CHAPTER 4

MHC-DRB EVOLUTION

Chapter 4: MHC-DRB Characterization and Evolution

Introduction

The major histocompatibility complex (MHC) is a multigene family of tightly linked homologous genes. The genes of the MHC encode for glycoproteins that function as receptors presenting self and foreign peptides to circulating T cells (Klein 1986). The complex is divided into three classes (I, II, III) of tightly linked genes (Figure 20). Class I genes encode receptor molecules which are comprised of an α chain and non-covalently associated β microglobulin component. These receptors are found on all nucleated cells and function to present self-peptides to the immune system for self-recognition. If a somatic cell is infected by a virus or becomes neoplastic, the foreign derived peptide is loaded into the class I receptor so that circulating cytotoxic T cells (CD8+) can recognize the "diseased" cell and eliminate it. Class II genes encode for a heterodimers (composed of α and β chains) that are exhibited on the cell surface of antigen presenting cells such as macrophages, B-cells, dendritic cells (Figure 21). Once an antigen-presenting cell, for instance a macrophage, is infected by a pathogen (bacteria, fungi, protist), the cell internally processes the parasite and presents a small antigen in the peptide-binding region (PBR) on the MHC receptor. This binding groove is composed of approximately 50 amino acid residues and presents peptides that are usually 11-19 residues in length (Blum and Creswell 1998). The presentation of the foreign peptide is then recognized by a T-cell receptor (TCR) that is specific to the antigen and MHC molecule. The interaction of receptors stimulates the proliferation of T cell expansion, release of cytokines and secretion of antibodies. The immune cascade described enables for the clearing of infection in the animal. Class II genes also have important roles for the processing of antigen (peptide) and immune response. For example, peptide degradation

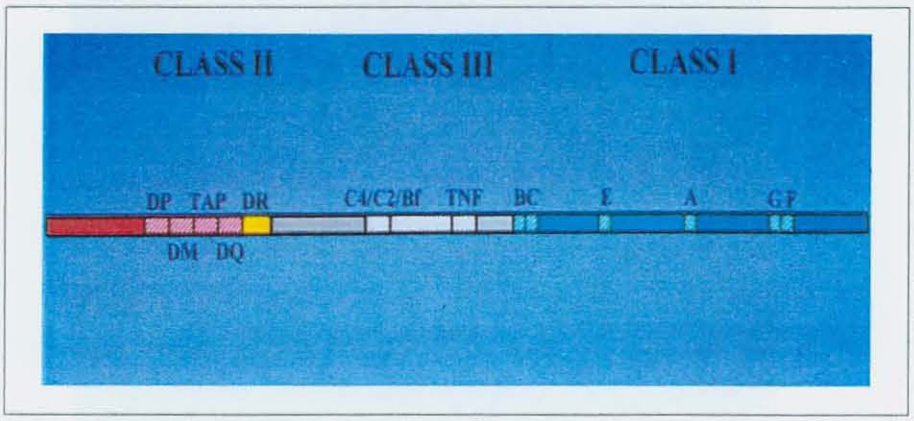


Figure 20. Major Histocompatibility Complex. The illustration depicts the gene organization of mammalian MHC genes. The DR locus is located within Class II.

Figure 21. Three-dimentional structure of the DRB allele Dama*16. The amino acid chain consists of an α-helix and β-sheet. A pocket of amino acids are involved in antigen binding.

and processing is enabled through the binding of a low molecular mass polypeptide (LMP). After the peptide is processed, it is transported into the endoplasmic reticulum (ER) all of which is facilitated by the proteins encoded by the TAP genes of class III (Neefjes et al. 1993).

The genes that encode the MHC protein receptors are the most polymorphic loci of all the nuclear encoding genes in vertebrate species (Hughes and Hughes 1995). The extreme polymorphism takes the form in a high number of alleles found at an individual locus, a high number of amino acid substitutions and large sequence divergence between alleles. Recovery from disease may be related to the MHC profile of the animal. This dependence or "restriction" of MHC is often characterized as co-evolving with parasites. Although the receptors are restricted in a sense, the binding grooves are able to display a range of peptides and are not limited to only one peptide. The binding is most likely dependent upon the physiochemical nature of the amino acids of the foreign peptide.

A selection of HLA alleles in humans has been associated with susceptibility to autoimmune diseases such as lupus (Gladman et al. 1999), Grave's disease (Chen et al. 1999) and subtypes of multiple sclerosis (Yamasaki et al. 1999). Other studies have found strong evidence for specific HLA alleles offering protection against infectious diseases including malaria in humans (Hill et al. 1991, Gilbert et al. 1998), Marek's disease in chickens (Briles et al. 1977) and bovine leukemia in cattle (Xu et al. 1993). In contrast, other MHC haplotypes tend to accelerate disease progression of hepatitis B (Thurz et al.1997) and AIDS (Carrington et al. 1999). Specific MHC allele polymorphisms are also being used to design polyepitope vaccines against foot and mouth disease virus in cattle (Glass and Miller 1994) and melanoma in humans (Mateo et al. 1999).

Balancing selection has been regarded as the strongest force acting on MHC loci to maintenance of high allelic diversity. The two types of balancing selection are

overdominant selection (Hughes and Nei 1989) and frequency dependent selection (Takahata and Nei 1990). Positive (overdominant) selection is evident when the number of nonsynonymous (d_N) substitutions exceeds the number of synonymous (d_S) substitutions (Maruyama et al. 1981), which can be represented as ($d_N > d_S$). The peptide binding region (PBR) of MHC molecules display a large number of amino acid substitutions (Hughes and Nei 1989, Hughes et al. 1990). The diversity of amino acids at these binding positions enables a wide variety of pathogen-derived peptides to be displayed to the immune system (Doherty and Zinkernagel 1975). Selection pressure on peptide binding residues should be greatest during speciation, when organisms move into new environments and their MHC molecules encounter new pathogens (Dixon et al 1996).

There are two approaches to the theory of polymorphism by way of frequency dependent or pathogen-driven selection. The heterosis model explains that heterozygous individuals have a selective advantage over those that are homozygous because a wider array of pathogens can be presented to the immune system. This type of selection has been shown in the study of hepatitis B prevalence in West Africans (Gambia) where individuals heterozygous for specific alleles were resistant (Thurz et al. 1997). In theory, this selection pressure seems likely given the constant co-evolution between host and parasite over time. The other approach to parasite driven selection assumes that an individual can acquire a selective advantage if he/she possesses an allele that has arisen recently in the population. In this case, the new antigen protects the host because the pathogen has not had time to evolve against it (Bodmer 1972). Although, parasite driven selection has been offered as an important force that has shaped HLA frequencies in human populations (Hill et. al. 1991), there is a paucity of field data that supports this theory (Hughes and Nei 1992). Moreover, it is difficult to separate the forces of over-

dominant (heterozygote advantage) and parasite driven selection that may work together in maintaining polymorphisms in populations.

The hypothesis that polymorphism of alleles is maintained by some form of balancing selection is able to explain both the diversity and ancient persistence of MHC alleles (Hedrick and Thompson 1983, Hughes and Nei 1989). The discovery of ancient polymorphisms within MHC allelic lineages provides more evidence for their evolution under balancing selection rather than neutrality. Polymorphism is maintained in allelic lineages that evolve in a trans-specific manner, whereby MHC alleles are passed from one species to another descendant species (Klein 1980). Evidence of ancient polymorphisms from primates (Mayer et al. 1988), rodents (Figueroa et al. 1988), pinnipeds (Hoelzel et al. 1999) and felids (Yuki and O'Brien 1999) supports the trans-specific mode of evolution of MHC loci.

Much of research that has resulted in the elucidation of MHC evolution and organization has been carried out of humans and domestic animals (Bodmer et al. 1998, Lewin et al. 1999). The extremely polymorphic nature of the MHC loci makes them appropriate makers for estimating genetic diversity in populations (Potts et al. 1992). Recently, these genes have been investigated to assess genetic diversity in populations of exotic species. High allelic diversity has been documented in MHC genes in most outbred terrestrial species (see review in Edwards and Hedrick 1998), with some exceptions (e.g. Murray et al. 1995). Demographic crashes can result in either severe loss of MHC alleles as demonstrated in populations of moose (Mikko and Andersson 1996), cheetah (Yuhki and O'Brien 1990) and northern elephant seal (Hoelzel et al. 1999), or a slight reduction in polymorphism as seen in populations of African buffalo (Wenink et al. 1998). Given the paramount role of MHC molecules to species survival, it has been

recommended that diversity at these loci should be maintained in natural populations (Hughes1991, Potts et al. 1994).

In the previous chapters, non-coding loci were examined to assess genetic diversity in each subspecies. The aim of this study was to characterize the DRB locus in order to reveal the pattern of polymorphism found in the exotic African bovid species, *Damaliscus pygargus*. Variation at the sequence level was investigated to define the regions of functional significance involved in antigen presentation. DRB sequences derived from distantly related ungulate species were aligned with Dama alleles so that the mode of evolution and relationships of alleles could be revealed in a phylogenetic analysis.

MATERIALS AND METHODS

Samples and Collection Sites

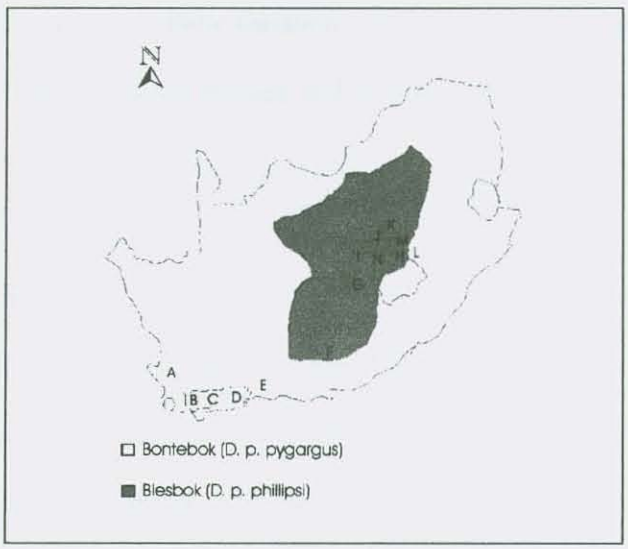
See Figure 22.

PCR-SSCP of the DRB locus

Cattle-specific primers HL030 and HL031 were used to amplify a 284 bp fragment of exon 2 of the DRB locus from genomic DNA as described by van Eijk et al. (1992). PCR product by thermocycling for one cycle at 94° C for 4 minutes, 30 cycles at 94° C for 1 minute, 30 seconds at 65° C, and 72° C, followed by an extension for 5 minutes at 72 ° C.

PCR products (8 µl) were mixed with a 32 µl volume of a low ionic strength buffer (LIS) and subjected to heat denaturation at 97 C ° for 2 minutes. The single strands were formed within the LIS sugar matrix (10% saccharose, 0.01% bromophenol blue, 0.01% xylene cyanol FF) and remained stable at room temperature (Maruya et al. 1996).

Figure 22. Sample Distribution



Sampling locations and number of animals genotyped are as follows: A = West Coast National Park (10), B = Overberg Farm (8), C = Bontebok National Park (14), D = Heidleberg Farm (2), E = Elandsberg Farm (4), F = Cradock Farm (2), G = De Brug Farm (2), H = Golden Gate National Park (3), I = Bloemhof Farm (3), J = Parys Farm (25), K = Suikersbosrand Reserve (2), L = Sterkfontein Dam Reserve (2), M = TDR Farm (5) and N = Fairview Farm (2).

Single-stranded products were then subjected to electrophoresis through a 10% non-denaturing polyacrylamide gel (1.4% cross-linking) in 1X TBE (Tris Borate EDTA) buffer at room temperature for 18 hours (Glavac and Dean 1993). After electrophoresis, the gel was incubated for 30 minutes with the fluorescent GelStar Nucleic Acid Gel Stain (BioWhittaker). Alleles were visualized by UV illumination and image capture with a Kodak digital camera. Allele patterns were scored for each animal and genotypes assigned. Unique patterns of allelic variation were investigated by sequence analysis. In most cases, gel fragments were excised and purified for sequencing (Sambrook et al. 1989).

Cloning and Sequencing

A subset of alleles as revealed by SSCP mobilities were chosen for cloning and sequencing. Homozygous individuals were directly sequenced from purified DRB PCR products (see Chapter 2-Materials and Methods). Heterozygous genotypes were cloned into a pGEM vector (Promega) using a TA cloning kit. Ten recombinant clones were purified by mini-prep filter columns (Qiagen) and sequenced in the forward and reverse directions using the BigDye Terminator Cycle Sequencing Ready Reaction Kit (ABI). The labelled PCR fragments were analysed using an ABI 377 automated sequencer.

Statistical and Phylogenetic Analyses

All Dama-DRB nucleotide sequences and one cow reference sequence (Bota-1) were aligned and translated into the corresponding amino acid sequences using MEGA software package (Kumar et al. 1993). Amino acid positions involved in peptide binding were identified by comparison with the peptide binding groove structure of the human class II molecule (Brown et al. 1993). Relative frequencies of nonsynonymous (dN) and

synonymous (dS) substitutions were calculated for the peptide binding region (PBR) and non-PBR according to Nei and Gojobori (1986) using the Jukes-Cantor correction. (1969). All dN and dS frequencies plus standard errors were estimated using the MEGA computer package. Significance levels were determined by the student's t-test where t = d/ s(d). Heterozygosity levels were determined assuming random mating within subpopulations and deviation from Hardy-Weinberg equilibrium was tested using Markov chain permutation test of 100,000 steps in Genepop. Genetic distances were estimated for the nucleic acid sequences by applying the Kimura two-parameter method. Neighbor-joining trees were constructed from distance matrices of ungulate alleles using the MEGA program and PAUP* (Swofford et al. 1998).

Genebank Sequences

Cattle sequences representing 39 of the major BLA allelic lineages were acquired through the BoLA Nomenclature Website: http://www2.ri.bbsrc.ac.uk.bola/) and used as in a phylogenetic analysis of cattle DRB exon 2 sequences: AF144544, AF144545, AJ001999, AJ002001, U77067, U78548, X87643 – X87645, X87647 – X87650, X87652, X87655, X87656, X87658, X87659, X87662, X87665 – X87668, X87670, Z30650, Z30649, Z36538, Z36542, Z82023 – Z82025, Z82027, Z82029, Z82032 - Z82035. DRB sequences derived from horse, ELA-DRB*9 (L7674) and moose, ALA1DRB1*8 (X83284) were used as outgroup taxa for the phylogenetic analyses.

RESULTS

The Dama-DRB homolog to the cow BoLA-DRB3 locus was characterized by SSCP analysis followed by cloning and direct sequencing of alleles. I am confident that only one DRB locus was amplified as no more than two sequences were revealed in each animal. Cloned PCR products revealed only two alleles through sequence analysis of a minimum of eight recombinant clones for any heterozygote. Alignment of all Dama alleles demonstrates close homology to the cattle (BoLA) DRB3 expressed locus. This close homology and the absence of stop codons within the nucleotide sequences provide evidence that the expressed DRB locus was isolated in *D. pygargus*.

Sequence variation of exon 2 was examined by SSCP analysis revealing 27 different alleles in the typed individuals (Figure 23, Table 13). Genotype frequencies are given in Table 14. Extensive polymorphism was found for the blesbok subspecies (22 alleles) compared to the level of polymorphism found for the bontebok subspecies (6 alleles). Two alleles are shared between subspecies (Dama*7 and *8). The sample Bb1 from the Fairview hybrid population displayed alleles from both subspecies. Heterozygosity values are shown in Table 15.

Polymorphism at class II loci occurs predominately in exon 2, which encodes a majority of the peptide region (PBR). The exon translates into a sequence length of 83 amino acids having 16 possible binding sites for foreign peptide presentation (Brown et al. 1993). In an alignment of all inferred exon amino acid sequences, 39 sites (47.0%) were variable while the other sites were conserved. The majority of amino acid site changes were found within the PBR or in positions that are adjacent to these binding sites, such as sites 23, 24 and 52 (Figure 24). Table 16 illustrates the number of amino acid differences between each allele. The relative frequency of nonsynonymous substitutions (dN = 0.299; SE = 0.066) was significantly higher than the frequency of synonymous

Figure 23.a.

SSCP analysis of DRB alleles in blesbok

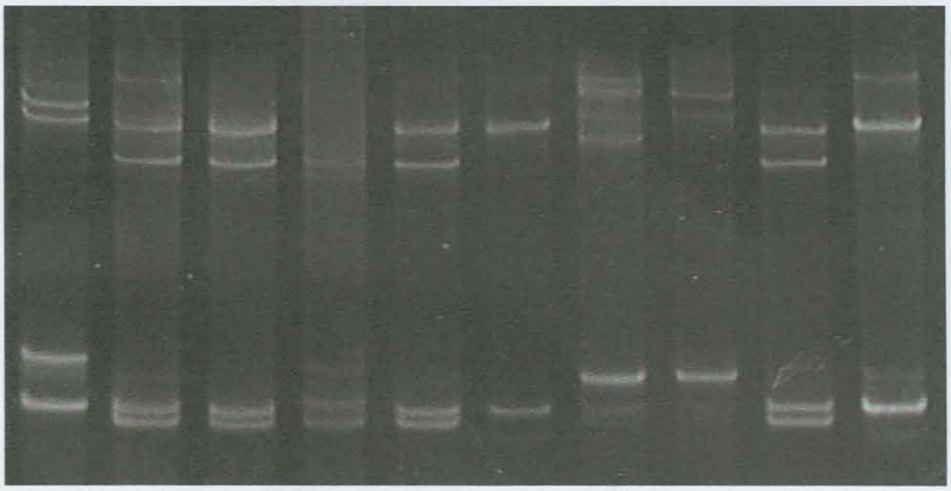


Figure 23.b.

SSCP analysis of DRB alleles in bontebok

Animal	Allele 1	Allele 2
JP1	1	19
JP2	1	19
JP3	19	19
JP4	3	4
JP5	18	20
JP6	13	21
JP7	20	20
JP8	4	8
JP9	2	3
JP10	1	6
JP11	1	4
JP13	1	4
JP14	1	4
JP15	1	4
JP16	4	5
JP17	1	2
JP18	1	2
JP19	1	4
JP20	2	4
JP21	2	4
JP22	1	4
JP23	2	4
JP24	1	1
JP26	1	2
SR1	1	4
SR2	1	1
JT2	26	27
JT3	15	15
JT4	17	26
JT5	15	17
JT6	15	15
BL2	6	12
BL3	15	15
BL4	1	4
DB2	2	3
DB3	17	3
ST3	4	9
ST4	10	11
Bb2	24	25
CR1	2	3
CR3	2	3
GG26	7	8
GG27	8	8
GG31	8	8

Table 13.a. DRB genotypes of blesbok samples.

Animal	Allele 1	Allele 2
BB1	8	8
BB2	8	8
BB3	8	8
BB4	8	8
BB5	7	23
BB6	14	16
BB7	14	22
BB9	14	22
BB10	14	22
EL3	8	8
EL5	7	8
EL8	7	14
EL10	8	8
JO1	7	8
JO2	8	8
JO4	8	8
JO5	7	8
JO6	7	14
JO7	8	8
JO8	8	8
HB1	8	8
HB2	8	8
DH2	14	16
DH8	8	8
DH6	14	16
DH11	14	14
WC1	14	23
WC2	14	16
WC3	14	23
WC4	14	23
WC5	14	16
WC6	14	14
WC8	14	14
WC10	14	14
WC11	14	14
WC12	8	8
BN1	14	23
BN2	14	16
BN3	8	7
BN4	8	7
BN5	14	14
BN6	14	14
WB1	8	8
WB8	14	23
Bb1	14	23

Table 13.b. DRB genotypes for bontebok samples.

Table14. DRB allele frequencies found within each subspecies.

DRB Allele	Blesbok	Bontebok
Dama*01	0.204	0.000
Dama*02	0.113	0.000
Dama*03	0.068	0.000
Dama*04	0.170	0.000
Dama*05	0.011	0.000
Dama*06	0.023	0.000
Dama*07	0.011	0.090
Dama*08	0.068	0.390
Dama*09	0.011	0.000
Dama*10	0.011	0.000
Dama*11	0.011	0.000
Dama*12	0.011	0.000
Dama*13	0.011	0.000
Dama*14	0.000	0.344
Dama*15	0.079	0.000
Dama*16	0.000	0.067
Dama*17	0.034	0.000
Dama*18	0.011	0.000
Dama*19	0.045	0.000
Dama*20	0.034	0.000
Dama*21	0.011	0.000
Dama*22	0.000	0.033
Dama*23	0.000	0.078
Dama*24	0.011	0.000
Dama*25	0.011	0.000
Dama*26	0.023	0.000
Dama*27	0.011	0.000

Table 15. Allelic diversity values estimated for blesbok and bontebok.

Subspecies	N	# of Alleles	D	H_{exp}	H_{obs}
blesbok	44	22	10.7(2.5)	0.89	0.98
bontebok	45	6	5.0(1.6)	0.71	0.41

N indicates the total number of samples genotyped for each subspecies. The average genetic distance between alleles (Jukes-Cantor) are represented as percentage D values by MEGA software (Kumar et al. 1993). Standard errors of calculated distances are indicated in parentheses. Heterozygosity levels are given for observed and expect (Nei 1987).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	0.0																									
2	10.0	0.0																								
3	9.0	1.0	0.0																							
4	14.0	16.0	15.0	0.0																						
5	15.0	17.0	16.0	1.0	0.0																					
6	14.0	16.0	15.0	0.0	1.0	0.0																				
7	8.0	7.0	6.0	16.0	17.0	16.0	0.0																			
8	7.0	16.0	15.0	14.0	15.0	14.0	12.0	0.0																		
9	1.0	11.0	10.0	15.0	16.0	15.0	9.0	8.0	0.0																	
10	15.0	17.0	16.0	1.0	2.0	1.0	17.0	15.0	16.0	0.0																
11	14.0	15.0	14.0	3.0	4.0	3.0	15.0	13.0	15.0	4.0	0.0															
12	14.0	16.0	15.0	0.0	1.0	0.0	16.0	14.0	15.0	1.0	3.0	0.0														
13	14.0	16.0	15.0	1.0	2.0	1.0	16.0	14.0	15.0	2.0	3.0	1.0	0.0													
14	8.0	16.0	15.0	14.0	15.0	14.0	13.0	1.0	9.0	15.0	13.0	14.0	14.0	0.0												
15	8.0	15.0	14.0	13.0	14.0	13.0	12.0	4.0	9.0	14.0	13.0	13.0	13.0	3.0	0.0											
16	8.0	17.0	16.0	15.0	16.0	15.0	13.0	1.0	9.0	16.0	14.0	15.0	15.0	2.0	5.0	0.0										
17	5.0	11.0	10.0	15.0	16.0	15.0	8.0	8.0	6.0	16.0	13.0	15.0	15.0	7.0	8.0	9.0	0.0									
18	5.0	14.0	13.0	11.0	12.0	11.0	12.0	8.0	6.0	12.0	11.0	11.0	11.0	9.0	9.0	9.0	8.0	0.0								
19	9.0	2.0	1.0	15.0	16.0	15.0	6.0	15.0	10.0	16.0	14.0	15.0	15.0	15.0	14.0	16.0	10.0	13.0	0.0							
20	14.0	16.0	15.0	0.0	1.0	0.0	16.0	14.0	15.0	1.0	3.0	0.0	1.0	14.0	13.0	15.0	15.0	11.0	15.0	0.0						
21	14.0	17.0	16.0	2.0	3.0	2.0	16.0	14.0	14.0	3.0	5.0	2.0	3.0	14.0	13.0	15.0	15.0	11.0	16.0	2.0	0.0					
22	8.0	16.0	15.0	14.0	15.0	14.0	13.0	2.0	9.0	15.0	14.0	14.0	14.0	1.0	3.0	3.0	8.0	9.0	15.0	14.0	14.0	0.0				
23	7.0	16.0	15.0	14.0	15.0	14.0	12.0	1.0	8.0	15.0	14.0	14.0	14.0	2.0	4.0	2.0	9.0	8.0	15.0	14.0	14.0	1.0	0.0			
24	7.0	15.0	14.0	13.0	14.0	13.0	12.0	2.0	8.0	14.0	12.0	13.0	13.0	1.0	2.0	3.0	6.0	8.0	14.0	13.0	13.0	2.0	3.0	0.0		
25	4.0	13.0	12.0	14.0	15.0	14.0	11.0	3.0	5.0	15.0	12.0	14.0	14.0	4.0	7.0	4.0	5.0	5.0	12.0	14.0	14.0	5.0	4.0	5.0	0.0	
26	8.0	15.0	14.0	13.0	14.0	13.0	12.0	3.0	9.0	14.0	12.0	13.0	13.0	2.0	1.0	4.0	7.0	9.0	14.0	13.0	13.0	3.0	4.0	1.0	6.0	0.0
27	6.0	12.0	11.0	14.0	15.0	14.0	9.0	7.0	7.0	15.0	12.0	14.0	14.0	6.0	7.0	8.0	1.0	7.0	11.0	14.0	14.0	7.0	8.0	5.0	4.0	6.0

Table 16. Distance matrix computed using the number of amino acids differences found between pairwise comparisons of each OTU. All Dama alleles are num 1 – 27 and highlighted in colors: (yellow = private blesbok alleles, light blue = private bontebok alleles, and green = shared alleles.

substitutions (dS=0.047; SE = 0.040) in the PBR for all alleles (Table 17). In the non-PBR region, substitutions were found to occur much less frequently (dN = 0.049; SE = 0.011 : dS = 0.019; SE = 0.012).

The number of replacement sites in the α-helix (positions 174-250) of exon 2 was compared to that of the β-sheet (positions 1-168) in order to define the region where selection pressure is greatest. In both subspecies, the number of replacement sites in the PBR was nearly equal in both the α-helix and the β-sheet. Moreover, the number of synonymous changes within exon 2 of Dama alleles falls far below the rate substitution of cattle and most other ruminants (Mikko et al. 1999). This result suggests that diversity of Dama alleles has evolved rapidly compared to cattle and other species (Table 18).

Neighbor-joining (minimum evolution), maximum likelihood and maximum parsimony trees based on 249 bp of exon 2 were constructed using Dama alleles and outgroup sequences derived from horse, cow and moose. Table 19 lists the sequence parameters derived from a maximum likelihood analysis that were incorporated for subsequent phylogenetic inference. Genetic distances were estimated by using the HKY model (Hasegawa et al. 1985) and the Kimura 2-parameter model (Kimura et al. 1980) for the minimum evolution (ME) analysis. Both distance measures produced strongly supported trees of equal topology revealing three evolutionary lineages of Dama alleles (Figure 25). The phylogeny constructed from the maximum likelihood estimation recapitulates this evolution. Most of the alleles found in the bontebok subspecies cluster together within group A (Dama*8, 14, 16, 22, 23), while Dama*7 clusters in C. The maximum parsimony analysis estimated 1005 equally parsimonious trees due to the lack of informative sites (39 out of 249 total sites). Although the number of informative sites was low, the MP analysis estimated a consensus tree topology (using 50% majority rule) similar to the ME and ML trees (tree not shown).

Table 17. Jukes-Cantor corrected proportion of dS and dN substitutions.

Exon 2	PBR dN	PBR dS	P-value	Non-PBR dN	Non-PBR dS	P-value
Full length	29.8(6.6)	4.7(4.0)	**	4.9(1.0)	1.9(1.1)	ns
β sheet	32.1(11.1)	1.9(2.7)	**	5.4(1.4)	2.3(1.9)	ns
α helix	32.8(11.6)	8.0(8.3)	*	3.0(1.6)	0.5(0.5)	ns

*Note - The positions involved in the peptide binding region (PBR) were inferred using human DRB1 sequence data (Brown et al. 1993). Exon 2 was divided according to the coding regions for the β sheet (postions 1-168) and the α helix (postions 174-249). Significance values were estimated using the student t-test and are indicated by * <0.1, ** <0.05, *** <0.001 and ns (non-significant).

Table 18. Species comparison of exon 2 diveristy within homologous DRB loci.

Species	# of alleles	PBR dN	PBR dS	Non-PBR dN	Non-PBR dS	P value
Cattle	68	41.1(5.1)	13.0(6.6)	4.4(0.7)	2.2(1.2)	<0.001
Sheep	20	29.8(5.6)	5.2(5.5)	3.1(0.6)	1.5(0.7)	<0.01
Moose	10	11.3(3.2)	0	0	1.1(1.1)	<0.001
Roe deer	3	23.5(7.4)	2.5(4.6)	1.8(0.9)	2.8(2.0)	<0.05
Reindeer	11	28.0(5.9)	9.8(7.4)	3.9(1.0)	3.1(1.7)	<0.05
Red deer	49*	42.7(5.5)	10.5(4.8)	6.4(0.9)	6.0(1.8)	<0.001
Blesbok/Bontebok	27	29.9(6.6)	4.7(4.0)	4.9(1.0)	1.9(1.1)	<0.01

*Note - Percentage relative frequencies of non-synonymous (replacement site) and synonomous (silent) substitutions within the peptide binding region (PBR) and non binding regions (non-PBR). The * indicates that alleles were derived from more than one locus. Adapted from Mikko et al. 1999.

Table 19. Estimation of nucleotide sequence parameters of Dama alleles for phylogenetic analysis.

Parameters	Estimates
Base frequencies	
A	0.225760
C	0.234422
G	0.348934
T	0.190884
Ts:Tv ratio	0.778616
Proportion of invariant sites	0.290924
α shape parameter	0.811716
# informative sites	39.0

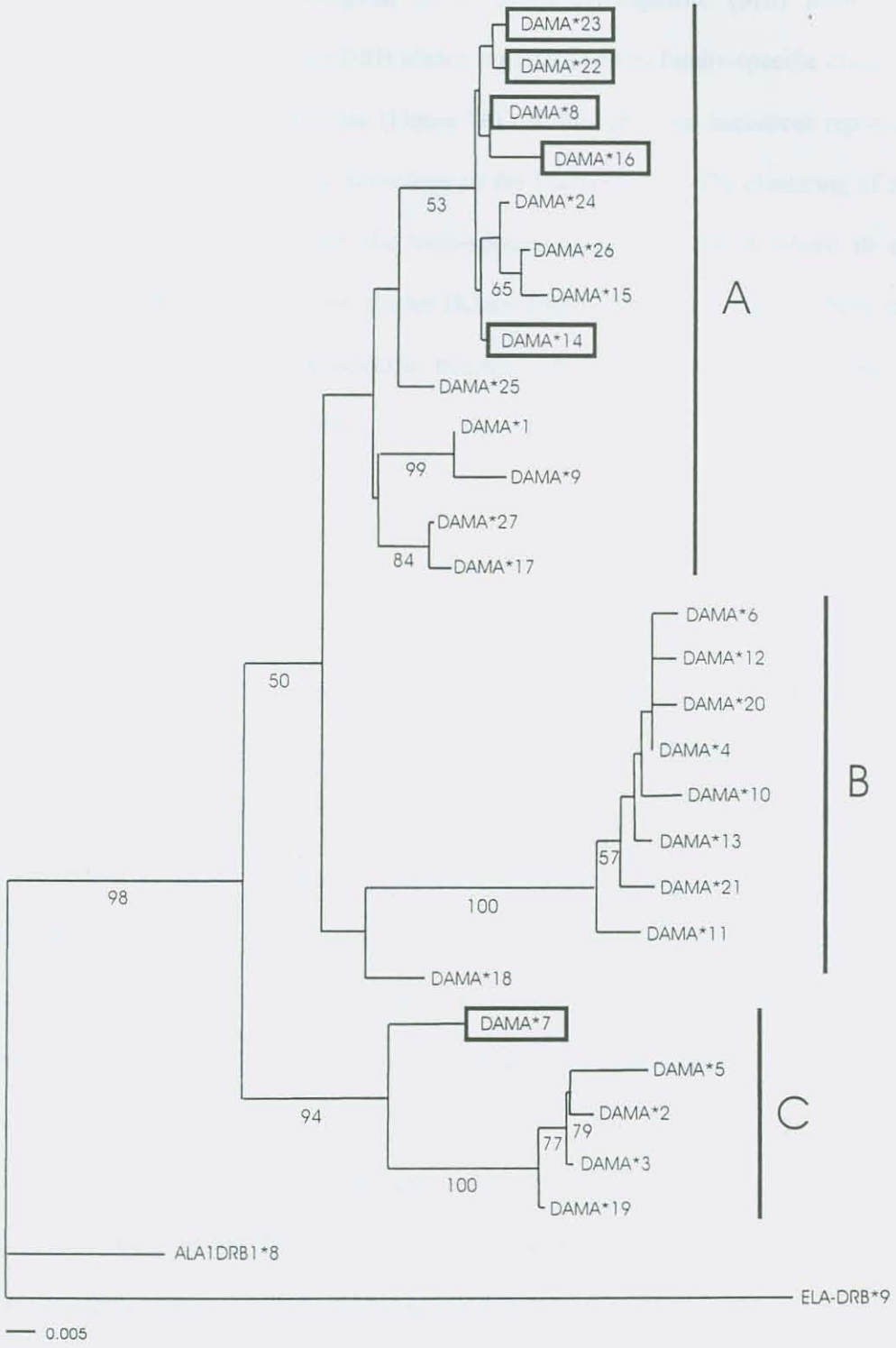


Figure 25. Minimum evolution tree of Dama alleles rooted by horse (ELA-DRB*9) and moose (ALA1DRB1*8) sequences. The tree topology displays three evolutionary lineages of Dama alleles. Bootstrap vaules (1000 iterations) are indicated at the nodes. Bontebok alleles are indicated by boxes.

The relationship between Dama alleles and the domestic cow (*Bos indicus* and *Bos tarus* = BoLA) alleles was investigated in a second phylogenetic (ME) analysis. The evolutionary relationship of the DRB alleles is represented by family-specific cluster as all sequences cluster into Bovidae (Figure 26). In this analysis, sequences representing 39 major BoLA lineages strong homology to the Dama alleles. The clustering of alleles from the bovid species displays the trans-species mode of evolution where all alleles derived from Bovidae form one cluster (Klein 1986). However, Dama alleles cluster strongly together in a species-specific manner, which is indicative of the large time divergence between these species.

DISCUSSION

The DRB3 homologue was characterized in *Damaliscus pygargus* in order to study the evolution of this MHC class II locus in an exotic species of bovid, and to reveal the impact of population collapse on DRB3 polymorphism in the bontebok subspecies. In this study I have presented 27 novel ungulate DRB alleles from a threatened species of antelope that are distantly related to the cow. The relative levels of allelic diversity in the two subspecies were consistent with their different demographic histories as well as their geographic subdivision. Furthermore, this investigation provides evidence for the mechanisms of trans-species evolution and balancing selection acting on the DRB locus in a new species.

The DRB profiles of each subspecies reflect past demographic histories and geographic partitioning. The blesbok subspecies had large and sustainable populations throughout the South African grasslands since the first recorded European settlement in the 17thcentury. However, the Great Trek brought the settlers and their guns northward, and with this human migration, blesbok numbers sharply declined due to over hunting.

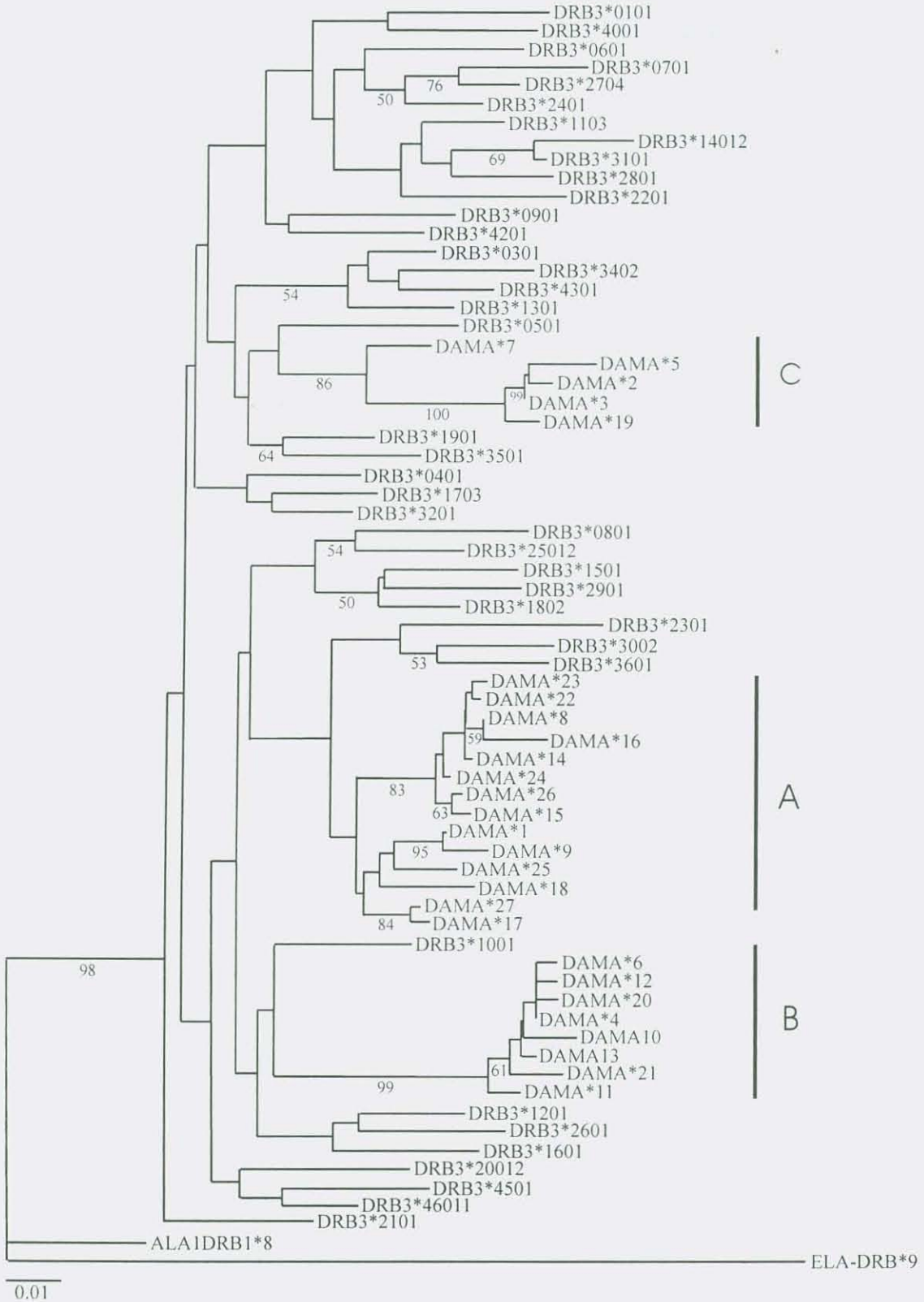


Figure 26. Minimum evolution tree of Dama and Bola-DRB alleles. The tree topology demonstrates the trans-species mode of evolution at this locus as well as species specific clustering of Dama alleles. Bootstrap values <50 are indicated at tree nodes.

The DRB3 data suggests that much of the allelic variation, unlike the mitochondrial diversity, has been retained. The considerable degree of polymorphism observed from our analysis suggests that the DRB diversity must have been generated through large numbers of breeding herds over a long period of time. Historical records have documented a drastic reduction of blesbok numbers throughout their range during the late 1800's; however, an exact census was never known. The founding population was most likely in the hundreds or even thousands in order for the population to increase to 64,000 individuals in the 1960's. The populations that survived into the 20th century were fenced off to prevent further hunting pressure but in doing so, also prohibited the exchange of new alleles between populations. Although there are no existing natural migrations of blesbok, movement of animals has been facilitated by translocation events. This type of management has led to the introduction of alleles from allopatric populations thus increasing diversity.

However, the relict population of bontebok in the southwestern region of the country appears to have maintained only six alleles from the recent founding population. Historic bontebok populations were restricted to a small region bound by a mountain range and ocean with no opportunity for migration. More recently, genetic variation was depleted by population collapse leaving only a residual amount of diversity from the 27 founding individuals that were saved in 1839. The effective population size (*Ne*) was not sufficient enough to re-establish polymorphism (Nei, 1975).

The allelic diversity of each subspecies demonstrates a significant amount of genetic substructure within the species. Two alleles (Dama*7 and *8) were shared between the subspecies while all others (n = 25) were private to each subspecies. MHC diversity measured in other species including chinook salmon runs (Kim et al. 1999) and beluga whale populations (Murry et al. 1999) was reported as significantly differentiated.

The genetic differences at this locus can be used to differentiate between bontebok and blesbok subspecies as well as identify hybrid animals. In one case, blesbok and bontebok alleles were found in a hybrid population indicating that at some point there had been genetic admixture in that herd.

There are several possible explanations for the robust structuring of the DRB locus, an unexpected result given the trans-specific pattern of evolution observed for MHC loci (see Klein 1987;below). First, private alleles in one subspecies may be present in the other but failed to be sampled in this analysis. However, the sample size in this study was adequate (178 chromosomes examined) and the populations sampled were representative of the species distribution across South Africa. Of course, extensive sampling could ultimately uncover more shared alleles as well as additional novel alleles. Second, the partitioning of DRB alleles may represent the phylogeography of the species as bontebok and blesbok populations have been separated by the expanse of the Karoo Desert. Some lineage sorting could have occurred in these geographically isolated populations as a result of genetic drift, especially given the population contractions experienced by the bontebok (cf. Yuhki and O'Brien 1990). This scenario is probably the most likely, given the relatively small number of alleles found in bontebok. Some novel alleles may represent new mutations, possibly maintained by heterotic selection. Lastly, the private alleles may have had adaptive value to the subspecies and signify historic selective event. Because the bontebok subspecies experienced a second bottleneck event that was caused partly by parasitic infection, certain alleles may have been selected in response to presentation of the foreign antigen (Hamilton 1980).

A comparison of amino acid sequences showed extensive polymorphism within the PBR. The relative frequencies of nonsynonymous substitutions exceed that of synonymous substitutions in the PBR, suggesting that positive selection has maintained

the allelic polymorphism in this species (Hughes and Nei 1989). The number of replacement sites in the α-helix (positions 174-250) of exon 2 was compared to that of the β-sheet (positions 1-168) in order to define the region where selection pressure is greatest (Schwaiger et al. 1993). The similar values and tree topologies (trees not shown) indicate that selection pressure is nearly equal in both sequence regions. Moreover, the bovid phylogeny illustrates the conservation of ancestral polymorphisms that could not exist under a neutral model of evolution (Takahata and Nei 1990). The close homology of Dama alleles to other bovid sequences reflects the transpecies mode of evolution of this MHC locus that has been documented in primates (Mayer et al. 1988), ungulates (Andersson 1994), felines (Yuhki and O'Brien 1997), pinnepeds (Hoelzel et al. 1999) and rodents (Figuera et al. 1988).

These results further indicate that the selective retention of diversity may be through heterosis whereby heterozygous individuals have a selective advantage. It is assumed that allelic diversity at Class II MHC genes evolves to promote pathogen resistance (Doherty and Zinkernagel 1975, O'Brien and Evermann 1988). The striking difference between the levels of the neutral nuclear loci and the coding DRB locus indicates that MHC heterozygosity is generated and maintained by a selective force (Hedrick and Thompson 1983). This pattern of diversity, displaying elevated numbers of heterozygotes for polymorphic MHC class II genes, has been well documented in many outbred species and in humans (Hedrick 1994, Markow et al. 1993).

Disassortative (non-random) mating has been proposed to play a role in selection of alleles whereby animals with dissimilar genotypes mate in order to avoid inbreeding within the population (Potts et al. 1994). Mate selection according to MHC genotypes has been documented in several vertebrate species (see Jordon and Bruford 1998 for review); however, MHC-based mating patterns have been excluded as a mechanism for

generating allelic diversity in Soay sheep (Paterson and Pemberton 1997). There is no evidence to describe disassortive mating in bontebok, and other authors question the capacity for this selective force to account for the large number of amino acid site changes found within the PBR (Yeager and Hughes 1999). It is most likely that a combination of evolutionary mechanisms has contributed to the enrichment of diversity found within both subspecies.

The phylogeny of full-length exon 2 Dama alleles revealed three distinct lineages of the DRB locus in this species. Alleles of each lineage are characterized by short branch lengths, which indicate small genetic distances and a recent origin and evolution. The genetic distances among the clusters were greater. Exon 2 was divided into α-helix and the β-sheet coding regions. An equal number of replacement substitutions were found within the α-helix and the β-sheet coding regions of exon 2. This finding contrasts that of the comparisons made in other ruminant species where far more replacement sites are found within the α-helix regions of cattle, sheep and goats (Schwaiger et al. 1994). It is expected that if the selective force drives the evolution of each region equally, then similar tree topologies will be reconstructed for each region (Schwaiger et al. 1993, Swarbrick et al. 1996).

In a further phylogenetic analysis, Dama alleles were aligned with other ruminant DRB alleles and it was showed that all bovid alleles are monophyletic. The inter-leaving of bovid sequences suggests that the Bovidae lineages were generated prior to the divergence of bovids from a common ancestor 20 million years ago (Vrba 1979). Dama alleles, although interleaved throughout the phylogeny, displayed more species-specific clustering, reminiscent of a phylogeny reconstructed with other ruminant species by Mikko and colleagues (1999). This species-specific clustering suggests that gene diversity of the Dama lineages evolved after the split between the other bovid taxa. This suggests a

possible difference in the selective environment for this species, perhaps especially during the relatively recent radiation of alleles.

High allelic diversity at MHC loci is thought to establish a stronger host defence, thereby increasing individual fitness (Hughes and Nei 1988, Takahata and Nei 1990). So far there are few empirical data sets to substantiate this assumption, and many studies have shown that species having low MHC diversity are still viable (Hoelzel et al. 1999, Mikko et al. 1999). Although bontebok populations are indeed viable and show no apparent morphological signs of inbreeding depression, they are still at risk to diseases that may be transmitted by cattle existing in close proximity to the reserves. Given this possible threat of cattle disease transmission and the lack of MHC variation, bontebok may have a smaller chance of mounting immune responses against pathogens in the long term. One very important difference between the bontebok and other species with low MHC diversity is that population size of this antelope has only expanded to 2,500 – 3,000 individuals since the first bottleneck (1800s). Other species that have undergone severe population crashes such as the Northern elephant seal, American bison and African buffalo have recovered to roughly 100,000 to 1,000,000 animals (Hoelzel et al. 1998; Mikko et al. 1997, Wenink et al. 1998). Although, bontebok are protected against hunting pressure, the status of the subspecies is still critical. Stochastic events (disease, drought) as well as human mediated events (hybridization, poaching) are possible threats to the survival of the remaining populations.