African buffalo maintain high genetic diversity in the major histocompatibility complex in spite of historically known population bottlenecks

P. W. WENINK,* † A. F. GROEN, M. E. ROELKE-PARKER,‡ and H. H. T. PRINS*

*Department of Environmental Sciences, Tropical and Vertebrate Ecology Group, Wageningen Agricultural University, The Netherlands, †Department of Animal Sciences, Animal Breeding and Genetics Group, Wageningen Agricultural University, The Netherlands, ‡Laboratory of Genomic Diversity, National Cancer Institute, Frederick, USA

Abstract

Historical population collapses caused by rinderpest epidemics are hypothesized to have resulted in notable genetic losses in populations of the African buffalo. Polymorphism in the major histocompatibility complex (MHC) DRB3 gene was probed by means of restriction analysis of the sequence encoding the peptide-binding region. Nucleotide substitution patterns agreed with a positive selection acting on this fitness-relevant locus. Buffalo populations from four National Parks, situated in eastern and southern Africa, each revealed a surprisingly high allelic diversity. Current high levels of heterozygosity may be reconciled with historical bottlenecks by assuming that local extinctions were followed by fast recolonization, in accordance with the high dispersive capabilities of buffalo. The specific amplification of DRB3 alleles also enabled the assignment of individual genotypes. For each population sample a deficiency in the expected number of heterozygous animals was found. As overdominant selection on the MHC is predicted to yield an excess of heterozygous individuals, this may not be a locus-specific effect. Several other explanations are discussed, of which increased homozygosity caused by nonrandom mating of buffalo in populations seems the most probable.

Keywords: African buffalo, bottleneck, MHC-DRB, PCR–RFLP, rinderpest, social organization

Received 16 January 1998; revision received 17 April 1998; accepted 20 April 1998

Introduction

The major histocompatibility complex (MHC) forms an essential part of the immune system by guiding the discrimination between self and nonself. Specifically, MHC proteins bind foreign peptides and present these to T cells in order to stimulate the immune response. High allelic diversity of the MHC, maintained by balancing selection, is typical for most vertebrate populations (Klein 1986; Hughes & Nei 1988). Individual alleles can bind foreign antigens with differential affinities, and causal relationships have been established between particular genotypes and increased resistance or susceptibility to infectious disease agents (e.g. Hill et al. 1991; Xu et al. 1993). Although it has been stated that MHC diversity is of general adaptive significance in the defence of populations against infectious disease (Doherty & Zinkernagel 1975; O’Brien & Evermann 1988), the implicit assumptions of a higher average fitness for MHC-heterozygous individuals and a better long-term survival prospect for MHC-diverse populations remain to be demonstrated (Potts & Slev 1995). Consequently, when observations of unusually low MHC molecular variability in natural populations were interpreted as indications of increased genetic risk much dispute was raised (May 1995). A sensitive factor in some of these most-prominent studies is a poor knowledge of the underlying cause for low MHC genetic variation. For example, both for cheetah (*Acinonyx jubatus*) and moose (*Alces alces*) a population bottleneck in the distant past could only be inferred from variation in neutral molecular markers (Menotti-Raymond & O’Brien 1993; Mikko & Andersson 1995a). Alternative explanations have been put forward, however, suggesting a continuously small effective population size for cheetah and a reduced parasitic exposure of moose due to their solitary existence (Ellegren et al. 1996; Hedrick & Gilpin 1996).
Here, we present results on MHC genetic variation in natural populations of the African buffalo. Wild populations of Cape buffalo (*Syncerus caffer caffer*) occur scattered over the savannahs and open woodlands of eastern and southern Africa. Together with elephant (*Loxodonta africana*) and wildebeest (*Connochaetes spp*.), buffaloes are the most dominant species in terms of biomass in these natural ecosystems (Prins & Douglas-Hamilton 1990). A central aspect in the life history of buffalo is their social organization in herds (Sinclair 1977; Prins & Van der Jeugd 1993). Although not endangered today, African buffalo almost became extinct in the course of a morbillivirus rinderpest pandemic that ravaged the African continent at the end of the last century. Rinderpest was initially introduced around 1889 with cattle, and while many other wild ruminant populations collapsed, buffaloes were affected most severely overall (Mack 1970; Plowright 1982; Prins & Van der Jeugd 1993). The extreme sensitivity of buffalo can best be explained by their evolutionary relatedness to cattle and their gregarious life style, facilitating fast lateral transmission of the virus. For example, where buffalo had been common before the epidemic, they were nearly wiped out in the Serengeti ecosystem of East Africa (Sinclair 1977) and in the Kruger National Park area of South Africa, where only 20 buffaloes in one herd were reported to survive in 1902 (Stevenson-Hamilton 1911). Rinderpest could be expelled from southern Africa after the initial outbreak, but epidemics kept reoccurring in eastern Africa throughout this century, most recently in 1996 in Tsavo National Park, Kenya. As a consequence, in many parts of Africa rinderpest has been the major determinant of the dynamics of buffalo and wildebeest populations, and indeed of ecosystems as a whole (Sinclair 1977; Prins & Van der Jeugd 1993). Therefore, on the basis of extreme regional fluctuations and many local extinctions, it can be postulated that a considerable loss of genetic variability occurred in African buffalo and that modern populations are marked by low molecular diversities (analogous to the suggestion by O’Brien & Evermann (1988) for wildebeest).

We have previously shown that an MHC class II DRB gene, orthologous to the cattle DRB3 and the human DRB1 genes, is expressed in African buffalo (Wenink & Prins 1997). This gene is highly polymorphic in humans and cattle (Davies *et al.* 1994; Bodmer *et al.* 1997) and seems to be a most proper candidate for probing a loss of MHC molecular variation in buffalo. Nearly all studies on MHC variation in natural populations to date have used Southern blotting to estimate restriction polymorphism from several loci (including pseudogenes) simultaneously (Edwards & Potts 1996; but see Mikko & Andersson 1995a). Here, by specific amplification of the sequence coding for the DRB3 antigen-binding site, we are able to study functional MHC allelic variation in a single locus on an individual and population basis.

**Materials and methods**

**National Parks sampled**

A total of 150 African buffalo samples were obtained from four National Parks (NPs) in eastern and southern Africa. In 1994–96, 42 individuals (18 males, 20 females, four individuals not defined; sampling design spread over the park) were sampled in Serengeti NP (Tanzania). In 1991, 23 individuals (sex and sampling design not defined) were sampled in Amboseli NP (Kenya). In 1996, 17 males and 26 females from five different herds were sampled in Kruger NP (South Africa). In 1992–96, 25 males and 17 females from all over the park were sampled in Lake Nakuru NP (Kenya). Samples were collected either from tranquillized animals as whole blood (Serengeti NP: 29 samples), from darted animals as skin-muscle biopsy (Lake Nakuru NP, Amboseli NP, Serengeti NP: 13 samples), or from culled animals as muscle biopsy (Kruger NP). Biopsies were kept in 80% ethanol, 1 mm EDTA (pH 8.0) and stored at room temperature, whereas blood samples were heparinized and kept frozen. DNA was extracted using the Puregene kit (Gentra Systems) according to the manufacturer’s instructions. Samples were imported under special permission of the Dutch Ministry of Agriculture, Nature Conservation and Fisheries.

Serengeti NP has a total land cover of 17 000 km² and is an integral part of the larger Serengeti ecosystem; Amboseli NP covers a 392 km² area and is open to migration with Tsavo NP (21 000 km²); Kruger NP covers a 20 000 km² area and is open along its eastern border; Lake Nakuru NP measures 147 km² of land surrounding its central lake and has been electrically fenced since 1987. In all analyses the total number of samples from one National Park was treated as a single population.

**PCR–RFLP analysis and sequencing**

The method employed for determination of restriction fragment length polymorphism in the PCR-generated exon 2 fragment from a single DRB3-like gene of buffalo has already been described (Van Eijk *et al.* 1992; Wenink & Prins 1997). Amplification of a 284 bp fragment proved highly robust using primers derived from cattle sequence. PCR primers HL030 and HL031 are overlapping the 5' conserved borders of exon 2 by seven and eight nucleotides, respectively, and an additional 3' primer (HL032) was nested in the two-step PCR reaction. Moreover, the exon border sequence was determined from a reverse transcript of forest buffalo (*Syncerus caffer nanus*) and proved identical to that of the cattle DRB3 gene (Wenink & Prins 1997). Nearly identical primers have also been successfully employed in the more distantly related moose (Mikko & Andersson 1995a). Thus, we feel confident about the specific amplification of a single orthologous DRB3 locus.
from buffalo and about the absence of nonamplifying alleles. Alltogether, 236 bp of individual-specific sequence (out of 268 bp exon 2 sequence) was probed for polymorphism in both alleles, by restriction with the four-cutter enzymes HhaI and Rsal and electrophoresis in nondenaturing polyacrylamide gels. Whereas HhaI revealed three different restriction patterns, Rsal could discriminate a total of eight. HhaI and Rsal patterns were combined into single alleles by inference from individuals that showed a homozygous pattern for one or both enzymes (111 cases out of 150), by double-restriction digestion (37 cases) or by cloning and digestion of alleles (two cases).

Sequencing of nine exon 2 alleles was performed as described previously (Wenink & Prins 1997). Alleles were identified with the PCR–RFLP method and subsequently cloned. For each individual at least two clones were sequenced. Sequences were determined from single buffalo in the case of five alleles; for one allele (2h) clones from two unrelated individuals were sequenced and two alleles (1a, 2f) were verified in three different buffalo. Sequences from the same RFLP-defined allele proved identical between these individuals. Alignment of aminoacid and nucleotide sequences was performed using the Microgenie software (Beckman instruments). A phylogenetic tree relating the nine buffalo DRB3 sequences was constructed using the neighbour-joining algorithm in the MEGA package (Kumar et al. 1993). Ten cattle DRB3 sequences were included for comparison (Mikko & Andersson 1995b). Pairwise genetic distances were calculated from the number of nucleotide differences in 236 bp of exon 2 sequence, corrected by the Kimura 2-parameter. Bootstrapping of the tree topology was performed using 500 replications.

Population statistics

Population differentiation was expressed by the fixation index $G_{ST}$ (analogous to $F_{ST}$), appropriate for one locus with many alleles (Nei 1987) and also by $\gamma$, suitable when a large proportion of the total allelic variation resides within populations (Wenink et al. 1993). Expected heterozygosity was calculated from the allelic frequencies in the sample, assuming an infinite allele model of mutation and random mating within populations (Nei & Tajima 1981). The overall inbreeding coefficient $F$ is measured as the fractional reduction in heterozygosity of individuals relative to the total population and is analogous to Wright’s $F_{IT}$. $F$ measures the combined effects of population subdivision ($F_{ST}$) and nonrandom mating within subpopulations ($F_{IS}$) according to the relationship $(1 - F_{IT}) = (1 - F_{ST})(1 - F_{IS})$ (Hartl & Clark 1989). Deviation from Hardy–Weinberg expectations was tested in ARLEQUIN version 1.1 (Schneider et al. 1997), using a Markov chain permutation test of 100 000 steps and 1000 dememorization steps. The exact probability of the observed table of genotypes is expressed under the null hypothesis of a random joining of alleles.

Results

To determine the extent of expressed polymorphism in the buffalo DRB3 gene, a 284 bp exon 2 sequence was amplified for the total of 150 wild Cape buffalo. Sequence polymorphism in this segment was probed by restriction digestion with two four-cutter enzymes. HhaI revealed three different restriction patterns (1–3), whereas Rsal detected eight different patterns (a to h) (Wenink & Prins 1997). A higher resolution of DRB3 polymorphism was obtained by combining both restriction patterns into single alleles. Thus, a total of 13 exon 2 alleles (1a to 2h) could be discerned in the total sample.

To assess the extent of sequence polymorphism underlying the different PCR–RFLP patterns, 236 bp were determined for nine different alleles (results not shown). Extensive polymorphism was observed between these allelic sequences with an average number of 20.5 nucleotide differences (minimum 7, maximum 37). Nonsynonymous nucleotide substitutions occurred on average 4.2 times more often than synonymous substitutions. This bias is reflected in the amino acid sequences that differ on average at 13.1 out of the 78 positions compared.

Most amino acid substitutions are located at or near sites that are known to be directly involved in antigen binding or presentation (Fig. 1). The buffalo nucleotide sequences were related in a phylogenetic tree together with 10 DRB3 sequences from cattle. Overall, a species-specific clustering of alleles can be recognized, whereas some alleles (SyLA*2a, SyLA*2c, BoLA*0801) occur intermingled with alleles from the other species (Fig. 2). A similar topology was obtained when alleles were joined on the basis of non-synonymous substitutions only (results not shown).

Fig. 1 Amino acid sequence comparison for DRB3 exon 2 alleles identified in buffalo. Identity to the top sequence is indicated with dots. ’Bos’ represents a DRB3 sequence from cattle (Groenen et al. 1990). Positions in the peptide-binding site are marked with an arrow. GenBank sequence accession numbers Syn1a–Syn3b are: AF0590233–AF0590241.
Out of 91 possible combinations from 13 alleles, no less than 56 different genotypes were observed in the combined population sample. Two-thirds of all individuals \((n = 101)\) proved heterozygous based on this restriction assay (Fig. 3). The frequency distribution of alleles for each NP-population sample is presented in Table 1. Nine to 10 different alleles can be observed within each park. Only a minority of alleles is rare, with on average 3.3 alleles per park that occur at a frequency lower than 5%. Common alleles are 3b and 2f, that are each found at an overall frequency of 18%. Together, these two alleles represent 61% of the chromosomes analysed for Serengeti NP. A similar skewed distribution is found for Kruger NP, where 51% of alleles is either 3b or 2g. Differences in composition between NPs are most obvious from the distribution of the most frequent alleles: 3b is absent in the Amboseli NP sample, 2f is apparently lacking in Kruger NP and 2g is not found in the Lake Nakuru NP sample. Measures of population differentiation indicate that Kruger NP is most divergent in its overall allelic composition from the other three parks: \(\Delta G^2 = 0.11, \gamma = 0.64\), compared to \(\Delta G^2 = 0.04, \gamma = 0.29\) among the three eastern Africa parks. Pairwise differentiation values between all four populations are given in Table 2.

From the observed DRB3 allelic frequencies per park the within-population expected heterozygosities \((H_e)\) were calculated. This diversity measure was found to be high for each of the four NPs (Table 1). The slightly reduced heterozygosities for Serengeti NP and Kruger NP agree with the less-even frequency distribution of alleles within these parks. When the expected heterozygosity is compared to the observed heterozygosity per population sample \((H_o)\), a consistent deficiency of heterozygous individuals is noted. The fractional reduction of heterozygosity \((F)\) is positive with high significance for three out of four NPs (Table 1). The highest deviation from expectation is observed in Lake Nakuru NP, where no less than 43% of individuals is homozygous at the DRB3 locus.

**Discussion**

Polymorphism in the DRB3 gene of African buffalo is fully consistent with established characteristics of MHC variation (Hedrick 1994). The PCR–RFLP method could discern 13 alleles in a sample of 150 buffalo. The equalized frequency distribution of alleles in four different populations is consistent with some form of balancing selection acting...
on the buffalo DRB3 locus, even though the distributions for Serengeti NP and Kruger NP are more skewed. Allelic sequences proved highly divergent, with most nucleotide substitutions being nonsynonymous. Variable amino acids were predominantly at or near positions known to be of functional importance in the MHC molecule. A phylogeny revealed paraplectic relationships between buffalo and cattle, attesting to the trans-specific inheritance of MHC polymorphism (Klein et al. 1993). The species-specific clustering of alleles fits the substantial time of evolutionary divergence between buffalo and cattle, approximately estimated at 10 Myr (Janecek et al. 1996).

The suitability of the PCR–RFLP method to detect DRB3 polymorphisms is based on the high overall

\[ H_E = 1 - n \left( \sum p_i^2 \right) / n - 1 \], where \( p_i \) is the frequency of the \( i \)th allele in the NP.

\[ F = 1 - (H_O / H_E) \].
between populations, and 1320 P. W. WENINK
satellite for diagnosis of expressed DRB3 polymorphism
allelic sequences based on the two restriction enzymes
have defined a significant share of expressed DRB3 polymorphisms in African buffalo. In cattle, and in some other domestic artiodactyl species, exon 2 alleles can be efficiently typed by analysis of length variation in a microsatellite that is present in the DRB3 second intron (Ellegren et al. 1993; Schaiger et al. 1993). Whereas this microsatellite is both present and highly polymorphic in African buffalo, loose association between the exonic and intronic polymorphisms precludes the use of this microsatellite for diagnosis of expressed DRB3 polymorphism in buffalo (P. Wenink, unpublished results).

Allelic differentiation, as expressed by the fixation indices \( G_{ST} \) and \( \gamma \), was most pronounced between Kruger NP and the three NPs located in eastern Africa. This is not surprising given the distant location of the South African population (average distance over 2000 km) and fits the notion of genetic isolation by distance in buffalo, previously inferred from the distribution of polymorphisms in the mitochondrial genome (Templeton & Georgiadis 1995). Population differentiation measures should not be overvaluated, as they are derived from only a single locus. Balancing selection is expected to oppose DRB3 genetic differentiation between populations and values for \( G_{ST} \) and \( \gamma \) may therefore be underestimates. Alternatively, balancing selection over the course of only few generations is unlikely to have nearly compensated the effects of strong genetic drift associated with the historical epidemic population crashes in buffalo. The allelic frequency differences between parks can thus perhaps best be explained by enhanced genetic drift in combination with limited long-distance gene flow. Strong directional selection by rinderpest virus epitopes for particular DRB3 alleles seems unlikely, given an apparent absence or low frequency for every individual allele in at least one NP population (Table 1).

The level of heterozygosity is high for each NP. This finding was not anticipated, given our a priori assumption of substantial genetic losses caused by the occurrence of widespread bottlenecks in buffalo populations one century ago and thereafter. No obvious link exists between MHC heterozygosity on the one hand and incidence of rinderpest epidemics (once in southern Africa, over five times in eastern Africa), park size (Lake Nakuru NP area is less than 1% that of Serengeti NP or Kruger NP), or current population census size (some 1500 head in Lake Nakuru NP vs. some 25 000 and 35 000 head in Kruger NP and Serengeti NP, respectively) on the other hand. Based on reports of moderate allozymic and extensive microsatellite variation, Kruger NP currently presents the best-studied case of a buffalo population retaining extensive and genome-wide genetic variation (Bruford et al. 1996; Grobler & Van der Bank 1996). Two explanations can reconcile this consolidated observation with the severe historical population crash, known to have occurred in the Kruger NP area (Stevenson-Hamilton 1911). First, the reported number of 20 surviving buffalo may be an underestimate. This is not unlikely given the bushy habitat of Kruger NP and the difficulty of comprehensive census counting. However, buffalo in the Lake Nakuru NP area have been rare quite recently, with repeated countings of \( \sim 30 \) head in the 1970s, probably as a combined result of disease and overhunting (Warden reports, G. Grootenhuis, S. Mwasi, personal communication). The Lake Nakuru NP population sample nevertheless shows a similar high level of heterozygosity as observed for Kruger NP. Second, local extinction of herds and concomitant genetic loss in populations may have been overcome by recolonization of buffalo from elsewhere, introducing new alleles and thus re-establishing high genetic diversity (McCausley 1991). Although buffalo herds are known to occupy home ranges with high fidelity (Sinclair 1977; Prins 1996), settlement of herds into a new area has been observed on several occasions, probably as a result of displacement through cultivation (Prins 1996). Individual bachelor bulls have been recorded to travel between herds over more than 50 km, apparently meeting no difficulty in crossing geographical barriers (Prins 1996). This notion is supported by the low DRB3 genetic differentiation between Lake Nakuru NP, situated in the Rift Valley itself, and both Serengeti NP and Amboseli NP that are located on the plains to the west and to the east, respectively (Table 2). Confirmation of the extinction/recolonization hypothesis may become possible through a detailed quantification of gene flow between regional populations, e.g. by means of neutral genetic variation residing in microsatellite markers and mitochondrial DNA.

A marked departure from Hardy–Weinberg (HW) proportions was observed within each population. Whereas a number of assumptions for HW equilibrium are certainly not met, such as neutrality of the DRB3 locus and a very

© 1998 Blackwell Science Ltd, Molecular Ecology, 7, 1315–1322
large size for each population, the deviation is positive for every park tested (between 7% and 35%), with high significance for three parks (Table 1). A number of explanations can be put forward to account for the observed fractional reductions of heterozygosity. (i) The presence of nonamplifying alleles would cause heterozygous individuals to be marked as homozygous. Overall, we are confident about the robustness of the heterologous PCR primers employed. (ii) Positive assortative mating for the DRB3 genotype would increase the fraction of homozygous individuals in a locus-specific manner. Although this is able to be tested with other nuclear markers, this possibility is in contrast with hypotheses of disassortative mating and selective advantage of heterozygotes to account for observations of an excess of MHC heterozygotes in populations (Hedrick 1994). (iii) Unconscious sampling of genetically differentiated subpopulations existing within the boundaries of an NP would cause a Wahlund effect (Hartl & Clark 1989). The sampling of buffalo within parks was overall geographically spread, from many different herds, and directed towards both sexes. Population subdivision, even within the vast areas of Kruger NP and Serengeti NP, seems unlikely based on the dispersive abilities of buffalo as argued above. The low measures of population differentiation between the three East African parks ($G_{ST} = 0.04$), coupled to a pronounced heterozygote deficiency in the pooled samples from these three populations ($F = 0.22$, exact $P = 0.000$), effectively rules out geographical subdivision within these parks as an explanation for the observed $F$-values. (iv) The social organization of buffalo in herds could cause an increased correlation of identical alleles within the population (‘inbreeding’ in the sense of system of mating in a population, Templeton & Read 1994). Buffalo cows are supposed to maintain a life-long fidelity to the herd in which they are born, whereas bachelor bulls maintain a constant flux between herds in a region, with apparent low fidelity to any particular herd (Prins 1996). The mating system of buffalo has been described as a re-entrant consecutive polygyny, implying a temporal stay within the herd of a fraction of all bulls at any time. Thus, matrilineal relatedness is predicted between the permanent members of a herd (cows and subadult bulls). Assuming strict bonding of buffalo cows in herds juxtaposed to a random dispersal and mating of bulls between the herds that are present in a NP, the population as a whole would be randomly mating and the inbreeding coefficient $F$ would be expected to be near zero (Chesser 1991). It is proposed that the observed homozygosity levels reflect a higher than (population) average relatedness between mates, induced by re-entry of bulls into their natal herd at a higher than random frequency. Heterozygote deficiencies observed among chimpanzees (Pan troglodytes), where, by contrast, females are the dispersing sex, have been interpreted as evidence of kin selection within social communities (Morin et al. 1994). More genetic data are needed to reconcile the suggested effect of social organization in African buffalo.

In conclusion, our analysis of MHC genetic variation revealed high diversities in African buffalo populations in four NPs. This finding was not anticipated, given the historical collapses of these populations from epidemics of rinderpest disease. A clue to this paradoxical finding may be the high dispersive capability of buffalo that is allowing a quick re-occupancy of empty ranges, perhaps in combination with an overestimation of historical population size reductions. Observed deficiencies in the number of MHC-heterozygous individuals may be related to the mating system of buffalo, although other possibilities cannot be formally excluded at present.

Acknowledgements

We are most grateful to all parties that contributed to this study. Samples were kindly provided by the Department of Molecular Genetics, National museums of Kenya (R. Amman), Kruger NP Scientific Services (V. De Vos), and the Serengeti Research Center (M.R.-P). Sampling in Kenya was permitted by the Kenya Wildlife Service and collection was coordinated by J. Grootenhuis, Nairobi. We thank J. van der Poel, L. van der Waaij and W. van Hooft for helpful discussions and A. Baker and A. Schreiber for critical reading of the manuscript. Research was funded by the Netherlands Organization for Scientific Research (NWO) as a postdoctoral grant to P.W.W.

References


Kumar S, Tamura K, Nei M (1993) MEGA: Molecular Evolutionary Genetics Analysis, version 1.01. The Pennsylvania State University, University Park, PA.


This work springs from our initiative to incorporate genetic data into the existing body of knowledge on tropical herbivore communities and is part of a national research program focusing on biodiversity in disturbed ecosystems. Paul Wenink obtained his PhD degree by studying the evolutionary genetics of shorebird populations on a global scale and is currently a postdoctoral researcher. Ab Groen is Associate Professor in Animal Breeding and Genetics. Herbert Prins is Professor in Tropical Nature Conservation and Vertebrate Ecology and has dedicated many years to studying the ecology and behaviour of African buffalo in Lake Manyara National Park, Tanzania.