Spatial and temporal patterns of neutral and adaptive genetic variation in the endangered African wild dog (*Lycaon pictus*)

CLARE D. MARSDEN,*† ROSIE WOODROFFE,‡ MICHAEL G. L. MILLS,§ J. WELDON McNUTT,* SCOTT CREEL,** ROSEMARY GROOM,†† MASENGA EMMANUEL,‡‡ SARAH CLEAVELAND,* PIETER KAT,§§ GREGORY S.A. RASMUSSEN,¶¶ JOSHUA GINSBERG,*** ROBIN LINES,††† JEAN-MARC ANDRÉ,‡‡‡ COLLEEN BEGG,¶¶¶ ROBERT K. WAYNE and BARBARA K. MABLE*

*Institute of Biodiversity, Animal Health & Comparative Medicine, University of Glasgow, Glasgow G12 8QQ, UK, †Royal Zoological Society of Scotland – Edinburgh Zoo, Edinburgh EH12 6TS, UK, ‡Institute of Zoology, Zoological Society of London, Regent’s Park, London NW1 4RY, UK, §The Tony and Lisette Lewis Foundation, Private Bag X5890, Upington 8800, South Africa, ¶Botswana Predator Conservation Trust, Private Bag 13, Maun, Botswana, **Department of Ecology, Montana State University, 310 Lewis Hall, Bozeman, MT 59717, USA, ††Department of Zoology & Entomology, Mammal Research Institute, University of Pretoria, Pretoria 0002, South Africa, †‡Tanzania Wildlife Research Institute, P.O.Box 661 Arusha, Tanzania, †§Investigação Veterinária Independente, R. Constantino Fernandes, 20 3Fte, 1700-11913 Lisbon, Portugal, ¶¶ Wildlife Conservation Research Unit, Department of Zoology, University of Oxford, Tubney House, Tubney, Oxon OX13 5QL, UK, ¶¶¶Wildlife Conservation Society, 2300 Southern Boulevard, Bronx, NY 10460, USA, †††Namibia Nature Foundation, PO Box 245 Windhoek, Namibia, ‡‡‡Museum of Natural History, Eduardo Mondlane University, Maputo, Mozambique, ¶¶¶Niassa Carnivore Project, The Ratel Trust, Postnet Suite 230, Private Bag X18, Ronderebosch 7701, South Africa, ¶¶¶¶Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA

Abstract

Deciphering patterns of genetic variation within a species is essential for understanding population structure, local adaptation and differences in diversity between populations. Whilst neutrally evolving genetic markers can be used to elucidate demographic processes and genetic structure, they are not subject to selection and therefore are not informative about patterns of adaptive variation. As such, assessments of pertinent adaptive loci, such as the immunity genes of the major histocompatibility complex (MHC), are increasingly being incorporated into genetic studies. In this study, we combined neutral (microsatellite, mtDNA) and adaptive (MHC class II DLA-DRB1 locus) markers to elucidate the factors influencing patterns of genetic variation in the African wild dog (*Lycaon pictus*); an endangered canid that has suffered extensive declines in distribution and abundance. Our genetic analyses found all extant wild dog populations to be relatively small (*N* < 30). Furthermore, through coalescent modelling, we detected a genetic signature of a recent and substantial demographic decline, which correlates with human expansion, but contrasts with findings in some other African mammals. We found strong structuring of wild dog populations, indicating the negative influence of extensive habitat fragmentation and loss of gene flow between habitat patches. Across populations, we found that the spatial and temporal structure of microsatellite diversity and MHC diversity were correlated and strongly influenced by demographic stability and population size, indicating the effects of genetic drift in these small populations. Despite this correlation, we detected signatures of selection at the MHC, implying that selection has not been completely overwhelmed by genetic drift.

Keywords: bottleneck, *Lycaon pictus*, major histocompatibility complex, spatial, temporal

Received 25 September 2011; revision received 19 December 2011; accepted 24 December 2011

Correspondence: Clare Marsden, Fax: 1 530 754 0299; E-mail: cmarsdenresearch@gmail.com

© 2012 Blackwell Publishing Ltd
Introduction

Assessing patterns of genetic variation within a species is essential for understanding population structure, local adaptation and differences in diversity between populations (Bos et al. 2008). Furthermore, in conservation genetics, this information is critical for setting conservation priorities and identifying management units. Demographic history and gene flow are key factors that influence patterns of genetic variation. For example, when populations become small and isolated, genetic divergence between populations increases, and genetic diversity is reduced within populations because of higher rates of genetic drift and inbreeding, and lower rates of gene flow (Frankham 1996). Neutrally evolving genetic markers are the most appropriate for elucidating demographic change and genetic structure, and therefore, they have been widely implemented in population genetic studies (Bos et al. 2008). However, the ability of neutral markers to reflect adaptive genetic diversity, which forms the basis of evolutionary change and local adaptation, has been questioned because neutral markers are not subject to selection and therefore may not be correlated with adaptive markers (Hedrick 2001; Reed & Frankham 2001; Aguilar et al. 2004). Thus, adaptive markers must be assessed directly to understand the distribution and level of adaptive variation, as well as the forces that govern it (Reed & Frankham 2001).

Adaptive markers are increasingly being incorporated, alongside neutral loci, to elucidate local adaptation and differences in evolutionary potential between populations (e.g. Campos et al. 2006; Evans et al. 2010; Miller et al. 2010). The major histocompatibility complex (MHC) includes multiple genes involved in the immune response (Piertney & Oliver 2006). MHC loci have extraordinarily high levels of variation, which is thought to be maintained by balancing selection (reviewed in Garrigan & Hedrick 2003). Associations between specific MHC alleles and susceptibility or resistance to specific diseases have been repeatedly shown (reviewed in Piertney & Oliver 2006), suggesting that pathogens can be a primary selective force maintaining MHC variation (reviewed in Spurgin & Richardson 2010). As one of the most well understood adaptive loci, and with clear relevance to population and evolutionary viability, the MHC is currently one of the best markers available to use as a proxy for adaptive variation (Piertney & Oliver 2006; Hoglund 2009; Miller et al. 2010).

Many studies that have assessed both neutral and MHC variation have been based on samples from a small number of populations at a single point in time (e.g. tiger salamander, Ambystoma tigrinum, Bos et al. 2008; brown trout, Salmo trutta, Campos et al. 2006) However, as selective and neutral forces vary both spatially and temporally, genetic variation is also predicted to vary in both space and time (Seddon & Ellegren 2004; Oliver et al. 2009; Evans et al. 2010). For example, temporal changes in pathogen communities within a population are predicted to result in genetic divergence in MHC variation within a population over time (Westerdahl et al. 2004; Charbonnel & Pemberton 2005; Oliver et al. 2009; Fraser et al. 2010). To achieve a more thorough insight into patterns of genetic variation, there is a need for studies based not only on neutral and adaptive markers, but also with samples collected at ecologically appropriate spatial scales (e.g. tuatara, Sphenodon spp. Miller et al. 2010) and from more than one time period (e.g. guppy, Poecilia reticulata, Fraser et al. 2010; salmon, Salmo salar, Consuegra et al. 2011; water vole, Arvicola terrestris, Oliver et al. 2009; great reed warblers, Acrocephalus arundinaceus, Westerdahl et al. 2004). This is particularly important in endangered species, where anthropogenic factors can alter the relative importance of neutral and selective forces that shape patterns of genetic diversity. Understanding the extent of demographic declines and its impact on selection, genetic diversity and population structuring, as well as determining which populations are adaptively different or suffering genetic declines, is critical for identifying threats and prioritizing conservation efforts. However, the level of sampling required for such an extensive study is especially challenging in endangered species, where sample sizes are inherently limited. In addition, few studies have focused on endangered carnivores with large home ranges and social pack structure, where maintenance of corridors to dispersal is critical. The aim of our study was to assess the forces that shape patterns of genetic variation in a species that has suffered large-scale population fragmentation and declines. To this end, we screened for variation at both neutral (microsatellite and mitochondrial DNA) and potentially adaptive (MHC) markers in a spatially and temporally variable set of endangered African wild dog samples (Lycaon pictus, hereafter, wild dog).

Wild dogs are a highly social canid that hunts and breeds cooperatively in packs averaging 5–15 adults (Creel & Creel 2002). They can disperse over considerable distances (>250 km) and occupy large home ranges, which can exceed 2000 km² (Fuller et al. 1992; McNutt 1996). Historically, wild dogs ranged across most of sub-Saharan Africa (Woodroffe et al. 2004b). However, extensive habitat loss and persecution have reduced their range to just 7% of its former distribution (Woodroffe et al. 1997; IUCN/SSC 2009). Today, <8000 individuals remain in the wild, and with few exceptions, extant populations are small and exist in locations isolated from each other by wide stretches of human...
modified habitats (IUCN/SSC 2009). Moreover, disease represents a significant threat to wild dogs, outbreaks of which have resulted in significant population declines in the past (reviewed in Woodroffe et al. 2004a). Considering the endangered status of this species, the sample set upon which our study is based is unique in terms of: (i) the number of samples (>300), which enabled us to accurately assess a variety of genetic parameters including demographic history and signatures of selection; (ii) the spatial scale (13 monitoring areas distributed throughout Eastern & Southern Africa), which allowed us to attain a detailed understanding of gene flow between populations, as well as differences in their genetic status; and (iii) the inclusion of temporal samples (three populations were sampled at two time points), which permitted us to assess the stability of genetic parameters over time in relation to demography. Specifically, we used these samples to address three questions: (i) Is there a genetic signature of demographic decline in wild dogs? (ii) How are neutral and MHC diversity in wild dogs structured temporally and spatially? and (iii) Is there evidence of selection at the MHC?

Methods

Sampling and DNA extraction

Blood, tissue, hair and serum samples were collected from thirteen monitoring areas in Africa (Fig. 1; Tables 1, S1 and S2, Supporting information): (i) Kruger; (ii) Lowveld; (iii) Okavango; (iv) Hwange; (v) Selous; (vi) Masai-Steppe; (vii) Serengeti-Mara; (viii) Lalkipia; (ix) Ghanzi; (x) NE Namibia; (xi) Sofala; (xii) Niassa; and (xiii) Kajiado. However, five or fewer samples were available for the last five localities; therefore, these samples were only used in descriptive analyses of genetic diversity to gain a broader geographical picture of variation. For Kruger, Okavango and Serengeti, samples were available from two temporal periods, hereafter referred to as ‘Old’ and ‘Recent’ (Old = 1980–1995; Recent = post-2000; Tables 1 and S1, Supporting information). Samples were extracted using DNeasy tissue and blood extraction kits (Qiagen, Crawley) according to manufacturer’s instructions. Hair was extracted according to a Qiagen user-developed protocol.

Genetic data collection

To enable comparison with previous work (Girman et al. 2001), we amplified a 327-bp segment of the mitochondrial D-Loop control region, using canid-specific primers (Thr-L and DLH), which were redesigned specifically for wild dogs (Leigh 2005): forward 5’ ACT-ATTCCCTGATCTCCCC 3’; reverse CCTGAAGTAA GAACCAGATGCC. Polymerase chain reactions (PCRs) were performed in a 20-lL reaction volume containing: 1.25 × Q solution (Qiagen); 1.25 × PCR buffer (containing 15 m M MgCl2); 3.1 m M MgCl2; 0.2 m M of each dNTP (Invitrogen, Paisley); 0.19 lM of each primer; 1 unit of

Fig. 1 Mitochondrial DNA variation in African wild dogs. (a) Bayesian geneology demonstrating the relationship between 10 wild dog mtDNA haplotypes, rooted using the grey wolf as an out-group. Node support is shown via Bayesian posterior probability values (%). Haplotype colours correspond with those in (b). (b) Distribution of haplotypes across Southern and Eastern Africa. Frequency of haplotypes per sampling location is depicted by pie charts, with colours corresponding to those shown in (a), and sample sizes shown in brackets. Geographical clustering, as indicated by AMOVA, is shown with coloured ovals. The former and current distribution of wild dogs according to McNutt et al. (2008) is depicted by light and dark grey shading, respectively.

© 2012 Blackwell Publishing Ltd
Table 1 Summary of diversity statistics for mtDNA (mt), microsatellite (ms) and MHC markers

<table>
<thead>
<tr>
<th>Population</th>
<th>n (mt/ms/MHC)</th>
<th>No. of pk</th>
<th>mtDNA</th>
<th>Microsatellite</th>
<th>MHC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>h</td>
<td>$\pi \times 10^3$</td>
<td>$R_s$</td>
</tr>
<tr>
<td>Kruger old</td>
<td>94/67/67</td>
<td>≥9</td>
<td>2</td>
<td>1.45 (±1.4)</td>
<td>4.60</td>
</tr>
<tr>
<td>Kruger recent</td>
<td>–/20/24</td>
<td>6</td>
<td>–</td>
<td>–</td>
<td>4.53</td>
</tr>
<tr>
<td>Lowveld</td>
<td>15/14/15</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>4.09</td>
</tr>
<tr>
<td>Okavango old</td>
<td>90*/12/19</td>
<td>?</td>
<td>6</td>
<td>22.20 (±11.6)</td>
<td>4.54</td>
</tr>
<tr>
<td>Okavango recent</td>
<td>–/42/42</td>
<td>15</td>
<td>5</td>
<td>11.20 (±6.4)</td>
<td>5.68</td>
</tr>
<tr>
<td>Hwange</td>
<td>47/14/19</td>
<td>?</td>
<td>3</td>
<td>19.42 (±10.5)</td>
<td>4.90</td>
</tr>
<tr>
<td>Selous</td>
<td>37/23/22</td>
<td>8</td>
<td>2</td>
<td>6.18 (±4.0)</td>
<td>4.31</td>
</tr>
<tr>
<td>Masai-Steppe</td>
<td>32/32/17</td>
<td>3</td>
<td>1</td>
<td>9.32 (±5.5)</td>
<td>4.97</td>
</tr>
<tr>
<td>Serengeti-Mara old</td>
<td>41*/20/18</td>
<td>≥5</td>
<td>3</td>
<td>7.50 (±4.6)</td>
<td>4.41</td>
</tr>
<tr>
<td>Serengeti-Mara recent</td>
<td>–/13/14</td>
<td>5</td>
<td>2</td>
<td>7.45 (±2.3)</td>
<td>4.67</td>
</tr>
</tbody>
</table>

*This value includes old and recent samples for that population.
Listed are the number of animals (n) and packs (pk) genotyped at each marker for each population, number of haplotypes (h), nucleotide diversity (\(\pi\) ± SE), allelic richness (\(R_s\)), observed heterozygosity (\(H_O\)), and expected heterozygosity (\(H_E\)).

Hot Star tag (Qiagen); and approximately 10 ng of template DNA. We sequenced the following PCR cycles: 5 min at 95 °C; 30 cycles of 95 °C for 30 s, 55 °C for 30 s and 72 °C for 30 s; and a final extension of 72 °C for 10 min. PCR products were sequenced using ExoSAP-IT (USB, OH, USA), sequenced using an ABI 3730 automated sequencer and analysed using GENEIOUS Pro v. 4.5.5 (Biomatters Ltd).

For the mtDNA control region, we resequenced PCR products from 116 of the samples that had been sequenced manually by Girman et al. (2001) to confirm haplotype designations using automated sequencing. However, we included the data for all wild sourced samples (n = 410) presented in Girman et al. (2001) (corrected for errors that we detected) in data analyses. The sample sizes for mtDNA (Table 1; n = 410) are thus larger than for microsatellite and MHC loci (n = 320 and 331, respectively).

We genotyped samples for variation at 10 physically unlinked microsatellite loci (Table S3, Supporting information) according to the protocol outlined in Marsden et al. (2011). In brief, we used ABI fluorescent dyes on the forward primer and amplified the 10 loci as two multiplexes using Qiagen Multiplex PCR mixes. Samples were then run on an ABI 3730 and analysed using GENEMAPPER 4.7 (Applied Biosystems). Low concentration and poorly amplifying samples were genotyped in triplicate. Samples with >30% missing data were excluded from analyses. For each population and locus, we tested for deviations from Hardy–Weinberg equilibrium using GENALEX6 (Raymond & Rousset 1995) and assessed for significance after Bonferroni correction.

Lastly, we assessed variation at the canine MHC, known as dog leucocyte antigen (DLA). The DLA class II includes two functional DQ genes (DLA-DQA1 and DLA-DQB1) and two functional DR genes (DLA-DRA1, DLA-DRB1), which form a tightly linked haplotype and together encode the α and β chains of class II molecules (Wagner et al. 1996). Previous assessments found DLA-DRA and DLA-DQA1 to be monomorphic and DLA-DQB1 to be biallelic in wild dogs (Marsden et al. 2009); therefore, we restricted our analyses to the more polymorphic DLA-DRB1 locus. Specifically, we conducted sequence-based typing at exon two of DLA-DRB1 using primers DRBln1 (CCGTCCCCACACGACATTTT, Wagner et al. 1996) and DRBln2 (CAGGAAACAGCTATGACCTGTGTCACACACCTCAGCACCA, Wagner et al. 1996; M13 tag underlined), and the PCR protocol outlined in Marsden et al. (2009). PCR products were then purified using ExoSAPIT (USB) and sequenced directly using an ABI3730. Sequences were analysed using Match Tools and Match Tools Navigator (Applied Biosystems), with alleles determined using a deductive approach, as described by Kennedy et al. 2002. In brief, this method uses an allele library consisting of previously identified alleles to determine the allele or pair of alleles present in homozgyotes and heterozygotes, with any new alleles added to the library. We cloned products that did not match any pair of known alleles found in two individuals using the TOPO TA cloning system and One Shot Competent cells (Invitrogen). The alleles present in all other individuals could be resolved using the allele library.

Summary statistics

We calculated nucleotide diversity (\(\pi\)) amongst mtDNA haplotypes using ARLEQUIN v3.11 (Excoffier et al. 2006).
Microsatellite diversity and DLA-DRB1 diversity were measured as expected (\(H_E\)) and observed heterozygosity (\(H_O\)), as calculated in *GENALEX* (Peakall & Smouse 2006), and allelic richness standardized for sample sizes (\(R_s\)) as calculated in *HP-RARE* (Kalinowski 2005). To assess whether populations were differentiated in terms of DLA-DRB1 alleles, we computed pairwise Fisher’s exact tests in *GENEPOP* (Raymond & Rousset 1995) and adjusted significance values according to the Bonferroni correction.

### Demographic history

As genes at the MHC are expected to evolve under balancing selection, where trans-species polymorphism can result in extensive allele sharing over long time periods (Klein 1987), only the microsatellite data were used to trace historical demographic patterns. Contemporary estimates of effective population size (\(N_e\)) were calculated in *NEESTIMATOR* v1.3 (Ovenden et al. 2007) with the linkage disequilibrium (LD) test (Hill 1981) applied to all populations, and the moment-based approach applied to populations where temporally separated samples were available (Waples & Yokota 2007).

We tested for evidence of recent bottlenecks using the \(M\) ratio test implemented in the programme MPVAL (Garza & Williamson 2001). As bottleneck tests are sensitive to sampling (Luikart et al. 1998; Piry et al. 1999; Garza & Williamson 2001), we only analysed populations where \(n \geq 30\). The \(M\) ratio test was run with the parameters suggested by the programme authors: \(p_s = 90\%\); \(\Delta_R = 3.5\). We assessed a range of \(\theta\) values (0.2, 0.6, 1.2 and 2.0) corresponding to pre-bottleneck \(N_e\) of 100, 300, 600 and 1000, respectively. We used 10 000 simulations for each run and assessed significance after Bonferroni correction.

To assess the scale and timing of demographic changes, we used the Bayesian coalescent hierarchical model-based approach implemented in *MSVAR* 1.3 (Storz & Beaumont 2002). *MSVAR* uses Markov chain Monte Carlo (MCMC) simulations to estimate the posterior probability distribution of a set of population log normal parameters: current (\(N_0\)) and ancestral (\(N_1\)) effective population size, time since demographic change (\(T\)), and microsatellite allele size distribution assuming a stepwise mutation rate (\(\mu\)). For *MSVAR*, we could only analyse data from Kruger (old), Okavango (recent) and Selous, as the other populations, or sampling periods (e.g. Kruger new), had small sample sizes, population substructure or recolonization history. We used generation times estimated by Creel et al. (2004) for Kruger (6.2 years), Okavango (4.5 years) and Selous (5.4 years) and assumed an exponential demographic model, which is predicted to be most accurate for recent declines (Beaumont 1999). For each population, we ran five chains with different starting points, 50 000 updates and a thinning interval of 50 000. We used wide-ranging, uninformative priors with large variances to limit the effect of the prior on posterior distributions (Goossens et al. 2006), and we tested a range of prior values (\(10^2\)–\(10^6\)). *MSVAR* output was assessed using the BOA package in R v. 2.10 (R core development team). The first 25 000 iterations were discarded as burn-in. Convergence of the remaining 25 000 iterations between the five chains was assessed using the Brooks, Gelman and Rubin statistic (Brooks & Gelman 1998). The last 25 000 updates from each of the five chains were combined to calculate the lower (5\%), median (50\%) and upper (95\%) quantiles of the posterior distributions of the parameters \(N_0, N_1\) and \(T\).

### Spatial patterns of genetic diversity

The genealogy of mtDNA haplotypes was reconstructed in *MrBayes* v3.1.2 (Huelsenbeck & Ronquist 2001) using the best-fit nucleotide substitution model as indicated by *MsModelTest* v2.2 (Nylander 2004; HKY substitution model, Hasegawa et al. 1985, no rate variation between sites). Four chains were run for 3 000 000 generations, with trees sampled every 100 generations. The first 5000 trees were discarded as burn-in. Based on these settings, two independent runs were conducted and we checked for convergence by assessing the standard deviation of split frequencies.

We conducted a hierarchical analysis of molecular variance (AMOVA) with 1000 permutations in *ARLEQUIN* v3.1 (Excoffier et al. 2006). AMOVA’s were conducted on mtDNA haplotype data and DRB and microsatellite allele frequency data, based on \(F_{ST}\) values to allow direct comparison between markers. We further assessed population structure of microsatellite data through principle coordinates analysis (PCoA) and Bayesian clustering analysis. PCoA was conducted in *GENALEX* using genetic distances between individuals and with a median value computed for each population. We used an individual level Bayesian clustering model-based algorithm implemented in *STRUCTURE* 2.3 to elucidate the number of genetic clusters (\(K\)) within our sample set (Pritchard et al. 2000). We assumed no prior population information for 200 000 burn-in and 2 000 000 MCMC iterations for \(K = 1–10\), with 10 replicates per \(K\). To identify the most likely \(K\) value, we followed the Evanno et al. (2005) method. Additionally, we used *STRUCTURE* to test for the presence of migrants (Pritchard et al. 2000), by implementing the USEPOPINFO function, which estimates posterior probabilities of membership per individual to each predefined cluster. Individuals with posterior probability values lower than
70% were identified as putative immigrants to that population.

Mantel tests were used to test for a pattern of isolation by distance at microsatellite and MHC loci. Pairwise population differentiation values were based on the estimator $D_{est}$ (Jost 2008) and calculated with 1000 bootstraps in the programme SMOGD (Crawford 2009). Pairwise geographical distances (km) were based on straight-line distances between population pairs. Mantel tests were computed in the VEGAN package of R (R core development team) with 10 000 permutations for significance testing after Bonferroni correction.

Temporal patterns of genetic diversity

For Kruger, Okavango and Serengeti, where we had samples from two time periods, we tested for temporal changes in microsatellite diversity ($H_E$, $H_O$ and $R_{ST}$) within populations using paired t-tests in MINITAB. At the MHC, we examined trends in genetic diversity ($H_E$, $H_O$ and $R_{ST}$). Fisher’s exact tests were performed in GENEPOP (Raymond & Rousset 1995) to test for significant changes in DLA-DRB1 and microsatellite allele frequencies over time.

Tests for selection

We tested for evidence of historical selection at DLA-DRB1 using the fixed effects likelihood method (FEL) implemented in HYPHY (Kosakovsky Pond & Frost 2005). FEL uses a maximum likelihood approach to estimate ratios of nonsynonymous ($d_N$) to synonymous ($d_S$) substitutions independently for each codon. A likelihood ratio test then determines whether there is evidence of positive ($d_N \geq d_S$) or purifying selection ($d_N \leq d_S$) (Kosakovsky Pond et al. 2006). Prior to implementing FEL, we tested for evidence of recombination amongst our sequences using the CARD method (Kosakovsky Pond & Frost 2005) and partitioned the sequenced according to detected break points where recombination was found. Codons involved in peptide binding (PBR) were taken from Bondinas et al. (2007).

We assessed spatial patterns of genetic variation to detect contemporary selection based on the expectation that selection should result in incongruent patterns of genetic differentiation between neutral and MHC markers (Garrigan & Hedrick 2003). Specifically, we conducted an analysis of covariance (ANCOVA) implemented in R to test whether there was a significant difference in the slopes describing the relationship between $D_{est}$ (genetic differentiation) and geographical distance, when based on DLA-DRB1 relative to microsatellite data. Lastly, we tested whether genetic differentiation at the DLA-DRB1 locus was larger than expected under neutrality, by controlling for microsatellite differentiation using partial Mantel tests.

Results

Summary statistics

We detected 10 mtDNA haplotypes, eight identified by Girman et al. (2001) and two that were new, S4 and S5 (Table S4, Supporting information). Errors were detected by resequencing in three of the eight haplotypes previously identified by Girman et al. (2001). The new and corrected haplotypes are provided in Table S5 (Supporting information). We found Okavango, Hwange and Selous to have the highest mtDNA diversity ($\pi > 0.01$; Table 1), whereas Laikipia and Lowveld had the lowest mtDNA diversity ($\pi < 0.001$; Table 1).

We identified between five and 30 microsatellite alleles per locus (mean = 13) across 10 loci. Only one locus deviated from Hardy–Weinberg equilibrium in two populations (Kruger-Old, FH2611; Okavango Old, FH2658). As for mtDNA, Okavango, Hwange and Selous showed the highest levels of microsatellite diversity ($R_S = 4.54–5.68$, $H_E = 0.68–0.76$; Table 1), and Laikipia and Lowveld the lowest diversity ($R_S = 4.09–4.41$, $H_E = 0.59–0.67$).

We identified 19 DLA-DRB1 alleles with unique amino acid sequences (Fig. S1; Tables 2 and S6, Supporting information). This included 14 alleles described in Marsden et al. (2009), three alleles from Marsden et al. (2011) and two new alleles (DLA-DRB1*90901 and DLA-DRB1*91001). All alleles were unique to wild dogs, and there was no evidence of pseudogenes (stop codons or frameshift mutations) or more than two sequences per individual, suggesting that we had amplified a single functional locus. As for mtDNA and microsatellite loci, the highest DLA-DRB1 diversity was found in Selous, Okavango and Hwange ($R_S = 6.6–9.4$; $H_E = 0.79–0.88$) and the lowest in Laikipia and Lowveld ($R_S = 3$, $H_E = 0.57–0.60$; Table 1).

Demographic history

Contemporary estimates of $N_e$ based on microsatellites were consistently small across populations (mean <30), regardless of the analysis method used (Table 3). This gave $N_e/N$ estimates ranging from 0.02 to 0.21 (Table 3). For populations with temporal data, we compared changes in $N_e$ estimates over time (Table 3) with changes in census sizes (Table S1, Supporting information). In Kruger, $N_e$ estimates decreased by ~50%, which corresponds with a 70% decline in census sizes, and in Okavango, which had stable census sizes, confidence intervals for $N_e$ estimates overlapped. In Seren-
Table 2 Frequencies of DLA-DRB1 alleles in each population across the Eastern and Southern Africa

<table>
<thead>
<tr>
<th>Population (n)</th>
<th>Region</th>
<th>90101</th>
<th>90102</th>
<th>90201</th>
<th>90202</th>
<th>90203</th>
<th>90204</th>
<th>90301</th>
<th>90205</th>
<th>90206</th>
<th>90401</th>
<th>90501</th>
<th>90601</th>
<th>90602</th>
<th>90701</th>
<th>90801</th>
<th>91101</th>
<th>91001</th>
<th>90403</th>
<th>90901</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kruger old (67)</td>
<td>Southern</td>
<td>0.08</td>
<td>0.23</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kruger recent (24)</td>
<td>Southern</td>
<td>0.06</td>
<td>0.17</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.13</td>
</tr>
<tr>
<td>Lowveld (15)</td>
<td>Southern</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Okavango old (19)</td>
<td>Southern</td>
<td>0.16</td>
<td>0.24</td>
<td>0.13</td>
<td>0.03</td>
<td></td>
<td>0.21</td>
<td>0.11</td>
<td>0.05</td>
<td>0.03</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Okavango recent (42)</td>
<td>Southern</td>
<td>0.1</td>
<td>0.21</td>
<td>0.24</td>
<td>0.06</td>
<td></td>
<td>0.05</td>
<td>0.18</td>
<td>0.02</td>
<td>0.08</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hwange (19)</td>
<td>Southern</td>
<td>0.08</td>
<td>0.05</td>
<td>0.13</td>
<td></td>
<td></td>
<td>0.29</td>
<td>0.03</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selous (22)</td>
<td>Eastern</td>
<td>0.02</td>
<td>0.11</td>
<td>0.18</td>
<td></td>
<td>0.05</td>
<td>0.11</td>
<td>0.02</td>
<td>0.09</td>
<td>0.02</td>
<td>0.18</td>
<td>0.18</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masai-Steppe (18)</td>
<td>Eastern</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.06</td>
<td>0.36</td>
<td>0.14</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serengeti-Mara old (18)</td>
<td>Eastern</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.36</td>
<td>0.11</td>
<td></td>
<td>0.33</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serengeti-Mara recent (14)</td>
<td>Eastern</td>
<td>0.14</td>
<td>0.43</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.04</td>
<td>0.18</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laikipia (63)</td>
<td>Eastern</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
<td>0.16</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Areas with ≤5 samples**

| Ne Namibia (4)       | Eastern    | 0.33  | 0.17  | 0.33  |       |       |       |       |       |       | 0.17  |       |       |       |       |       |       |       |       |       |
| Ghanzi (1)          | Southern   |       |       |       |       |       | 0.33  |       |       |       |       |       |       |       |       |       |       |       |       |
| Niassa (1)          | Eastern    |       |       |       |       |       |       |       |       |       |       | 0.5   |       |       |       |       |       |       |       |
| Sofala (3)          | Eastern    |       |       |       |       |       |       |       |       |       |       |       | 0.5   |       |       |       |       |       |       |
| Kajado (1)          | Eastern    |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       | 0.5   |

DLA, dog leucocyte antigen.
geti-Mara, however, Ne estimates decreased by a factor of three, whereas census sizes increased by the same factor, which may reflect inaccurate variation estimates (sample size <30 individuals), a time lag in the genetic signal or recent migration into the population (Marsden et al. 2011).

All populations showed significantly lower M ratios than those expected under mutation–drift equilibrium (Table 4), for at least some values of h (i.e. pre-bottleneck population size), which is indicative of a bottleneck (Garza & Williamson 2001). Bayesian coalescence simulations implemented in MSVAR detected evidence of a large population decline from an ancestral Ne of approximately 600–900, to a current Ne of <10, within the last 100 years, within each of three populations assessed (Table 5). Estimates of time since decline (T) and ancestral (Ne) and current (N0) effective population sizes were consistent across a range of prior values (10²–10⁴), and the signal of decline was still apparent when a prior of population expansion was used (data not shown), thus indicating that the signal from the data is robust and not the result of the model converging on the priors.

### Spatial patterns of genetic diversity

mtDNA. As found previously (Girman et al. 2001), genealogical analysis revealed that the mtDNA haplotypes were derived from two highly divergent clades (Eastern and Southern, Fig. 1; Table S4, Supporting information). The two clades showed strong geographical structuring, with the most easterly (Laikipia, Serengeti and Masai-Steppe) and southerly (Kruger and Lowveld) populations exhibiting exclusively Eastern and Southern mtDNA clade haplotypes, respectively. Our more comprehensive sampling highlighted that the E3 haplotype is not restricted to the Serengeti-Mara but is common to populations in Eastern Africa (except

<table>
<thead>
<tr>
<th>Population</th>
<th>n</th>
<th>Ne temporal method</th>
<th>Ne linkage disequilibrium</th>
<th>N0/Npop linkage disequilibrium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kruger all</td>
<td>87</td>
<td>21.8 (11.6–43.5)</td>
<td>19.6 (17.3–22.2)</td>
<td>0.05</td>
</tr>
<tr>
<td>Old</td>
<td>67</td>
<td></td>
<td>10.2 (8.4–12.7)</td>
<td>0.10</td>
</tr>
<tr>
<td>Recent</td>
<td>20</td>
<td></td>
<td>7.8 (6.1–10.5)</td>
<td>0.06</td>
</tr>
<tr>
<td>Lowveld</td>
<td>14</td>
<td></td>
<td>10.4 (7.7–15.1)</td>
<td>0.05</td>
</tr>
<tr>
<td>Okavango all</td>
<td>54</td>
<td>28.3 (13.1–92.7)</td>
<td>16.6 (14.7–18.8)</td>
<td>0.09</td>
</tr>
<tr>
<td>Old</td>
<td>12</td>
<td></td>
<td>12.4 (9.9–15.9)</td>
<td>0.05</td>
</tr>
<tr>
<td>Recent</td>
<td>42</td>
<td></td>
<td>20.5 (16.2–27.0)</td>
<td>0.21</td>
</tr>
<tr>
<td>Selous</td>
<td>23</td>
<td></td>
<td>12.6 (10.4–15.7)</td>
<td>0.25</td>
</tr>
<tr>
<td>Hwange</td>
<td>14</td>
<td></td>
<td>3.0 (2.6–3.6)</td>
<td>0.02</td>
</tr>
<tr>
<td>Masai-Steppe</td>
<td>32</td>
<td></td>
<td>115 (9.8–13.6)</td>
<td>NA</td>
</tr>
<tr>
<td>Okavango all</td>
<td>33</td>
<td>21.5 (6.8–20.8)</td>
<td>21.3 (18.7–24.8)</td>
<td>0.07</td>
</tr>
</tbody>
</table>

*Ne/Npop estimates are provided for the linkage disequilibrium method; Npop is the number of animals in the monitoring area, and values are taken from Table S1.

<table>
<thead>
<tr>
<th>Population</th>
<th>M ratio</th>
<th>P values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0 = 0.2</td>
</tr>
<tr>
<td>Kruger</td>
<td>87</td>
<td>0.677</td>
</tr>
<tr>
<td>Okavango</td>
<td>54</td>
<td>0.754</td>
</tr>
<tr>
<td>Masai-Steppe</td>
<td>32</td>
<td>0.636</td>
</tr>
<tr>
<td>Serengeti</td>
<td>33</td>
<td>0.674</td>
</tr>
<tr>
<td>Laikipia</td>
<td>67</td>
<td>0.741</td>
</tr>
</tbody>
</table>

Shown is the M ratio statistic and P values for 0 = 0.2, 0.6, 1.2 and 2, which correspond to pre-bottleneck population sizes of 100, 300, 600 and 1000. Underlined P values indicate where the M ratio is significantly lower than values expected under mutation-drift equilibrium.

*Kruger, Okavango and Serengeti are the combined temporal sample set.
Selous), and we found new private haplotypes in both Selous (S4) and Okavango (S5). A hierarchical AMOVA of mtDNA haplotypes indicated significant population structure, with 46% of variation being apportioned between four regions: (i) Southern cluster (Kruger, Lowveld); (ii) South Western cluster (Okavango, Hwange); (iii) Central cluster (Selous); and (iv) Eastern cluster (Masai-Steppe, Serengeti-Mara and 4 = Laikipia) (Table S7, Supporting information), which corresponds with geographical expectations (Fig. 1b).

Microsatellites. Hierarchical AMOVA detected significant differentiation between populations (P < 0.001). However, as the majority of microsatellite variation (~85%) was apportioned within populations, AMOVA was not informative in identifying population groupings (Table S7, Supporting information). Bayesian clustering analysis in STRUCTURE also indicated strong population subdivision, with the ΔK statistic (Evanno et al. 2005) identifying two peaks at K = 4 and K = 7. The former represented the deepest level of population subdivision, whereas K = 7 represented finer-scale population structuring (Figs 2b and S2, Supporting information). The four clusters at K = 4 were 1 = Kruger, 2 = Okavango-Hwange, (Lowveld was admixed with 1&2), 3 = Selous, Masai-Steppe, Serengeti-Mara and 4 = Laikipia. At K = 7, every sampled population was a distinct cluster except for Okavango and Hwange, which appear as a single genetic population (Fig. 2b). Only at K = 4 and 7 were clustering results consistent between replicates.

Principle coordinates analysis also indicated strong structuring, but not complete isolation, of wild dog populations (Fig. 2a). PC1 and PC2 appear to correspond to northsouth and eastwest axes and accounted for considerable variation (28% and 18%, respectively). Overall, there was a striking concordance between population medians and geographical sampling locations (Fig. 2a). These results are consistent with Mantel tests, which found significant evidence of isolation by distance (r = 0.417, P < 0.001).

Posterior probability tests conducted in STRUCTURE indicated that the majority of individuals originated from the population from which they were sampled as 98% of individuals were assigned to their sampled population with more than 70% probability (316/321). Five putative migrants identified were from Okavango (1), Masai-Steppe (1), Kruger (1) and Selous (2).

Table 5 Estimates of ancestral (N1) and current (N0) effective population sizes and time in years since decline (T), using MSVAR

<table>
<thead>
<tr>
<th>Population</th>
<th>N1 (95-5%)</th>
<th>N0 (95-5%)</th>
<th>T (95-5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kruger</td>
<td>865 (225–3278)</td>
<td>1 (0.1–7)</td>
<td>10 (1–82)</td>
</tr>
<tr>
<td>Okavango</td>
<td>938 (284–3031)</td>
<td>2 (0.1–24)</td>
<td>14 (1–154)</td>
</tr>
<tr>
<td>Selous</td>
<td>670 (181–2439)</td>
<td>1 (0.1–20)</td>
<td>9 (1–123)</td>
</tr>
</tbody>
</table>

Results show the median value and 5-95% quantiles.

Fig. 2 Clustering analyses of African wild dogs based on 10 microsatellite loci. Colours correspond to sampling location as displayed in panel b. (a) Principle coordinates analysis (PCoA) computed in GENALEX (Peakall & Smouse 2006). X symbols represent individuals, and large circles represent population medians. The axes have been rotated to reflect the resemblance between the PCoA and geographical sampling location. PC 1 appears to represent a northsouth axis and PC 2 an eastwest axis. (b) Bayesian clustering analyses computed in STRUCTURE (Pritchard et al. 2000). Shown is the uppermost level of structuring (K = 7) as indicated by the ΔK statistic (Fig. S2, Supporting information). Columns are individuals, with the proportion of an individual’s genotype assigned to each cluster (K) denoted by different colours.
case, the source population is probably an unsampled population, as posterior probability values for the putative migrants were <70% for the populations we sampled.

**DLA-DRB1.** At DLA-DRB1, seven of 19 alleles were private alleles (Table 2). Fisher’s exact tests showed that DLA-DRB1 allele frequencies differed significantly between all populations (P < 0.01). Hierarchical AMOVA detected significant differentiation between populations (P < 0.001), but was not informative in identifying population grouping as ~80% DLA-DRB1 variation was apportioned within populations (Table S7, Supporting information). DLA-DRB1 was found to show evidence of significant isolation by distance based on Mantel tests (r = 0.506, P = 0.01).

**Temporal patterns of genetic diversity**

In Okavango, we detected no significant temporal changes in genetic diversity at microsatellite loci or DLA-DRB1 (msats, paired T test, \(H_E-T = -1.09\), \(P = 0.304\); \(H_O-T = -2.14\), \(P = 0.058\); \(R_S-T = -2.29\), \(P = 0.053\); DLA-DRB1, \(H_E = 0.86\) and 0.85; \(H_O = 0.84\) and 0.81, \(R_S = 8.3\) and 8.0). Furthermore, there was little change in allele frequencies between old and recent samples (Table S8, Supporting information, msat \(D_{est} = 0.03\); DLA-DRB1 \(D_{est} = 0.06\), Fisher’s exact test, \(P = 0.252\)). In Kruger, \(H_E\) and \(R_S\) were temporally stable at both microsatellite loci and DLA-DRB1 (Table 1: msats, paired T test, \(H_E-T = -0.78\), \(P = 0.457\), \(R_S-T = 0.42\), \(P = 0.684\); DLA-DRB1, \(R_S = 5.1\) and 5.5; \(H_E = 0.76\) and 0.66) as were allele frequencies (Table S8, Supporting information, msat \(D_{est} = 0.02\); DLA-DRB1 \(D_{est} = 0.07\), Fisher’s exact test, \(P = 0.252\)). However, there was a significant reduction in \(H_O\) at both microsatellite (15.4%; paired T test, \(T = 2.24\), \(P < 0.05\)) and DLA-DRB1 loci (13.1%) in Kruger, which is indicative of inbreeding. In the Serengeti, there was no significant change in microsatellite diversity (paired T test, \(H_E\), \(T = 1.60\), \(P = 0.145\); \(H_O\), \(T = 0.72\), \(P = 0.486\); \(R_S\), \(T = 1.70\), \(P = 0.123\)). However, at DLA-DRB1, \(H_O\) and \(R_S\) increased, from 67% to 93% and 4.8 to 6, respectively, and there were notable changes in allele frequencies (Table 2, Table S8, Supporting information, msat \(D_{est} = 0.14\); DLA-DRB1 \(D_{est} = 0.68\), Fisher’s exact test \(P < 0.00001\)), which could reflect sampling effects, immigration or selection.

**Tests for selection**

Analyses using FEL, which tests for both positive and purifying selection at individual codons, indicated no evidence of positive selection at DLA-DRB1. However, purifying selection was detected at two sites: codon 56, which is a PBR site (\(d_N/d_S = 0.001, P = 0.038\)), and codon 72, which is adjacent to a PBR site (\(d_N/d_S = 0.000, P = 0.003\)).

Genetic differentiation (\(D_{est}\)) was relatively high (>0.2) amongst all populations at both microsatellite and DLA-DRB1 markers, except between Okavango and Hwange (Table S8, Supporting information; DLA-DRB1 \(D_{est} = 0.722\); Msats, \(D_{est} = 0.419\)), and across populations, genetic differentiation at DLA-DRB1 and microsatellite loci were significantly correlated. However, the observed increase in genetic divergence with geographical distance noted above was significantly stronger at DLA-DRB1 than microsatellites (ANCOVA, \(P < 0.05\), Fig. 3). Moreover, the relationship between DLA-DRB1 and geographical distance remained significant even after controlling for microsatellite differentiation (partial Mantel test; \(r = 0.416\), \(P = 0.0075\)), whereas the relationship between microsatellite differentiation and geographical distance was not significant when DLA-DRB1 differentiation was controlled for (partial Mantel test, Bonferroni-corrected \(P = 0.01\); \(r = 0.284\), \(P = 0.0132\)). These results suggest that genetic differentiation at the DLA-DRB1 is greater than expected under neutrality.

**Discussion**

**Demographic history**

Despite sampling some of the largest wild dog populations, our estimates of current \(N_e\) were small across all populations (\(N_e < 30\); Table 3) and, with the exception...
of Selous, $N_e/N$ estimates were ≤0.1. Population declines, inbreeding and reduced gene flow are probably the primary cause of the small $N_e$. However, the cooperatively structured breeding system of wild dogs will also have reduced $N_e$ (Creel et al. 2004). Overall, our $N_e$ estimates are similar to other cooperatively breeding endangered canids (e.g. Ethiopian wolves, *Canis simensis*, Bale Mountains, $N_e$ 10–25, Randall et al. 2010).

We found evidence of recent bottlenecks in all populations ($M$ ratio test, $M = 0.64–0.75$, $P < 0.05$; Table 4). Furthermore, Bayesian coalescent simulations of microsatellite data found evidence for a large demographic reduction within the last 100 years (Table 5). The timing of this population decrease is concurrent with extensive anthropogenic habitat loss and persecution during the 20th century (Woodroffe et al. 1997) and is consistent with other endangered species such as the orangutan (*Pongo pygmaeus*, Goossens et al. 2006). However, it differs from some other African mammals where major declines correspond with mid-Holocene climatic change several thousands of years ago as found in the African elephant (*Loxodonta africana*, Okello et al. 2008), African buffalo (*Syncerus caffer*, Heller et al. 2008) and Walia Ibex (*Capra walie*, Gebremedhin et al. 2009).

In addition, the mtDNA data may suggest older declines prior to that predicted by microsatellites. Mitochondrial DNA haplotypes showed evidence of strong eastern-southern regional clustering. A similar phylogenetic pattern has been observed in many other African mammals (reviewed in, Hewitt 2004) and is thought to represent contraction to refugia during the climatic transitions in the Pleistocene/Holocene boundary. Thus, the phylogenetic pattern in wild dogs may reflect a scenario of extinction of Eastern African populations and subsequent recolonization from Western or Central Africa, with secondary migrations between Eastern and Southern Africa, as suggested by Girman et al. (2001).

**Spatial and temporal patterns of genetic diversity**

Our results suggest that, despite the dispersal ability of wild dogs (>250 km), genetic isolation associated with fragmentation of populations has resulted in strong population subdivision. Each sampled population was found to be a separate genetic cluster (Fig. 2b), except for Hwange and Okavango; these two populations probably have high genetic connectivity because they are geographically proximate and occupy the largest continuous area supporting wild dogs. Further, there was a strong pattern of isolation with distance based on microsatellites (Mantel tests; Fig. 3), which is particularly evident in the PCoA (Fig. 2a). It seems unlikely that gaps in our sampling regime have resulted in an underestimation of connectivity between wild dog populations, as many of the gaps, reflect geographical areas that do not currently support wild dogs (Fig. 1). Furthermore, we detected few migrants and realized gene flow resulting from migration in wild dogs is probably reduced because of their cooperative breeding system and the high mortality of dispersers (Creel & Creel 2002).

We consistently found higher diversity across both neutral and DLA-DRB1 markers in Okavango, Hwange and Selous (Table 1), which are the monitoring areas located in some of the largest contiguous and most demographically stable wild dog populations ($n = 1300–2500$) (Woodroffe et al. 1997, 2004b). In contrast, populations that had recently experienced recolonization events following extirpation consistently showed lower microsatellite and MHC diversity (Lowveld and Laikipia; Table 1), which contrasts with the expectation that selection can retain adaptive genetic variation despite strong genetic drift, as found in African buffalos (Wenink et al. 1998) and Island foxes (Aguilar et al. 2004). However, our result is consistent with findings in some other species that have also undergone declines (reviewed in, Radwan et al. 2010).

The results from our temporally sampled populations were highly variable. Whilst diversity estimates were stable across sampling periods in Okavango, we detected a large decline in $H_0$ relative to $H_E$ at both microsatellite and DLA-DRB1 loci in Kruger between 1991–1995 and 2007 (Table 1). This relative decline in $H_0$ probably reflects increased inbreeding and genetic drift associated with a demographic decline of >70% within Kruger National Park between 1995 ($n = 434$) and 2005 ($n = 120$; EWT 2009). In the Serengeti-Mara, there were no significant temporal changes at microsatellite loci (Table 1). However, at DLA-DRB1, there were notable changes in allelic composition as well as an increase in $H_0$ and $R_S$ (Tables 1 and 2). Here also, the genetic changes were correlated with a demographic event, in this case related to a disease outbreak. As such, the differences at the MHC could be the result of changes in selective pressures, as observed in other systems (e.g. Fraser et al. 2010). Alternatively, they may reflect sampling effects and/or the effects of the immigration of a small number of migrants from other populations after the initial sampling period (Marsden et al. 2011). Overall, our temporal data indicate that demographic declines, stochastic processes and selection may contribute to temporal changes in genetic variation within populations.

Other studies that have assessed genetic variation both spatially and temporally have also found that MHC and microsatellite allele frequencies are temporally stable in some populations, but not others (e.g. Miller et al. 2001; Oliver et al. 2009; Fraser et al. 2010),
indicating that selection can be spatially and temporally variable (Hedrick 2002; Consuegra et al. 2011). Similar to the Serengeti-Mara, temporal studies of great reed warblers (Westerdahl et al. 2004) and Soay sheep (Charbonnel & Pemberton 2005) found higher temporal divergence at MHC loci than at microsatellite loci, which was attributed to temporal variation in parasite selective pressures. Interestingly, Soay sheep exhibited no significant changes in microsatellite or MHC divergence associated with 12 years of population fluctuations from ~2000 to ~600 animals (Charbonnel & Pemberton 2005). This directly contrasts with Kruger which may have been more vulnerable to genetic drift because it was reduced to smaller population sizes. A dramatic bottleneck (100 to 11) in a small population of song sparrows (Melospiza melodia) resulted in a significant decline in microsatellite diversity (MHC diversity was not measured; Keller et al. 2001), which corresponds more with our findings in Kruger. In the song sparrows, low levels of immigration contributed to demographic and genetic recovery to pre-bottleneck levels within 3 years (Keller et al. 2001). Similar genetic rescue of MHC and microsatellite diversity by migration has been reported elsewhere (e.g. African buffalo, Wenink et al. 1998; and Grey wolves, Canis lupus, Vilà et al. 2003; Seddon & Ellegren 2004). It is encouraging, therefore, that wild dogs are such a highly mobile species (Fuller et al. 1992). Nonetheless, genetic recovery in Kruger may be delayed by both the absence of demographic recovery to date (which will accentuate the impact of the bottleneck) and the apparent lack of genetic connectivity to nearby populations (Fig. 2).

Tests for selection

We found no significant evidence of historical positive selection at the DLA-DRB1 locus but did detect evidence of historical purifying selection at two codons located in or adjacent to the PBR. In terms of contemporary selection, we found patterns of DLA-DRB1 and microsatellite differentiation were correlated, indicating that current patterns of diversity are strongly influenced by genetic drift. However, overall, genetic differentiation and isolation by distance were more pronounced for DLA-DRB1 allelic data than for microsatellite data (Fig. 3b), and partial Mantel tests showed that patterns of genetic differentiation at DLA-DRB1 remained significant even after controlling for neutral microsatellite variation ($r = 0.4156$, $P = 0.0075$). These data show that genetic differentiation at DLA-DRB1 is larger than expected under neutrality, which is indicative of contemporary selection at DLA-DRB1. The higher divergence at DLA-DRB1 relative to the microsatellites has been observed in other taxa (e.g. atlantic salmon, Landry & Bernatchez 2001; great snipe, Gallinago media, Ekblom et al. 2007 and water voles Bryja et al. 2007) and is consistent with diversifying, rather than balancing, selection acting on DLA-DRB1 and may indicate local adaptation (Landry & Bernatchez 2001; Bryja et al. 2007; Ekblom et al. 2007).

Despite finding a number of indicators of selection at DLA-DRB1, like many other MHC studies, the overall footprint of selection in our data was low, which may reflect an erosion of historical signals owing to strong genetic drift in small populations (Campos et al. 2006; Oliver et al. 2009), a lack of statistical power where sample sizes are small, or the transient nature of selective events (reviewed in, Garrigan & Hedrick 2003). Alternatively, given that we assessed a single MHC locus and that patterns of selection can differ amongst MHC loci (Bryja et al. 2007; Evans et al. 2010; Consuegra et al. 2011), it is also possible that selection is stronger at other regions of the MHC such as MHC class I loci, which are involved in recognition of intracellular viruses, rather than extracellular pathogens and parasites recognised by class II loci.

Conservation implications

We have shown that extensive population declines, population subdivision and genetic isolation of wild dog populations are associated with habitat fragmentation and loss. Where continuous habitat remains (Okavango-Hwange), gene flow connects wild dogs over considerable distances (~400 km), resulting in higher genetic diversity and few indicators of inbreeding (Table 1). This result highlights the conservation value of habitat connectivity. Nonetheless, we also show that habitat connectivity does not guarantee genetic connectivity. We found monitoring areas in Southern Kruger and Lowveld to be highly differentiated ($D_{est} = 0.28–0.37$; Table S8, Supporting information) despite continuous habitat between them (~500 km apart). There are no clear topographic barriers to migration, and wild dogs are able to disperse long distances (>250 km, Fuller et al. 1992; McNutt 1996). However, low and declining wild dog densities in Northern Kruger, combined with fencing of Kruger National Park in the 1970s, may have contributed to separation. The establishment of the Greater Limpopo Transfrontier Park and removal of fences may help to re-establish connectivity.

Our temporal sampling has shown that genetic diversity is not always stable in small or declining wild dog populations and that these changes result from both population declines and changes in evolutionary forces. As such, our data indicate that genetic variation in a population may not be fully reflected in samples collected from a single time period. Therefore, continuous monitoring
through long-term field projects is important to detect changes in variation, especially considering that such changes can occur very quickly in endangered species as they are inherently unstable. Kruger National Park, for example, was previously described as a ‘stronghold’ for wild dogs (Woodroffe et al. 1997), yet suffered a 70% population decline in recent years, which significantly reduced observed heterozygosity (Table 1)

As a species, wild dogs have less MHC variation than other wolf-like canids (reviewed in Marsden et al. 2009), and we also found evidence for the loss of DLA-DRB1 diversity within some populations over time. Whether reduced MHC variation will impact population survival is unclear, as results from other species are equivocal. However, examples of species with depleted MHC variation showing increased disease susceptibility suggests that our findings could be a cause for concern (reviewed in, Radwan et al. 2010).

Overall, our study has shown that the demographic history is a very strong factor determining spatial and temporal patterns of both neutral and adaptive genetic variation in wild dogs as effective population size influences the relative strength of selection, rates of genetic drift, as well as inbreeding and extinction probability. Nonetheless, evidence of selection at the MHC suggests selection has not been completely compromised by demographic declines in this taxon. Our results highlight the need to enhance connectivity between wild dog populations and increase effective population size such that genetic variation can be preserved. Of special concern is the Kruger National Park population that although protected shows evidence of demographic declines and inbreeding. Efforts to enhance the connections between this population and elsewhere are needed.

Acknowledgements

We would like to thank Harriet Mostert-Davies, Paulette Bloomer and Janet Edwards for provision of recent Kruger National Park samples, Aart Viseé for Masai-Steppe samples, Bruce Fivaz for Malilangwe samples and Mark Swarner for providing carcass samples. We thank Bridgett vonHoldt, Daniel Greenfield, John Pollinger and Aileen Adam for assistance with laboratory work and analyses, Lorna Kennedy for training in MHC typing methodologies and three anonymous reviewers for helpful comments on this manuscript. This research was funded by a joint CASE studentship (NER/S/A/2006/14139) between the Natural Environment Research Council and the Royal Zoological Society of Scotland-Edinburgh Zoo (BK) as well as support from the US National Science Foundation (RKW).

References


Fraser BA, Ramnarine IW, Neff BD (2010) Temporal variation at the MHC class IIb in wild populations of the guppy (Poecilia reticulata). Evolution, 64, 2086–2096.


Data accessibility

DNA sequences are available from GenBank (MHC accessions JQ085963–JQ085964; mtDNA accessions JQ282681–JQ282690), and individual level MHC, mtDNA and microsatellite data have been deposited to DRYAD doi: 10.506 1/dryad.86 p82f43.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Wild dog DLA-DRB1 alleles aligned to allele DLA-DRB1*0101.

Fig. S2 (a) Genetic structure of wild dog populations based on Bayesian clustering analyses (STRUCTURE) of samples at 10 microsatellite loci. Shown is the uppermost level of structuring (K = 4) and finer scale population structuring (K = 7) as indicated by the ΔK statistic (panel c). Columns are individuals, with the proportion of an individual’s genotype assigned to each cluster (k) denoted by different colours. Colours correspond with location map in Figure 5 and population three letter codes with Table 5. (b) Likelihood probability profile estimated from STRUCTURE 2.2 at K=10 showing the mean and variance at each K. (c) ΔK at each value of K, averaged across 10 replicates.

Table S1 Details of the origin of the wild dog samples used in this study including the name, country and region of the monitoring area, the sampling years, the number of animals in the monitoring area, number of animals and packs sampled, and monitoring area three letter abbreviations.

Table S2 Details about the source of samples.

Table S3 Information on microsatellite primers used in this study isolated from the domestic dog (see Breen et al. 2001; Guyon et al. 2003; Neff et al. 1999).

Table S4 Sample sizes (n) and distribution of mtDNA haplotypes across wild dog populations.

Table S5 Original (as submitted by Girman et al. 2001), corrected and new mtDNA haplotype sequences detected in African wild dogs.

Table S6 Official DLA-DRB1 names alleles given in the paper.

Table S7 Results from hierarchial analysis of molecular variance (AMOVA) computed in Arlequin v 3.11 (Excoffier et al. 2006).

Table S8 Pairwise bootstrapped D_{est} (Jost 2008) estimates of genetic differentiation between wild dog populations.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.