D* would be 122 bp long and would not be visible on Southern hybridisation.

Duplication at the DQA2 locus

Animals containing alleles *F*, *G*, *D* and *E* were chosen for Southern hybridisation analysis with restriction enzyme *Eco*RV. Upon probing with the *DQA2* exon 2 probe, two hybridising bands were seen for alleles *G* and *F* while only one was seen for alleles *D* and *E* (Fig. 3a). No band was visible for the predicted 122 bp fragment with alleles *F*, *G* and *D*. Reprobing of the same membrane with the *DQA1* exon specific probe (Fig. 3b) confirmed that alleles *G* and *F* were associated with the *DQA1*-null allele and that the extra bands detected by the *DQA2* probe were not *DQA1*-like.
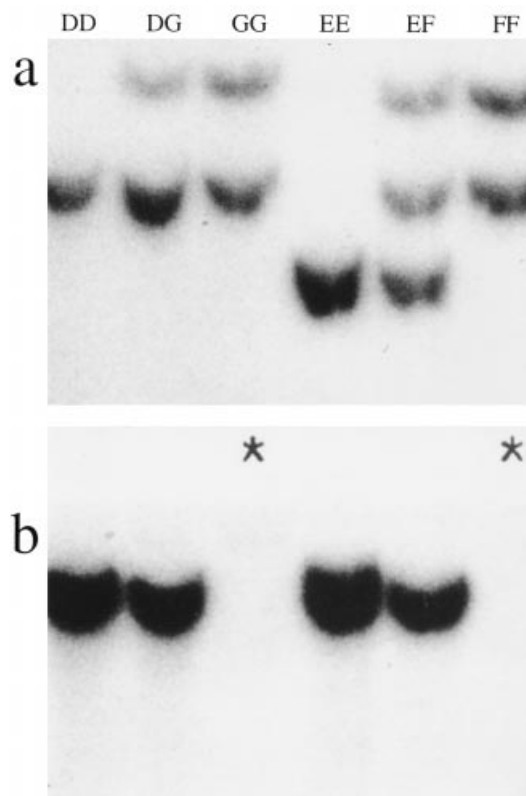
Discussion

This study has demonstrated the presence of haplotypes containing two *DQA*2-like sequences in some sheep. This duplication appears to be exclusively associated with a *DQA1* null allele, retaining the pattern of two *DQA* loci per haplotype. The two ovine *DQA*2-like sequences detected are similar given the hybridisation stringency of the exon 2-specific probe (Fig. 3a,b). This suggests that the two *DQA*2-like loci may have arisen from a gene duplication event.

Given that both the *DQA1/DQA2* and *DQA1 null/DQA2* (duplicated) haplotypes are found in sheep the question arises as to which is the ancestral state. The similarity between the *DQA2* sequences suggests that the duplication event giving rise to two *DQA*2 loci occurred recently. However, most sheep (82–89%) contain the *DQA1/DQA2* haplotype (Escayg *et al.* 1996). Furthermore, the *DQA1/DQA2* haplotype is extremely polymorphic with at least eight

*DQA*1 alleles and 16 *DQA*2 alleles detected to date (Escayg *et al.* 1996; Snibson *et al.* 1998). Most researchers agree that a major driving force behind MHC evolution is the generation of diversity. Given that the *DQA1/DQA2* haplotype is more diverse than the *DQA1 null/DQA2* it seems likely that the *DQA*1 null/*DQA*2 haplotype preceded this haplotype. However, the maintenance since breed development of 11–18% (Escayg *et al.* 1996) of sheep with the *DQA*1 null/*DQA2* haplotype suggests that there is little advantage in having a distinct *DQA*1 gene.

Comparison of the nucleotide sequence of exon 2 of the ovine *DQA*1 and *DQA*2 genes with those of cattle has shown that the ovine *DQA*1 and *DQA*2 genes are more similar to their bovine counterparts than to each other (Snibson *et al.* 1998). Furthermore, two *DQA*2-like sequences in a single haplotype have also been reported in cattle (Ballingall *et al.* 1997). As in the ovine *DQA*1null/*DQA2* haplotypes duplication of the *DQA*2 region was associated with a *DQA1* null gene (Ballingall *et al.* 1997). However, the two bovine *DQA*2-like loci were diverse and are categorised as a *DQA*2 and *DQA*3 sequence. Snibson *et al.* (1998) showed that some ovine *DQA*2 sequences bear a much closer similarity to the cattle *DQA*3 sequences than to other ovine *DQA*2 sequences. This suggests that the duplicated ovine *DQA*2 gene may be analogous to the cattle *DQA*3 gene. The presence of a duplicated *DQA*2 gene in both cattle and sheep suggests that the duplication event took place prior to their divergence ≈ 15 million years ago (Klein & Figueroa 1986). Following the duplication event these genes have undergone independent evolution to produce distinct *DQA*1 and *DQA*2 genes (Scott *et al.* 1991).

The *DQA* analogue in rabbits, rats and mice is a single copy gene (Flavell *et al.* 1985; LeGuern *et al.* 1987; Watters *et al.* 1987). Therefore, duplication of the *DQA* locus must have occurred after divergence of the rodent lineage. In comparison humans have a duplicated *DQA* gene, but this locus comprises a single functional HLA-*DQA1* gene and a *DQX* pseudogene. Divergence of the artiodactyl and primate lineages occurred about 75 million years ago (Shoshani 1986). If the duplication event that gave rise to the two human *DQA-like* genes is the same as that which gave rise to two *DQA* genes in sheep and cattle then HLA-*DQA1* will be more similar to the bovine and sheep *DQA*1 gene than to the *DQX* gene. This is not the case (Snibson *et al.* 1998), suggesting that the duplication event which gave rise to the two *DQA*-like genes in humans is independent of that which produced two *DQA* genes in cattle and sheep (Fabb *et al.* 1993; Snibson *et al.* 1998). This conclusion is supported by differences in arrangement of the human and sheep/cattle loci. A distance of 22 kb separates the ovine *DQ* loci whereas 70 kb separates the two human *DQ* loci (Campbell & Trowsdale 1993).

There is no evidence that either the ovine *DQA*2-like gene described here or the bovine *DQA*3 genes are expressed (Ballingall *et al.* 1998; Snibson *et al.* 1998). Detailed sequence analysis and expression studies are required to establish their functionality. Even if they prove to be pseudogenes they may still be the source from which the *DQA*1 gene arose. Opinion is divided over whether pseudogenes can become functional new genes or are simply slowly degenerating redundant genes. It has been reported that the ribonuclease gene cluster may have evolved from pseudogenes (Trabesinger-Ruef *et al.* 1996). However, this is in contrast to the more widely accepted theory that pseudogenes are irretrievably lost as functional genes. It is possible that in concerted evolution, gene conversion events repair pseudogenes to form fully functional genes.

References

Ballingall, K.T., Luyai, A. & McKeever, D.J. (1997) Analysis of genetic diversity at the *DQA* loci in African cattle: evidence for a *BoLA-DQA3* locus. *Immunogenetics* **46,** 237–44.

Ballingall, K.T., Marasa, B.S., Luyai, A. & McKeever, D.J. (1998) Identification of diverse *BoLA DQA3* genes consistent with non-allelic sequences. *Animal Genetics* **29,** 123–9.

Buitkamp, J., Filmether, R., Stear, M.J. & Epplen, J.T. (1996) Class I and class II major histocompatibility complex alleles are associated with faecal egg counts following natural, predominantly *Ostertagia circumcincta* infection. *Parasitology Research* **82,** 693–6.

Campbell, R.D. & Trowsdale, J. (1993) Map of the human MHC. *Immunology Today* **14,** 349–52.

Escayg, A.P., Hickford, J.G.H. & Bullock, D.W. (1997) Association between alleles of the ovine major histocompatibility complex and resistance to footrot. *Research in Veterinary Science* **63,** 283–7.

Escayg, A.P., Hickford, J.G.H., Montgomery, G.W., Dodds, K.G. & Bullock, D.W. (1996) Polymorphism at the ovine major histocompatibility complex class II loci. *Animal Genetics* **27,** 305–12.

Fabb, S.A., Maddox, J.F., Gogolin-Ewens, K.J., Baker, L., Wu, M.-J. & Brandon, M.R. (1993) Isolation, characterization and evolution of the ovine major histocompatibility complex class II *DRA* and *DQA* genes. *Animal Genetics* **24,** 249–55.

Flavell, R.A., Allen, H., Huber, B., Wake, C. & Widera, G. (1985) Organisation and expression of the MHC of the C57 Black/10 mouse. *Immunological Reviews* **84,** 29–50.

Horne, C., Goodfellow, P.J., McDonald, H.L., Junker, A.K., Keown, P.A. & Hoar, D.I. (1993) A new HLA-DRB1 allele formed by an intraexonic interallelic crossover. *Tissue Antigens* **42,** 141–3.

Klein, J. & Figueroa, F. (1986) Evolution of the major histocompatibility complex. *Critical Reviews in Immunology* **6,** 295–386.

Kuhner, M.K. & Peterson, M.J. (1992) Genetic exchange in the evolution of the human MHC class II loci. *Tissue Antigens* **39,** 209–15.

LeGuern, C., Weissman, J.D., Marche, P.N. *et al.* (1987) Sequence determination of a transcribed rabbit class I gene with homology to HLA-DQ$_\alpha$. *Immunogenetics* **25,** 104–9.

Montgomery, G.W. & Sise, S.A. (1990) Extraction of DNA from sheep white blood cells. *New Zealand Journal of Agricultural Research* **33,** 437–41.

Rask, L., Andersson, L., Gustafsson, K. & Jonsson, A.-K. (1990) Parsimony analysis of mammalian class II histocompatibility genes. *Immunological Reviews* **84,** 187–206.

Rask, L., Gustafsson, K., Larhammar, D., Ronne, H. & Peterson, P.A. (1985) Generation of class II antigen polymorphism. *Immunological Reviews* **84,** 123–43.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual. (*2nd edn.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Scott, P.C., Gogolin-Ewens, K.J., Adams, T.E. & Brandon, M.R. (1991) Nucleotide sequence, polymorphism and evolution of ovine MHC class II *DQA* genes. *Immunogenetics* **34,** 69–79.

Shoshani, J. (1986) Mammalian phylogeny: Comparison of morphological and molecular results. *Molecular Biology of Evolution* **3,** 222–42.

Snibson, K.J., Maddox, J.F., Fabb, S.A. & Brandon, M.R. (1998) Allelic variation of ovine MHC class II *DQA1* and *DQA2* genes. *Animal Genetics* **29,** 356–62.

Stear, M.J. & Wakelin, D. (1998) Genetic resistance to parasitic infection. *Review Scientiifique et Technique Office International Des Epizooties* **17,** 143–53.

Tokunaga, K., Saueracker, G., Kay, P.H., Christiansen, F.T., Anand, R. & Dawkins, R.L. (1988) Extensive deletion and insertions in different MHC supratypes detected by pulsed field gel electrophoresis. *The Journal of Experimental Medicine* **168,** 933–40.

Trabesinger-Ruef, N., Jermann, T., Zankel, T., Durrant, B., Frank, G. & Benner, S.A. (1996) Pseudogenes in ribonuclease evolution: a source of new biomacromolecular function? *FEBS Letters* **382,** 319–22.

Watters, J.W.F., Locker, J.D., Kunz, H.W. & Gill, T.J. (1987) Polymorphism and mapping of the class II genes in the rat. *Rt*1*:. B, RT*1*:. D* and *RT1.H*, a new *DP*-like region. *Immunogenetics* **26,** 220–9.

Weigel, K.A., Freeman, A.E., Kehrli, M.E., Stear, M.J. & Kelley, D.H. (1990) Association of class I bovine lymphocyte antigens complex alleles with health and production traits in dairy cattle. *Journal of Dairy Science* **73,** 2538–46.

Wright, H. & Ballingall, K.T. (1994) Mapping and characterization of the *DQ* subregion of the ovine MHC. *Animal Genetics* **25,** 243–9.

Xu, A., van Eijk, M.J.T., Park, C. & Lewin, H.A. (1993) Polymorphism in *BoLA-DRB3* exon 2 correlates with resistance to persistent lymphocytosis caused by bovine leukemia virus. *Journal of Immunology* **151,** 6977–85.