Evolutionary & Ecological Genetics of African Wild Dogs

Clare Diana Marsden

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy.

University of Glasgow
Faculty of Biomedical and Life Sciences
Division of Ecology and Evolutionary Biology

June 2010

Abstract

Loss of adaptive variation arising from population declines and fragmentation is a primary concern in conservation. However, many conservation programmes assess only neutral genetic variation. Whilst assessments of neutral variation are informative about demographic history, inbreeding and genetic structure, they do not provide information on adaptive variation. The Major Histocompatibility Complex (MHC) is a group of genes that has been extensively studied and are known to be important in effective immune responses. Given the threat posed by infectious diseases to wildlife, the MHC is increasingly being assessed in endangered species.

African wild dogs (*Lycaon pictus,* hereafter wild dog) are an endangered canid that has suffered extensive declines in the wild and now persist as small and fragmented populations totalling less than 8,000 individuals. The purpose of this study was to assess how neutral and MHC marker data genetic data can be used to assist conservation of this species. As such, I assessed sequence diversity across ~300bp of mitochondrial DNA, patterns of polymorphism and heterozygosity at 10 neutral microsatellite loci, compared to sequence variation and haplotype diversity at the MHC.

Wild dogs were found to be genetically depauperate at the MHC compared to other canids. Patterns of variation indicate a historical loss of variation, followed by more recent diversification. However, it was also shown that evolutionary history contributes to differences in diversity between species. The spatial and temporal structure of MHC diversity was found to be largely correlated with neutral markers, which may suggest that selection is unable to counter strong genetic drift in such small populations. Overall, genetic diversity of both neutral and MHC markers appeared to be largely determined by demographic stability and size of populations. Habitat fragmentation and loss were associated with genetic isolation of wild dog populations, which showed strong structuring. However, the barriers to, or corridors for, dispersal of wild dogs were not always clear. The European captive breeding population was found to have comparable diversity metrics to wild populations, and was found to contain a large proportion of the MHC variation from the Southern African populations from which they were originally sourced. Careful genetic management is now required to correct the severe over- and under-representation of some founder lineages in this captive population to reduce inbreeding and loss of genetic variation.

Table of contents

List of Tables

List of Figures

List of Accompanying Material

A full size (A1) pedigree for Chapter 6 (Figure 6.1) is provided as accompanying material.

Acknowledgement

Firstly, and most of all, I would like to thank my PhD supervisor, Barbara Mable; I am hugely grateful for all of the time and energy you have invested in me and also for your patience – alas I fear I required much of that. Thank you also for encouraging me to leave the office occasionally to join you on various hiking adventures, which certainly were the highlight of Glasgow. I also need to thank my co-supervisors: Dan Haydon for help with project management, statistics and multiple near death experiences, and Lorna Kennedy at the University of Manchester for training in MHC typing. I am also very grateful to Sarah Cleaveland and Greg Rasmussen for providing much useful advice with various wild dog aspects of the project. I would like to thank Rob Thomas, my CASE partner at The Royal Zoological Society of Scotland, who provided much funding and assistance. I am hugely grateful to Aileen Adam for much training and assistance in the lab, as well as providing significant chocolate supplies. To Liz Masden, I couldn't have done it without you mate, and Katie Prager, it was great to work with someone else who studies wild dogs (even if they deny it)! At UCLA, I would like to thank Bob Wayne, for allowing me to join his lab for 6 months and use his canid freezers, but also for his helpful perspective on manuscripts. Also at UCLA, I need to thank John Pollinger for his help with various evil genetic programmes, Klaus Koepfli for his unrivalled scientific passion and enthusiasm, and to Bridgett vonHoldt for just being so damn clever - you certainly raised the bar. Overall, Klaus, Bridgett, Daniel, James, Sergio and Rena, you made UCLA the most enjoyable and humorous 6 months of the PhD; I didn't realise North American's could do 'wit'.

I am indebted to Rosie Woodroffe for enabling me to tap into her expertise and knowledge of wild dogs, and for facilitating collaborations with the wild dog network. I am also hugely grateful to researchers from many wild dog field projects that entrusted their samples with me, as well as providing much field information and help with interpretation of findings: Gus Mills (Old Kruger); Harriet Davies, Paulette Bloomer and Janet Edwards (Recent Kruger and metapopulation); R.Lines (NE Namibia); F.Stander (NE Namibia); Tico McNutt (Old and Recent Okavango); Josh Ginsberg (Hwange); Greg Rasmussen (Hwange); Mark Swarner (Ghanzi), Rosemary Groom (Lowveld/Savé), Jean-Marc André (Sofala), Colleen and Keith Begg (Niassa), Scott Creel (Selous), Aart Viseé (Masai Steppe), Sarah Cleaveland (Old Serengeti), Emmanuel Masenga (Recent Serengeti), Pieter Kat (Masai-Mara), Rosie Woodroffe (Laikipia), A.Bastos, T.Matjila and H.Strydom (South African captive and metapopulation). Further samples were provided by Andrew Kitchener (museum specimens) and Arati Ivengar (Dhole samples), for which I am very grateful.

For the zoo research, I need to thank the studbook coordinator, Hanny Verberkmoes, who made it all possible. But overall, I need to thank all of the EEP Species Committee for African wild dogs for patiently teaching me the realities of captive breeding management: Wim and Hanny Verberkmoes (Kerkrade), Jens Lilleør (Aalborg), Stefan Stadler, (Frankfurt), Lennart Sunden (Kolmarden), Lars Versteege (Hilvarenbeek), Alexis Maillot (Amnéville), Hans van Weerd (Amsterdam). I am also hugely grateful to all of the zoos that collaborated with this research by providing samples and information: Aalborg zoo, Aalborg, Denmark; Artis zoo, Amsterdam, Netherlands; Attica Zoological Park, Spata, Greece; Beekse Bergen Safarai Park, Hilvarenbeek, Netherlands; Borås Djurpark Zoo, Alvsborg, Sweden; Bretagne Zoo Sarl, Pont-Scorff, France; Centre d'Etude Rech Zool Augeron, Lisieux, France; City of Belfast Zoo, Belfast, UK; Colchester zoo, Essex, UK; Ebeltoft zoo, Ebeltoft, Denmark; Edinburgh Zoo, Edinburgh, UK; Friguia Zoo, Hammam, Tunisia; Fondazione Bioparco di Roma, Rome, Italy; GaiaPark Kerkrade Zoo, Kerkrade, Netherlands; Kolmården Djurpark AB, Kolmården, Sweden; La Palmyre Zoo, Royan, France; Le Pal Parc Animalier, Dompierre-sur-Besbre, France; London zoo, London, UK; Munchener Tierpark; Hellabrunn, Muenchen, Germany; Parken Zoo Eskilstuna AB, Sodermanland, Sweden; Port Lymne Wild Animal Park, Hythe, UK; Quinta De Santo Inacio, Avintes, Portugal; Reserve Africaine De Sigean, Sigean, France; Rostock Zoologischer Garten, Rostock, Germany; Safari De Peaugres, Peaugres, France; Warsaw Zoological Garden, Warszawa, Poland; West Midland Safari & Leisure Park, Worcester, UK; Zoo Basel, Basel, Switzerland; Zoo Dortmund, Dormund, Germany; Zoo D'Amneville, Amneville, France; Zoo Dvůr Králové, Dvůr Králové nad Labem, Czech Republic; Zoo Duisburgh Ag, Duisburg, Germany; Zoological Center Tel Aviv, Ramat Gan, Israel; Zoological Society of Ireland-Dublin, Dublin, Ireland.

Author's declaration

I declare that the work recorded in this thesis is entirely my own, except where otherwise stated, and that it has not been submitted as part of a degree elsewhere. Much of this thesis has been produced in co-authorship, and my personal contribution to individual chapters is listed below.

Chapter 2: *Published as*: Marsden, C.D., Mable, B.K., Woodroffe, R., Rasmussen, G.S.A., Cleaveland, S., McNutt, J.W., Emmanuel, M., Thomas, R., Kennedy, L.J. (2009) **Highly endangered African wild dogs (Lycaon pictus) lack variation at the major histocompatibility complex**. Journal of Heredity 100: S45-65. Initial concept developed by BKM and CDM. Field sampling conducted by RW, GSAR, SC, JWM, EM, who also provided comments on the manuscript. LJK trained CDM in MHC typing and RSCA analysis. Study design, lab work, analyses of data and drafting of manuscript by CDM. Final draft was completed after detailed comments on manuscript from BKM and LJK.

Chapter 3: *Submitted to Molecular Ecology and currently in revision as*: Marsden, C.D., Wayne, R.K. Mable, B.K. **Puzzling persistance of African wild dogs in Serengeti.** *Molecular Ecology submitted*. S. Cleaveland proposed the project. Field samples provided by S. Cleaveland, M. Emmanuel, A. Visée, P. Kat, S. Creel, R. Woodroffe. CDM planned the study, conducted the lab work (except microsatellite PCRs, which were conducted by A. Adam) and analyses and drafted the first manuscript. The final draft was completed after detailed comments on manuscript by RKW, BKM, R. Woodroffe and S. Cleaveland.

Chapter 4: *In preparation for submission to Molecular Ecology as:* Marsden, C.D., Wayne, R.K. Woodroffe, R., Mills, M.G.L, McNutt, J.W. Creel, S., Groom, R., Masenga, E.H., Cleaveland, S., Kat, P., Rasmussen, G.S.A., Ginsberg, J. Lines, R., André, J-M., Begg, C., Mable, B.K.. **Demographic processes determine patterns of genetic variation across African wild dog populations**. Initial concept developed by BKM and CDM. RW, MGLM, JWM, SC, RG, EHM, SC, PK, GSAR, JG, JMA, CB, RKW, provided samples and field information, and commented on results. Study design, lab work (except microsatellite PCR's, which were conducted by A. Adam), analyses of data and drafting of manuscript by CDM. BKM and RKW provided advice with analyses. Final draft completed after detailed comments from BKM.

Chapter 5: *In preparation for submission to Molecular Phylogenetics and Evolution as*; Marsden, C.D., Mable, B.M, Koepfli, K-P., Kennedy, L.K., Wayne, R.K. **The importance of considering phylogenetic relatedness when interpreting patterns of**

MHC polymorphism in endangered species. CDM proposed and designed the study, conducted lab work, analyses of data and drafting of manuscript. Samples were provided by RKW and KPK. LJK provided Ethiopian wolf data. BKM provided advice with analyses and detailed comments on manuscript.

Chapter 6: *In preparation for submission for Conservation Genetics as;* Marsden, C.D., Wayne, R.K., Verberkmoes, H., Thomas, R., Mable, B.K. **Combining studbook, neutral markers and MHC data for an assessment of the genetic status of the European African wild dog captive breeding programme.** RT proposed the research. CDM developed the concept and designed the study. HV and CDM facilitated a sampling programme in zoos. Individual zoos collected the samples. South African zoo samples were provided through A. Bastos and RKW. CDM conducted lab work (except microsatellite PCR's, which were conducted by A. Adam), analysis of data and drafting of manuscript. HV provided the studbook and general information. BKM provided extensive comments on manuscript. The management suggestions were devised in Kerkrade in May 2008 with the EEP Species Committee for African wild dogs and have now been implemented and are in effect. The EEP committee consisted of Wim and Hanny Verberkmoes (Kerkrade), Jens Lilleør (Aalborg), Stefan Stadler, (Frankfurt), Lennart Sunden (Kolmarden), Lars Versteege (Hilvarenbeek), Alexis Maillot (Amnéville), Hans van Weerd (Amsterdam).

Chapter 1: General introduction

1.1 The need for conservation

Habitat destruction, unsustainable harvesting, persecution and environmental pollution have resulted in unprecedented rates of species decline and extinction (reviewed in Lande 1988; Stokstad 2005). The proportion of species (36%) currently threatened with extinction present a clear signal that effective conservation strategies are urgently required to prevent widespread extinctions (IUCN 2010; Wall et al. 2001).

Conservation biology is a field that draws from a variety of disciplines, including veterinary medicine, wildlife biology and ecology (Frankham et al. 2002; O'Brien 1994). Relatively speaking, conservation genetics is a new aspect of conservation biology associated with the use of genetic theory and genetic data to improve conservation of species. For example, genetic data have been used to help understand species biology, resolve taxonomy and identify management units (Frankham et al. 2002; Palsboll et al. 2007). Although many conservation programmes now incorporate conservation genetics, it is not routinely implemented for most systems (Amos & Balmford 2001; Moran 2002). However, consideration of genetic threats and the use of genetic information in these situations is vital. Firstly, because most species requiring conservation typically persist in small and isolated populations and thus are at high risk of genetic threats. Secondly, because implementation of conservation efforts can increase genetic threats (Ellstrand & Elam 1993; Frankham et al. 2002; Lacy 1997; Moran 2002).

1.2 Conservation genetics and genetic threats

Where populations become small and isolated, they are at risk to a number of genetic threats including inbreeding depression, genetic drift and loss of genetic variation. The occurrence of these genetic threats are dependent on effective population size (N_e), which is defined as "the number of breeding individuals in an idealised population that would show the same amount of dispersion of allele frequencies under random genetic drift or the same amount of inbreeding as the population under consideration" (Wright 1931). Since populations rarely meet the assumptions of an idealised populations (stable population sizes; equal sex ratios; equal reproductive success between individuals; non overlapping generations), census population sizes are typically much smaller than N_e (Frankham 1995a; Frankham et al. 2002).

Inbreeding depression

Inbreeding results from the mating of related individuals (Caughley 1994; Ellstrand & Elam 1993). Genetic theory suggests that inbreeding results in inbreeding depression, that is, general reductions in fitness such as reduced disease resistance and reproductive output, thus making populations/species more vulnerable to extinction (Charlesworth & Charlesworth 1999; Edmands 2007; Keller & Waller 2002). Two mechanisitic explainations have been proposed for inbreeding depression; overdominance and partial dominance (Charlesworth & Charlesworth 1999; Edmands 2007; Keller & Waller 2002). Overdominance is associated with the loss of heterzygote advantage occurring as a result of increased homozygosity that occurs as a consequence of inbreeding. Partial dominance is based on the expectation that inbreeding results in a higher chance that an individual will inherit (from its parents) two copies of the same allele that are identical by descent, thus increasing expression of random deleterious mutations/alleles (Charlesworth & Charlesworth 1999; Edmands 2007; Keller & Waller 2002). The probability of this occurring is known as the inbreeding coefficient (F), which is measured by the extent of loss of heterozygosity (i.e. loss of allelic diversity) relative to Hardy-Weinberg expectations, that is, for a population which is infinitely large, randomly mating, with no gene flow, mutation or selection (Hartl & Clark 2007). Genetic theory has shown that inbreeding increases according to N_e ; $\Delta F = 1/2\ N_e$ (Barrett & Kohn 1991; Frankham 1996; Wright 1931). Inbreeding depression has been demonstrated in a number of species. For example, a small population (N=40) of adders (*Vipera berus*) that became isolated in Sweden exhibited high and atypical levels of stillborn and malformed offspring indicative of inbreeding depression, which resulted in population declines (Madsen 1996). However, the introduction of 20 adult males from another population reduced the frequency of abnormalities and drastically increased recruitment. The transferring of

individuals between populations raises the subject of outbreeding depression. Outbreeding depression is caused by the mating of individuals from genetically divergent populations, resulting in offspring with reduced fitness because of the breakdown of co-adapted gene complexes or gene-environment interactions (Edmands 2007; Marshall & Spalton 2000). A severe example of this has been shown with salmonid fish, where hatchery stock has been used widely to augment declining natural populations. However, salmonid fish have strong local adaptation in migration behaviour and thus introgression between hatchery stock and wild born fish has resulted in lower return rates to natural breeding grounds and dramatic population declines (Allendorf & Waples 1996). Another example comes from Atlantic salmon (*Salmo salar*), where stock transfers from Baltic populations to East Atlantic populations led to the extinction of native stock in 26/212 Norwegian rivers (Johnsen & Jensen 1986). This was caused by differences in local adaptation to parasite communities; Baltic populations were resistant to a local monogenean parasite, whereas East Atlantic populations (i.e. Norway) had not been exposed and were susceptible. Thus, when the parasite was inadvertently transferred with the Baltic stock, it resulted in widespread declines to the Norwegian populations. Overall, the examples given here demonstrate that both inbreeding and outbreeding depression can increase vulnerability to extinction (Edmands 2007; Hedrick 2001; Marshall & Spalton 2000; Saccheri et al. 1998).

Genetic drift

Genetic drift causes random changes in allele frequencies due to the sampling effect of alleles transmitted between generations (Ellstrand & Elam 1993), which results in random fixation and/or loss of alleles (Lacy 1997; Lande 1988). Genetic theory states that the extent of genetic drift is determined by effective population size. In small populations, genetic drift replaces selection as the dominant evolutionary force (Wright 1931). As such, all alleles effectively behave neutrally, which increases the chance that deleterious alleles could become randomly fixed and beneficial alleles lost. Overall, genetic drift increases the rate of loss of genetic diversity; the fraction of neutral variation lost per generation = $1/(2N_e)$ (Barrett & Kohn 1991; Ellstrand & Elam 1993). Bottlenecks and founder effects are a form of genetic drift, and typically result in severe reductions in diversity. For example, the Mauritius Kestrel (*Falco punctatus*) suffered a bottleneck from hundreds to just one breeding pair, and this was associated with a large decline in genetic diversity (Groombridge et al. 2000).

Loss of genetic diversity

Loss of genetic diversity is inherent to population reductions and further exacerbated by the strength of genetic drift and inbreeding in small populations (Ellstrand

& Elam 1993; Higgins & Lynch 2001; Lacy 1997; Lande 1988). However, this is thought to be one of the greatest threats to the persistence of species (Lacy 1997) because evolutionary change is reliant on the presence of adaptive genetic variation (Briggs & Goldman 2006; Lavergne & Molofsky 2007; Reznick et al. 1997). For example, Bradshaw (1991) observed that resistance to heavy metals was only observed in genetically diverse populations of *Agrostis stolonifera*. Although evolution is most commonly associated with long time scales, it can occur over very short periods (reviewed in Stockwell et al. 2003). For example in Tasmanian devils (*Sarcophilus harrisii*), a shift from iteroparity to semelparity and dramatic increase in precocious sexual maturity have been observed within 10 years in response to high mortality arising from a lack of immunity to devil facial tumour disease (Jones et al. 2008). Rapid change is particularly important for endangered species, given that most instances of species extinction result from the inability of a species to respond to changes in their environment, such as new predators or diseases or changes in weather conditions (Lacy 1997).

Overall, the examples mentioned and many others (e.g. Antiguan racer snake, *Alsophis antiguae*, (Daltry et al. 2001); Florida panther, *Felis concolor coryi*, (Roelke et al. 1993); Tasmanian devil, ,(Hawkins et al. 2006), demonstrate that genetic factors can threaten species persistence. However, there are also examples of species showing survival despite low levels of genetic variation and high levels of inbreeding (e.g. Chatham Island Black robin, *Petroica traverse*, (Ardern & Lambert 1997; Miller & Lambert 2004); Northern elephant seal, *Mirounga angustirostris*, (Bonnell 1974). These varying results indicate that the relationship between extinction risk and genetic diversity and/or inbreeding is complex. However, they are not unexpected because the effects of bottlenecks are highly stochastic and depend on the speed, duration and size of the bottleneck (Bouzat 2010; Ejsmond & Radwan 2009). Furthermore, the effects of genetic drift and inbreeding accumulate over time (Frankham 1995b), and expression of deleterious alleles can be environmentally dependent (Barrett & Kohn 1991); for example, mortality of song sparrows (*Melospiza melodia*) on Mandarte island during severe weather conditions was much higher amongst inbred individuals than outbred ones (Keller 1998; Keller & Waller 2002). As such, populations that do not appear to be suffering genetic threats cannot be assumed to be unaffected, and thus continuous monitoring is important (Lacy 1997). Overall, there is a general consensus that genetic threats pose significant threats and that conservation genetics should be used to monitor and assess these. However, conservation genetics also provides a unique and powerful tool for addressing questions that can not be assessed using other approaches, such as elucidating geneflow between populations, determining parentage, investigating demographic history of populations and applying wildlife forensics (Frankham et al. 2002).

Conservation genetic studies typically rely on neutral genetic markers such as microsatellites and mitochondrial DNA (Ennos et al. 1997; Frankham et al. 2002; Lynch 1996). These are the most appropriate type of markers for addressing questions relating to population structure, inbreeding, relatedness and demographic history. However, the ability of neutral markers to reflect adaptive genetic diversity has been heavily criticised (Ennos et al. 1997; Frankham et al. 2002; Lynch 1996) because they are not subject to selection and therefore neutral markers are unlikely to be correlated with adaptive traits. The ability to measure adaptive genetic variation is vital to detect adaptive differences between individuals/populations and loss of evolutionary potential as well as to improve our understanding of the forces that govern adaptive variation so we can conserve it more effectively (Ennos et al. 1997; Lynch 1996; Moran 2002). As such there is clearly a need for assessments of loci directly under selection (Miller et al. 2001). As one of the most well understood adaptive loci (Bernatchez & Landry 2003; Miller et al. 2001), and with clear relevance to viability (Hoglund 2009; Piertney & Oliver 2006; Siddle et al. 2007), the major histocompatibility complex (MHC) is currently one of the best markers available to use as a proxy for adaptive genetic variation.

1.3 The Major Histocompatibility Complex (MHC)

Structure, function and diversity of the MHC

The MHC includes multiple highly polymorphic genes that code for a set of cell-surface molecules involved in the recognition of foreign antigens as part of the immune response (Klein 1980). The MHC consists of three gene regions (classes I, II and III). MHC III genes include complement proteins and some cytokines involved in immune pathways, but are not discussed further here. MHC class I and II molecules function in the recognition of intracellular protein antigens, such as viruses (class I), and extracellular protein antigens, such as bacteria and nematodes (class II) (Klein & Horejsi 1997; Klein & Sato 2000). MHC class I genes are expressed on all nucleated cells and are made up of a trans-membrane peptide known as the heavy chain and three extra-cellular domains (Jeffery & Bangham 2000), whereas MHC class II genes are mainly expressed on antigen-presenting cells such as B cells, dendritic cells and macrophages and consist of an α and β chain (Klein & Horejsi 1997; Watts 1997). Recognition of foreign antigens in MHC class I and II molecules occurs at the peptide binding residues (PBR), where peptides are loaded. If they are recognised, the MHC molecules then present the peptides to thymus-derived lymphocytes (T cells) which triggers an immune response (Klein et al. 2007) . Alleles at a MHC locus differ in their specificity in terms of what antigens they can bind to, although individual alleles are able to bind to a variety of peptides, and different alleles may be able to bind to the same peptides (Klein et al. 2007). Diversity at the MHC is thought to be important for diverse immune capabilities (Doherty & Zinkernagel 1975; Klein et al. 2007; Sommer et al. 2002), specifically functional diversity amongst alleles. However, determining functional diversity from sequence data presents a significant challenge. The divergent allele advantage hypothesis proposed by Wakeland (1990) predicts that highly divergent alleles are able to confer resistance to a wider range of pathogens, but single amino acid changes have been shown to be distinguish susceptible from resistant forms of proteins (e.g. Bryan et al. 2000). Furthermore, amongst MHC alleles, the ability to bind to different pathogens is thought to be largely attributable to variation in the functionally important PBR (Klein et al. 2007). Therefore, sequence diversity in the PBR is likely to be particularly important. Overall, whilst it will rarely be possible to test functional differences between alleles, it is important for diversity studies to consider divergence between alleles (Spurgin & Richardson 2010).

Structure of the MHC in canids

Due to its status as a model species, extensive research has been conducted on the domestic dog (*Canis familiaris*) MHC, known as the dog leukocyte antigen (DLA) (Kennedy et al. 2002a; Kennedy et al. 2002b; Kennedy et al. 1999; Kennedy et al. 2000; Wagner 2003; Wagner et al. 1999), enabling understanding of the genetic organisation and diversity of the DLA. Genomic structure analyses have shown that the DLA is split; ~3-Mbp of Class II, III and I genes are located on chromosome 12 and a further ~ 500 kbp of Class I genes and four further pseudogenes on chromosome 35 (Yuhki et al. 2007). The DLA class I region also consists of several genes, one of which (DLA-88) is known to be highly polymorphic (Wagner et al. 1999). The class III region has been less well characterised, but consists of a number of genes, including tumor necrosis factor (TNF) and lymphotoxin-alpha (LTA) (Yuhki et al. 2007). Overall, assessments of the DLA have primarily focused on four MHC class II loci; DLA-DRB1 (highly polymorphic), DLA-DQB1 (highly polymorphic), DLA-DQA (moderately polymorphic), and DLA-DRA (monomorphic), which together encode the two α and β chains of class II molecules. These genes have been found to be tightly linked and inherited as a haplotype (Kennedy 2007). Associations have been found between MHC class II haplotypes and a number of immune-mediated conditions in domestic dogs, including canine hypothyroidism (Kennedy et al. 2006c), haemolytic anaemia (Kennedy et al. 2006a), diabetes mellitus (Kennedy et al. 2006b) and visceral leishmaniasis (Quinnell et al. 2003). With the knowledge gained from domestic dog studies, it is possible to expand MHC research to other canids. Indeed, studies have already been conducted on the Grey wolf (*Canis lupus*) (Kennedy et al. 2007a; Seddon & Ellegren 2002), Ethiopian wolf (Canis simensis) (Kennedy *in prep*), Mexican wolf (*Canis lupus baileyi*) (Hedrick et al. 2000) and Red wolf (*Canis rufus*) (Hedrick et al. 2000).

Mechanisms driving diversity at the MHC

A defining feature of MHC genes is their high diversity (Garrigan & Hedrick 2003). This diversity is thought to be primarily the result of balancing selection, which is predicted to maintain a large number of medium frequency alleles, which reduces the likelihood that alleles will become fixed and overall increases levels of heterozygosity (Garrigan & Hedrick 2003; Muirhead 2001). Balancing selection is also thought to be responsible for the retention of ancestral polymorphism (trans-specific polymorphism), resulting in high sequence diversity amongst alleles within a species.

A large number of studies have shown associations between specific MHC alleles and susceptibility or resistance to specific diseases (reviewed in Hill 1998; Piertney & Oliver 2006). For example, Quinnell (2003) found that susceptibility to visceral

leishmaniasis in domestic dogs was associated with specific MHC alleles at the DRB locus, and a number of protective MHC haplotypes to malaria have been demonstrated in humans (reviewed in Hill 1998). Studies have also demonstrated that heterozyogous individuals (Oliver et al. 2009b) and individuals with highly divergent MHC alleles (Schwensow et al. 2010) have lower parasite burdens. Together, these data support the contention that the high levels of diversity at the MHC are maintained by some form of pathogen-driven balancing selection. Three mechanisms of pathogen-mediated selection have been proposed for the MHC: 1) heterozygote advantage (Doherty & Zinkernagel 1975); 2) negative frequency dependent selection (Slade & McCallum 1992); and 3) fluctuating selection (Hill 1991). Heterozygote advantage predicts that heterozygotes can respond to a wider range of pathogens than homozygotes, which have just one allele (over-dominance) (Sommer 2005). Overall, the higher fitness of heterozygotes results in a large number of MHC alleles being maintained in the population (Spurgin & Richardson 2010). Negative frequency dependent selection is based on the arms race between MHC alleles and pathogen resistance. Theory predicts that there should be strong pressure for pathogens to become resistant to the most common MHC alleles. As such, there should be less resistance against new or rare alleles and thus these should have a selective advantage and increase in frequency (Spurgin & Richardson 2010). Overall, changes in allele frequencies reduce the probability that alleles become fixed or lost, resulting in large numbers of alleles being maintained. Fluctuating selection is derived from the fact that pathogen pressures vary both spatially and temporally, which results in directional selection for different alleles in different locations, as well as changes in selection, and thus alleles, within an area over time (Hill 1991; Spurgin & Richardson 2010). This results in a large number of alleles being maintained across all populations. Many studies have tried to determine which of these three mechanisms of balancing selection are most important in a system, but this has been difficult because they are not exclusive (Spurgin & Richardson). For example, an excess of heterozygotes could be the result of either heterozygote advantage or negative frequency-dependent selection through selection for a specific allele in a heterozygote. Furthermore, changes in alleles over time could represent either directional selection or negative frequency dependent selection.

In addition to pathogen-mediated balancing selection, it has also been proposed that MHC diversity can be maintained by mate choice (reviewed in Milinski 2006). In some species, the same MHC molecules that function in immune responses have also been shown to function as odorants (Carroll et al. 2002; Eggert et al. 1998; Slev et al. 2006). As such, it is thought that the MHC might act as cue in mate choice to enable selection of mates with dissimilar or pathogen resistant MHC genotypes, or that the MHC may be used as a ‘marker of relatedness’ to avoid inbreeding (Arcaro & Eklund 1999; Landry & Bernatchez 2001; Penn & Potts 1998; Slev et al. 2006). Whilst this has been demonstrated in some species (reviewed in Bernatchez & Landry 2003; Piertney & Oliver

2006), the results have been highly variable, and more recent studies implicate the importance of other chemical signalling proteins (e.g. major urinary proteins in mice) (Hurst et al. 2001; Thom et al. 2008). Overall, the data from wild populations has proven limited, indicating that pathogen-driven selection is the primary selection mechanism at the MHC (Radwan et al. 2010).

Evidence of MHC selection in wild populations

There are a variety of tests available for detecting selection in wild populations which Garrigan (2003) distinguished according to the timescale at which they detect selection. Over historical time scales, evidence that balancing selection operates at the MHC has been demonstrated by an excess of non-synonymous variation ($d_N>d_S$) amongst MHC alleles, as well as the concentration of non-synonymous changes to the residues (PBR) intricately involved with the recognition of foreign antigens (Furlong & Yang 2008), i.e. the sites where amino acid changes are most likely to result in functional changes. However, Garrigan (2003) demonstrated that such signals of selection typically take a very long time to both appear and disappear (10,000 -100,000 years). As such, it is only possible to infer that selection has occurred, not when it occurred. It is also important to consider that the power of these tests is low when there are few sequences (Nozawa et al. 2009) or if strong genetic drift has eroded the signal of selection (Garrigan & Hedrick 2003). A related test is the McDonald-Kreitman test, which compares variation at synonymous and non-synonymous sites within species to divergence between species. If all nonsynonymous mutations are neutral, the ratio of non-synonymous to synonymous variation within species (d_N:d_S) should be equal to non-synonymous to synonymous variation within species between species (p_N:p_S), whereas if they are advantageous, the mutations should increase divergence between species, resulting in d_N:d_S > p_N:p_S (Egea et al. 2008). This is one of the most widely used tests for inferring selection; however, its application to the MHC is restricted by increased sharing of alleles or allelic lineages between species expected under balancing selection (trans-specific polymorphism), which decreases fixed differences between species which results in too many zero values to compute the test (Garrigan & Hedrick 2003). Another sequence-based test is the Tajima's D statistic, which assesses whether DNA sequences are evolving neutrally, or whether they are evolving under a non-random process (explained in more detail in Chapter 4). Where $D=0$, neutral evolution is indicated, whereas purifying selection is indicated where $D<0$, and balancing selection where $D>0$ (Hartl & Clark 2007). However, demographic factors, such as expansions and contractions affect D in a similar way to selection, and therefore it is not possible to conclusively disentangle selective from demographic events. Overall, one of the strongest indicators of balancing selection is evidence for trans-specific polymorphism, which is responsible for the high sequence divergence amongst alleles within a species (Klein 1980). Trans-specific polymorphism describes the phenomenon

whereby ancestral alleles/allelic lineages are retained due to strong balancing selection. To date, trans-specific polymorphism at the MHC has been most often inferred from the sharing of allelic lineages, as direct allele sharing is considerably less common. Amongst mammals it has only been observed in primates (Otting et al. 2002; Suárez et al. 2006), ungulates (Radwan et al. 2007), rodents (Cutrera & Lacey 2007), canids (Seddon & Ellegren 2002) and cetaceans (Xu et al. 2009), and rarely outside of genera (excluding DRA locus). In comparison to dN:dS, the signal of trans-specific polymorphism persists for longer; indeed Klein (1993) suggested the upper limit to be ~40 million years. However, it is worthy to note that evidence of trans-specific polymorphism is confounded by convergent evolution (Klein et al. 2007), although methods have been proposed to disentangle these two processes (e.g. Kriener et al. 2000; Yeager & Hughes 1999).

Selection in the recent past has been inferred by comparing patterns of genetic diversity at adaptive and neutral loci (Garrigan & Hedrick 2003). If selection pressures are uniform across populations, differentiation is expected to be lower at the MHC relative to neutral loci (i.e. due to balancing selection), whereas if selection differs between populations (directional selection), higher levels of divergence are expected at adaptive loci. Results from wild populations on this subject have been variable. For example, higher divergence at the MHC relative to microsatellites, which is indicative of directional selection, was detected in Atlantic salmon (*Salmo salar* (Landry & Bernatchez 2001), Great snipe (*Gallinago media*; (Ekblom et al. 2007) and Water voles (*Arvicola terrestris*; (Bryja 2007), whereas no difference was reported in brown trout (*Salmo trutta*; (Campos et al. 2006), Gila topminnow (*Poeciliopsis occidentalis*; (Hedrick et al. 2001b) and Bighorn sheep (*Ovis canadensis*; (Boyce et al. 1997), which indicates that demographic processes are the primary factor influencing MHC variation. Whilst these comparisons are widely used, it is important to consider that demographic history can confound the outcome (Garrigan & Hedrick 2003; Piertney & Oliver 2006); for example, random loss of alleles in small populations due to genetic drift would be expected to elevate differentiation between populations. It is also possible to test for selection in the recent past using the Ewen-Watterson test, which assesses allele frequency distributions to detect footprints of selection (Garrigan & Hedrick 2003). Under neutral expectations, it is expected that there will be one common allele, whilst all other alleles occur at low frequency. The Ewen-Watterson test compares the expected heterozygosity based on mutation-drift equilibrium with the expected heterozygosity based on the allele frequency distribution within a population (Garrigan & Hedrick 2003). A major drawback of the Ewen's Watterson test comes from the assumptions that populations are at equilibrium and of constant size, which are often not the case.

Lastly, selection in the current generation can be demonstrated by deviations from Hardy-Weinberg expectations or by disease-fitness associations, whereby susceptibility or

resistance to certain diseases is associated with specific MHC alleles (reviewed in Hill 1998; Piertney & Oliver 2006). Whilst these tests have the advantage of detecting current selection, they may not be able to detect weak selection, and are known to suffer low power where there are very high levels of heterozygosity or small sample sizes (Garrigan & Hedrick 2003).

Overall, a review of all of these tests by Garrigan (2003) found that selection is not detectible in every generation, population or evolutionary lineage, and attributed this to either variable selection pressures at the MHC or the lack of power of tests. As such, it was recommended that tests of selection for different timescales be applied.

MHC and conservation

The effects of balancing selection have been repeatedly demonstrated at the MHC. However, the ability of selection to maintain diversity is very sensitive to N_e (Richman 2000). Where N_e is small, genetic drift replaces selection as the dominant evolutionary force, which results in the random loss of variation (reviewed in Sommer 2005). As such, there is concern that levels of MHC diversity will be lower in endangered species and thus that they will potentially have increased susceptibility to disease (O'Brien & Evermann 1988). The best example of this comes from Tasmanian devils, where a catastrophic population crash has been linked to an inability to mount an immune response against an emerging disease as a direct result of a lack of MHC variation due to previous bottlenecks (Siddle et al. 2007). However, although many endangered species appear to have atypically low levels of variation (e.g. Galapagos penguin, *Spheniscus mendiculus*, (Bollmer et al. 2007), this is not always the case (reviewed in Radwan et al. 2010). For example, MHC diversity was found in the otherwise genetically monomorphic San Nicolas Island fox (*Urocyon littoralis dickeyi*) (Aguilar et al. 2004), and high sequence diversity has been found amongst small numbers of MHC alleles in other endangered species (e.g. Arabian oryx and Mexican wolf, (Hedrick 2003). This has led to the suggestion that selection may be able to maintain MHC diversity despite strong genetic drift. However, these varied results show that the relationship between endangered status and MHC diversity is complex, and highlights the need for research to improve understanding of the factors that influence MHC diversity in endangered species.

Overall, there are two clear roles for MHC data in endangered species research and conservation. Firstly, studies of MHC variation can be used to assess whether endangered species have lost diversity at these loci as a result of population declines or bottlenecks, and thus may potentially have increased disease susceptibility. This is increasingly important given the rising incidences of infectious diseases in endangered species (Daszak et al. 2000). Secondly, the MHC can act as a proxy of adaptive variation

so that it is possible to assess evidence of selection, elucidate local adaptation and understand differences in evolutionary potential between wild populations. In captivity, it also provides a method by which to assess conservation of adaptive diversity from wild populations maintained in zoos.

1.4 African wild dogs

African wild dogs (*Lycaon pictus*, hereafter wild dog) are distinguished from other canids by their unusual coat patterning of black, brown, white and caramel patches, their characteristically large ears and tall lean body (Creel & Creel 2002). Phylogenetic analyses show wild dogs to be a member of the wolf-like canid clade, that also includes Domestic dogs (*Canis familiaris*), Grey wolves (*Canis lupus*), Coyotes (*Canis latrans*), Golden jackals (*Canis aureus*), Ethiopian wolf (*Canis simensis*), Dholes (*Cuon alpinus*), Black-backed jackals (*Canis mesomelas*) and Side-striped jackals (*Canis adustus*) (Lindblad-Toh et al. 2005). However, they are phylogenetically distinct to all other extant canids, and belong to the monotypic genus, *Lycaon* (Girman et al. 1993).

Like many other canids (Macdonald & Sillero-Zubiri 2004), wild dogs have a cooperative breeding system. They live in packs averaging 5-15 adults (range 2-28) typically formed by the fusion of a group of closely related females, and a group of closely related males, which are unrelated to each other (Creel & Creel 2002). Within packs, normally only the alpha pair breed, although subordinate breeding of both sexes can occur (e.g. Creel & Creel 2002; e.g. Girman et al. 1997). However, all individuals help to raise the large litters (~10 pups, but up to 21 (Fuller 1992), by 'babysitting', regurgitating food and defending the young (Courchamp et al. 2000; Courchamp & Macdonald 2001). Overall, larger pack sizes are correlated with both higher hunting and reproductive success, with the Allee effect (i.e. inverse density dependence) detected in packs with less than 5 adults because of reduced cooperative interaction (Courchamp et al. 2000). In fact, the cooperative breeding system of wild dogs is considered almost obligate, as there is only one report of successful solitary breeding (Woodroffe et al. 2009). As such, offspring are commonly permanently recruited to the pack. However, others disperse at 18-24 months as single sex groups, with the bias in dispersal to a particular sex varying between populations (Creel & Creel 2002).

Historically, African wild dogs ranged across most of sub-Saharan Africa (Woodroffe et al. 2004b) but they now occupy just 7% of their former range (IUCN/SSC 2008, 2009; Woodroffe et al. 1997). In the wild, fewer than 8,000 individuals remain, scattered across a small number of fragmented populations (IUCN/SSC 2008, 2009), only nine of which are known to constitute more than 200 animals (Woodroffe et al. 1997; Woodroffe et al. 2004b). Like many other endangered species, extensive habitat loss has been a major factor in their demise. However, habitat loss is a particular problem for wild dogs due to their very large home ranges, which can be > 4,000km^2 (Lines *Pers comm*; (Woodroffe et al. 1997). As such, wild dogs suffer greatly from edge effects because their home ranges frequently overlap both protected and unprotected areas (Woodroffe &

Ginsberg 1998), increasing the risk of persecution, snaring, domestic dog diseases and road deaths which together account for >60% of wild dog mortality (Woodroffe et al. 2007a). Whilst habitat loss has clearly caused substantial declines in wild dogs, human persecution is thought to be the 'single most important' cause of decreasing numbers of wild dogs in the past century (Woodroffe & Ginsberg 1997). This persecution stems not from a threat to human life, rather, their persecution appears to be a response to their method of killing (disembowelment, which is typical for canids, rather than strangulation in felids). It also derives from their hunting success rates (~70%, range 39-85%) which are considerably greater that most other carnivores, and thus wild dogs are perceived to create a higher threat to ungulate populations, although in reality a large proportion of wild dog kills are taken by hyenas and lions (i.e. kleptoparasitism, (Courchamp & Macdonald 2001; Creel & Creel 1998). In some areas wild dogs are also perceived as a threat to livestock, although the data suggest that livestock depredation is restricted to areas where wild prey is significantly reduced (Woodroffe et al. 2005). Together, these factors have resulted in wild dogs being perceived as 'cruel, bloodthirsty killers,' that present a significant threat to wild game and livestock, and should be dealt with as vermin (Woodroffe et al. 2004b).

> *"Wild dogs hunt in packs, killing wantonly far more than they need for food, and by methods of utmost cruelty" (Bere 1956)*

> *"It will be an excellent day for African game and its preservation when means can be devised for [wild dogs'] complete destruction" (Maugham 1914)*

Indeed, between 1956-1975, there were 3404 wild dogs shot as part of a 'vermin control' programme in Zimbabwe (Childes 1988) and shooting of wild dogs was even endorsed in protected areas until as late as the mid 1980's in some countries (Creel & Creel 1998, 2002; Woodroffe & Ginsberg 1997).

Alongside habitat loss and persecution, disease represents a significant threat to remaining populations, as wild dogs share susceptibility to diseases of common sympatric canids such as jackals and domestic dogs (Alexander et al. 2010). Outbreaks of diseases have resulted in significant population declines in the past (reviewed in Woodroffe et al. 2004a) and such outbreaks present an increasing risk as human populations and associated domestic dogs spread further into wildlife areas (Randall et al. 2006). Consequently, knowledge of the MHC is particularly pertinent to wild dog conservation.

A number of *in situ* and *ex situ* conservation programmes have been initiated to improve the status of African wild dogs. Community education programmes are being used to change negative perceptions towards wild dogs (e.g. Laikipia wild dog project)

and ecotourism used to generate revenue to support wild dog conservation (Lindsey et al. 2005). Anti-poaching patrols and road-sign awareness has also reduced snare and road mortalities (e.g. Painted Dog Research Project, in Zimbabwe). Vaccination of domestic dogs and wild dogs themselves has been implemented in some areas to reduce disease, although this remains a contentious issue in wild dog conservation following the implication that vaccination was a causal factor in the extinction of a population of wild dogs in the Serengeti in 1991 (Burrows et al. 1994). In South Africa, an artificial metapopulation was set up so to overcome the problem of the large habitat requirements of wild dogs (Endangered Wildlife Trust, South Africa). This involves conserving small numbers of wild dogs (often just one pack) in several small reserves, with occasional transfers between reserves to imitate dispersal and prevent inbreeding (Gusset et al. 2006). Largely due to experience gained from the metapopulation, translocations of wild dogs are becoming more successful (Gusset et al. 2006) and have been used to reintroduce wild dogs to areas from which they have been extirpated or remove them from areas of persecution. Work is also being done to reduce livestock depredation using bio-boundaries (Botswana Predator Conservation Trust). Lastly, ~650 wild dogs form part of an *ex situ* captive breeding programmes which have been used in reintroduction efforts, and in education and fund raising programmes (ISIS 2010). The captive population is managed as four regional programmes in Europe, Australasia, North America and South Africa. More than 40 zoos contribute to the European captive breeding programme which collectively hold almost half of world's captive population (n~270 in 2008 (Verberkmoes & Verberkmoes 2008) and therefore are central to *ex situ* conservation efforts.

1.5 Aim of the thesis

The aim of my thesis was to investigate how genetic information can be used to assist conservation of African wild dogs, with a particular focus on MHC data. There were three overall objectives to my thesis:

1. To determine how much MHC variation there is in wild dogs, and how that compares to other canids.
2. To compare the spatial and temporal structure of neutral and MHC variation across free-ranging populations of wild dogs
3. To assess how much genetic variation from wild populations is conserved in captivity, specifically, the European zoo captive population

1.6 Chapter objectives

The first stage of my thesis was to conduct a baseline study that characterises MHC variation in wild dogs. Extensive research has been conducted on the MHC class II loci in domestic dogs, specifically on the DLA-DRB1, DLA-DRA, DLA-DQA1 and DLA-DQB1 loci. Therefore, in **Chapter 2** I conducted a preliminary study of MHC polymorphism at these four loci in wild dogs. These data were compared with published data from Grey Wolves, Mexican wolves and Ethiopian wolves to assess how much MHC variation wild dogs have in comparison to other canids.

Genetic information and molecular tools have many potential applications in conservation. However, conservation genetics is an emerging field and many still question how much genetics really adds to conservation. In **Chapter 3** I present a focused case study where molecular tools were used to provide novel data that could not have been elucidated with any other approach, and which revolutionises our understanding of one of the most high profile population extinctions in conservation history; the disappearance of endangered African wild dogs (*Lycaon pictus*) from the Serengeti-Mara in 1991.

Knowledge of patterns of genetic variation is critical to understanding of population structure, local adaptation and differences in levels of diversity between populations. To achieve a thorough insight into patterns of genetic variation it is therefore important that both neutral and adaptive markers are assessed. Furthermore, since selective and neutral forces vary both spatially and temporally, it is important that studies of variation are based on samples collected at an ecologically appropriate spatial scale for the species in question and from more than one temporal period. To this end, in **Chapter 4** I assessed the forces that shape patterns of genetic variation in the highly endangered African wild dog, by assessing a spatially and temporally variable set of samples with a combination of both neutral (microsatellite and mitochondrial DNA) and adaptive (MHC) markers. Specifically I assessed the following four questions: 1) Is there a genetic signature of demographic decline in wild dogs? 2) How are African wild dog populations currently structured? 3) How are neutral and MHC diversity in African wild dogs structured temporally and spatially? 4) Is there evidence of selection at the MHC and/or local adaptation of African wild dog populations?

Assessments of MHC diversity in endangered species are increasingly common due to concerns that 'low' MHC diversity may reduce immune competence. However, the level of diversity in endangered species is typically inferred by comparison with non-endangered reference taxa, which are often distantly related or domesticated species. As

such, any differences in levels of diversity may be the result of differences in evolutionary history rather than endangered status. For example, Chapters 2 showed that levels of MHC variation in the wild dog are low in comparison to Grey, Mexican and Ethiopian wolves. This difference may be result of demographic declines, but it could just reflect the fact that wild dogs are distantly related to the reference taxa. In **Chapter 5**, I assess MHC variation across eight species of wolf-like canid clade. These data were used to assess how endangered status, selection, hybridisation and phylogeny impact patterns of MHC diversity.

Genetic management of captive breeding programmes has traditionally relied on studbooks. However, increasingly, genetic data are being utilised to improve the status of captive populations. The aim of **Chapter 6** was to assess the genetic status of the European captive wild dog population by using a combination of studbook information and genetic data based on both neutral and MHC markers. These data were used to assess how much diversity from wild populations is represented in European zoos, and also to investigate how recent imports from South African captive facilities have affected the genetic status of the European captive population. I then used these data to formulate management suggestions to improve the genetic status of the zoo population. Finally, I assessed whether patterns of diversity based on MHC and neutral markers were correlated, as such comparisons are currently lacking in the literature.

Lastly, in **Chapter 7**, I discuss the challenges of applying genetic data in conservation, as well as the broader implications of my work.

Chapter 2: Highly endangered African wild dogs *(Lycaon pictus)* lack variation at the Major Histocompatibility Complex

2.1 Abstract

The Major Histocompatibility Complex (MHC) is a set of highly polymorphic genes involved in the immune response. Extensive research on the canid MHC has found moderate to high levels of diversity at the DLA-DRB1, DLA-DRA, DLA-DQA1 and DLA-DQB1 class II loci with frequent trans-specific polymorphism among Canis species. In this study I assessed MHC variation in the more distantly related and highly endangered African wild dog (Lycaon pictus). I screened 168 African wild dogs from Eastern and Southern Africa as well as 200 samples from the European captive population for variation at MHC class II loci. As for all other canids screened to date, I found a single allele at DLA-DRA, which was the same as that found in Canis species. In contrast, I found 17 DLA-DRB1 alleles, one DLA-DQA1 allele and two DLA-DQB1 alleles, all of which were unique to African wild dogs. At DLA-DRB1, African wild dogs were found to have comparable numbers of alleles but less overall amino acid variation than other canids. However, the low numbers of alleles at DLA-DQA1 and DLA-DQB1 are surprising, given that in other canids these loci are also highly variable. Overall, our data suggest that African wild dogs are genetically depauperate at the MHC relative to other canids. These data are indicative of a loss of genetic variation, possibly as a result of population bottlenecks and declines experienced by this species.

2.2 Introduction

The Major Histocompatibility Complex (MHC) is a highly diverse set of vertebrate genes that code for molecules involved in the recognition of intra- and extra-cellular antigens, and therefore form a fundamental component of immune responses (Eggert et al. 1998; Hedrick 2003; Piertney & Oliver 2006). MHC genes are renowned for their high allelic diversity and heterozygosity which is thought to be the result of pathogen driven balancing selection (van Den Bussche et al. 1999). Diversity at the MHC is adaptively significant in disease resistance; high diversity has been shown to allow response to a wider range of parasites and pathogens than low diversity (Hedrick et al. 2001a; Hedrick et al. 2003; Sommer et al. 2002). Given the importance of adaptive genetic variation for evolutionary change and rising concerns about infectious diseases in the conservation of endangered species (Daszak et al. 2000), assessments of MHC variation are increasingly incorporated into endangered species research (e.g. giant panda, *Ailuropoda melanoleuca* (Wan et al. 2006), crested ibis, *Nipponia nippon* (Zhang et al. 2006) and Mexican wolf Canis lupus baileyi (Hedrick et al. 2000).

Considerable research has been conducted on the canid MHC (known as the dog leukocyte antigen, DLA) in the domestic dog, *Canis familiaris,* and more recently, wild Canis species: Grey wolf, *Canis lupus*, (Kennedy et al. 2007a; Seddon & Ellegren 2002); Coyote, *Canis latrans*, (Seddon & Ellegren 2002); Ethiopian wolf, *Canis simensis*, (LJ Kennedy submitted); and Mexican wolf, *Canis lupus baileyi,* (Hedrick et al. 2000). Research has focused on variation at three MHC class II loci: DLA-DRB1, DQA1 and DQB1, which are physically tightly linked and inherited as a haplotype (Kennedy et al. 2007a). MHC class II loci are involved in the recognition of antigens of extracellular pathogens and parasites. However, strong linkage disequilbrium has been found between MHC class I and II loci in humans (Sanchez-Mazas et al. 2000), domestic dogs and many other species studied, suggesting that variation at MHC class II loci can also reflect variation at MHC class I loci, which are involved in the recognition of intracellular pathogens such as viruses (Piertney & Oliver 2006). To date, 134 DLA-DRB1 alleles, 26 DLA-DQA1 alleles and 68 DLA-DQB1 alleles have been assigned official names by the DLA Nomenclature Committee (LJ Kennedy, unpublished data). These genes have been shown to be polymorphic across the Canis genus, with particularly high levels of polymorphism in both the domestic dog and Grey wolf (Kennedy 2007; Kennedy et al. 2007a; Seddon & Ellegren 2002). Trans-specific polymorphism (allele sharing) has been found to be a recurring feature among Canis species at all three loci. A fourth locus, DLA-DRA, appears to be monomorphic for allele DLA-DRA*00101 in all canids screened to date (LJ Kennedy, unpublished data). Given the focus of research on the genus Canis, it

is not currently known if these patterns of MHC polymorphism are specific to these species or a characteristic of canids in general.

African wild dogs, *Lycaon pictus*, are the sole member of the Lycaon genus and a distantly related member of the wolf-like canid clade, to which the genus Canis belongs (Girman et al. 1993). This highly endangered social species has suffered extensive declines in the wild to <6,000 individuals distributed across a few remaining small and fragmented populations (Figure 2.1) (Sillero-Zubiri et al. 2004; Woodroffe & Ginsberg 1997). Disease is argued to represent a significant threat to African wild dogs, which share susceptibility to diseases of common sympatric canids such as jackals and domestic dogs (Alexander et al. 2010), outbreaks of which have resulted in both pack and population extinctions in the past (reviewed in Woodroffe et al. 2004a). Consequently, knowledge of the MHC is particularly pertinent to African wild dog conservation.

In this study I have characterised MHC class II DLA-DRB1, DRA, DQA1 and DQB1 variation in African wild dogs to extend knowledge of the canid MHC to more distantly related canid species (Lindblad-Toh et al. 2005). Specifically, I assessed levels of polymorphism at MHC class II loci, and looked for evidence of allele sharing between African wild dogs and species in the genus Canis.

2.3 Methods

Blood, tissue, hair and serum samples were provided from free ranging study populations in Eastern and Southern Africa (Figure 2.1): Laikipia, Central Kenya (n=56 from 13 packs; study population size ~300); Serengeti, Northern Tanzania (n=14 from 4 packs; study population size ~ 160); Okavango, Northern Botswana (n=53 from 8 packs; study population size ~200); Hwange, Western Zimbabwe (n=15 from 7 packs; study population size ~250). The sampled Serengeti population, hereafter referred to as New Serengeti, represents a population that is thought to have naturally re-established in the early 2000s, rather than the Serengeti population assessed in previous genetic studies (Girman et al. 1997; Girman et al. 2001), which was extirpated with the last pack disappearing in 1991 (Woodroffe et al. 1997). South African samples were derived from a set of animals artificially reintroduced and translocated between game reserves in South Africa and included some captive animals of South African origin (A.Bastos, n=43). This South African sample set is considered a managed group of animals, rather than a free-ranging population. A further six samples from this managed group were collected from a set of 16 wild dogs that were translocated from Pilansberg Game Reserve, South Africa to Hwange National Park, ZimbabI in 2006. These animals were analysed as part of the South African sample set, rather than the Hwange sample set. The 15 Hwange samples do not include any animals recently translocated from South Africa, or their offspring. Ear or muscle samples were also provided from carcasses collected in Kajiado district in Southern Kenya (R. Woodroffe n=1), Ghanzi district in Western Botswana (M. Swarner, n=1), Northern Sofala province in Central Mozambique (J-M. André, n=3 from one pack) and Mangetti district, North Western Namibia (F. Stander, n=1). I sampled 200 captive African wild dogs (75% of the total population) from the European Endangered Species Programme (EEP), which are derived from Southern Africa. This sample set was analyzed together and is hereafter referred to as EU zoos. Details of the 31 contributing institutions are given in Table 2.1.

DNA was extracted from samples using DNeasy extraction kits (Qiagen, Crawley, UK) according to the manufacturer's instructions, with the following modifications: tissue samples were lysed for 18 hours rather than 3; blood spots and hair samples were lysed for 3 hours rather than one. A negative control was conducted with all extractions to detect contamination.

Sequence-based typing was conducted on exon 2 of the DLA-DRB1, DLA-DQA1, and DLA-DQB1 loci using locus-specific intronic domestic dog primers that gave products of 303 bp (DLA-DRB1), 345 bp (DLA-DQA1) and 300 bp (DLA-DQB1). Primers were as follows (M13 and T7 tags are underlined): DRBIn1: ccg tcc cca cag cac att tc (Wagner et

al. 1996b); DRBln2M13r: cag gaa aca gct atg acc tgt gtc aca cac ctc agc acc a (Wagner et al. 1996b); DQAln1: taa ggt tct ttt ctc cct ct (Wagner et al. 1996a); DQAIn2: gga cag att cag tga aga ga (Wagner et al. 1996a). DQB1BT7 taa tac gac tca cta tag gg ctc act ggc ccg gct gtc tc (Wagner et al. 1996a); DQBR2: cac ctc gcc gct gca acg tg (Kennedy et al. 2002a). A fourth MHC class II locus DLA-DRA, which has been shown to be monomorphic in all other canids tested to date (LJ Kennedy, unpublished data) was examined using locus-specific exonic primers: DRAF: gag cac gta atc atc cag gc; DRAR: ggt gtg gtt gga gcg cgc ttt a (JL Wagner, personal communication) and gave products of approximately 261 bp.

Polymerase Chain Reactions (PCR) were performed in 25-µl reactions containing 1 x Q solution (Qiagen), 1 x PCR buffer containing, 1 mM MgCl (Qiagen), 0.4 mM of each DNTP (Invitrogen), 0.04 mM of each primer, 0.1 µg/µl BSA (Promega), 1 unit of Hot Startaq (Qiagen) and approximately 25ng of template DNA. To detect contamination, each PCR was run with both the DNA extraction negative and a PCR negative control containing no template DNA. Reactions were run on PTC-200 DNA engine machines (MJ Research Inc). PCR amplifications were conducted with a touchdown protocol: 15 min at 95°C, 14 touchdown cycles of 95°C for 30 s, followed by 1 min annealing, starting at 62°C (DLA-DRB1), 62°C (DLA-DRA), 52°C (DLA-DQA1), 68°C (DLA-DQB1) and reducing at 0.5°C per cycle, and 72°C for 1 min. This was followed by 20 cycles of 95°C for 30 s, 60°C (DLA-DRB1), 55°C (DLA-DRA), 50°C (DLA-DQA1), 65°C (DLA-DQB1) for 1 min, and 72°C for 1 min. The protocol ended with a final extension of 72°C for 10 minutes. The number of amplifications in the second stage of the PCR protocol was increased from 20 to 30 cycles for DNA derived from hair, blot spot and serum samples, which typically yielded lower quantities of DNA.

PCR products were cleaned using ExoSAP-IT (USB) according to the manufacturer's instructions, and sequenced on an ABI 3730 sequencer. Sequencing was conducted in both directions for DLA-DRB1, using primers DRBln1 and M13r (cag gaa aca gct atg acc). To reduce costs, unidirectional sequencing was used for DLA-DQA1 and DLA-DQB1, using primers DQAln1 and DQB1BT7, respectively. Sequence data were analysed using Match Tools and Match Tools Navigator (Applied Biosystems), as described in Kennedy (2002a). This method relies on an allele library built from homozygotes. I had 6 heterozygous individuals (S.Africa=3, EU zoos=3) that did not match any pair of known alleles. Therefore I cloned these six individuals using the TOPO TA™ cloning system (InVitrogen, San Diego, CA) and identified a single new allele DRB1*90301. This allele was subsequently found in a further 12 heterozygotes. There were three pairs of alleles which could not be distinguished using the Match tools analytical method because some allele pairs gave the same heterozygous sequence (DRB1*90601/90202 and DRB1*90602/90201; DRB1*90101/90201 and

DRB1*90102/90202; DRB1*90101/90601 and DRB1*90102/90602). These ambiguous combinations were resolved using a combination of Reference Strand-mediated Conformation Analysis (RSCA) and pedigree information from the zoo populations. RSCA is a genotyping method that separates allelic variants based on conformation-dependent mobility through a gel (Kennedy et al. 2005), and was used to distinguish between ambiguous DLA-DRB1 heterozygous sequences by running ambiguous samples alongside a set of candidate alleles in homozygous form. For the EU zoo samples, individuals with ambiguous allele combinations could be resolved using pedigree information to examine the alleles of siblings, parents and offspring. For example individual #P20791 was found to be heterozygous for either 1) DRB1*90601/DRB1*90202 or 2) DRB1*90602/DRB1*90201. Five of its siblings were found to have the following four alleles DRB1*90101, DRB1*90201, DRB1*90301 and DRB1*90602, which means that #P20791 must be heterozygous for DRB1*90602/DRB1*90201. Pedigree data were also used to examine segregation of DLA-DRB1 alleles and lineages within families. Chi-square goodness of fit tests were used to compare observed segregation patterns to expected genotype combinations under random segregation at a single locus. Pedigree information for the EU zoo samples were provided by H.Verberkmoes. Pedigrees were drawn using SmartDraw 2009.

Preliminary sequencing of 30 individuals for DLA-DQA1, DLA-DQB1 and DLA-DRA revealed just one, two and one alleles respectively. Consequently, I used RSCA (together with sequenced samples as controls), to screen for further variation at these loci. For DLA-DQA1, DLA-DQB1 and DLA-DRA, RSCA analysis was conducted on samples from EU zoos (n=92), Laikipia (n=56), New Serengeti (n=9), Okavango (n=53), Hwange (n=13), S.Africa (n=6) and the 6 carcass samples. DNA from 5 New Serengeti, 2 Hwange and 43 South Africa samples were not available in time for RSCA analysis, however sequence based typing detected no new DLA-DRB1 alleles in these samples. DLA-DQB1 typing was conducted on an additional 25 Okavango samples that were not successfully typed at the DLA-DRB1 due to low quality DNA and RSCA failures. Since RSCA was used to screen for new variants and the EU zoos included large family groups, I did not type offspring if I typed both parents, and screened a maximum of three animals per litter. In total, I typed 92 individuals representative of 36 sibling groups from the 200 captive samples.

The new alleles identified in this study were submitted to the DLA nomenclature committee. Those that met the appropriate criteria were recognized and assigned official names by the committee. Prior to this study, preliminary data (LJ Kennedy, H Bacon and A Radford, unpublished data) based on four African wild dog museum samples provided by the National Museums of Scotland (A Kitchener), had identified three DLA-DRB1 alleles (DLA-DRB1*90101, 90102 and 90201), one DLA-DQA1 allele (DLA-DQA1*01901)

and two DLA-DQB1 alleles (DLA-DQB1*90101, 90201). One allele did not fulfill the naming criteria, and is referred to by its local name, 'fmut'.

Sample sizes varied from 14-56 for non-managed populations. Therefore, I used rarefaction to compensate for sampling disparity between study populations by standardizing to a population size of 10 using the programme HP-Rare v.4.1 (Kalinowski 2005). I calculated nucleotide diversity in populations as the average number of segregating sites θ, and pairwise diversity π, in DnaSP 4.20 (Rozas & Rozas 1995), using a Jukes Cantor model of substitutions and standard errors calculated with 5,000 bootstrap replications. I tested for an excess of heterozygosity relative to Hardy Weinberg proportions, which is indicative of selection on the current generation, using the U test in Genepop 4.0 (Raymond & Rousset 1995). Synonymous and nonsynonymous genetic distances were calculated separately for putative peptide binding region (PBR) sites and non-PBR sites using the Nei-Gojobori method with a Jukes Cantor model of substitutions in Mega 4.0 (Tamura et al. 2007). Putative PBR sites were based on the human HLA-DRB1 (Brown et al. 1993). Due to the recombining nature of MHC genes, phylogenetic trees are not strictly appropriate for analysis of the MHC and there is too much variation to allow a network approach. However, MHC allele trees are a useful tool for displaying relationships among alleles. Phylogenetic trees were constructed using African wild dog sequences alongside 105 alleles from Canis species made available by LJ Kennedy, who collates these data on behalf of the DLA nomenclature committee (Kennedy et al. 2001). I also tested alternative phylogenetic models but these did not affect the resolved relationships within the tree. Therefore, I have only shown neighbor joining trees with Kimura's 2 parameter model as implemented in Mega 4.0 to demonstrate relationships. Following Seddon (2002), a human HLA sequence with ~80% similarity to dog DLA-DRB1 alleles was used as an outgroup (HLA-DRB1*03011, Accession number AF352294). Bootstrapping was conducted with 5,000 replicates.

2.4 Results

African wild dogs were found to have 17 DLA-DRB1 alleles (n=368), one DLA-DQA1 allele (n=234), two DLA-DQB1 alleles (n=234) and one DLA-DRA allele (n=234) (official and local names are provided in Appendix 1). Fewer samples were analyzed at DLA-DQB1, DLA-DQA1 and DLA-DRA because of the lack of variation found. However, I did type representative individuals for all DLA-DRB1 alleles. This is important because in domestic dogs and other Canis species there is strong linkage between MHC class II loci. Therefore, new DLA-DQB1, DLA-DQA1 and DLA-DRA variants would be most likely found in individuals with new DLA-DRB1 alleles. There was no evidence of pseudogenes (stop codons or frameshift mutations), indicating that functional genes were being amplified. All DLA-DRB1, DQA1 and DQB1 alleles detected in African wild dogs were new and have not been identified in any other canid species to date; accession numbers DQA1 (AM182470), DQB1 (FJ648575, FJ648576), DRB1 (FJ648559-FJ648574). As with all other surveyed wolf-like canids (LJ Kennedy, unpublished data), African wild dogs were monomorphic at DLA-DRA for allele DRA*00101, which was originally identified in domestic dogs (Wagner et al. 1995).

African wild dog DLA-DRB1 alleles varied at 31 polymorphic sites across 95 codons, with 14 substitutions at the 1st codon position, 10 at the 2nd codon position and seven at the 3rd codon position. These changes corresponded to 17 amino acid differences amongst alleles (Figure 2.2). This included unique amino acid residues at two codons not seen in other canids and one new polymorphic site at a putative PBR residue which is monomorphic in all other canids. All DLA-DRB1 alleles differed from each other at the amino acid level, except for DLA-DRB1*907011 and DRB1*907012, which indicates a high level of non-synonymous substitutions. The majority of nucleotide (22/31) and amino acid (14/17) differences between DLA-DRB1 alleles were found to occur within the three hypervariable regions (HVR) (Figure 2.2) (Kennedy et al. 2007). Nine of the 22 functionally important putative PBR sites of DLA-DRB1 based on human HLA-DRB1 were variable in African wild dogs (Figure 2.2). The ratio of non-synonymous to synonymous substitutions at the putative PBR sites was greater than 1.0, and larger than in non-PBR, but it was not found to be significant (PBR dN = 0.2, dS = 0.117, dN/dS = 1.709, p=0.073; non-PBR dN=0.031, dS = 0.022, dN/dS = 1.409, p=0.307).

DLA-DRB1 alleles consisted of two highly divergent allelic lineages, which I have called A (7 alleles) and B (10 alleles). Alleles within lineages were relatively similar whereas alleles from different lineages were highly divergent (Figure 2.2). Lineage A alleles have identical HVR1 and HVR2 sequences. Lineage B alleles have the same HVR1 sequence (which is different from that in lineage A), and one of two very similar

HVR2 sequences which differed by just one amino acid. At HVR3, there were five different sequences, three of which were shared amongst lineages, and two of which were specific to Lineage B. Overall, the average numbers of nucleotide differences within alleles of the same lineage were 6.0 (lineage A) and 6.8 (lineage B), compared with an average of 22.9 nucleotide differences between alleles from different lineages. Since RSCA analysis, DNA cloning and sequencing did not detect more than two alleles in any individual, and less than half of the individuals sampled (46%) had alleles from both lineages, I am confident that these two allelic lineages are derived from a single locus. Furthermore, pedigree data clearly show co-segregation of the two allelic lineages within families (Figure 2.3).

Phylogenetic analyses on African wild dog DLA-DRB1 alleles were conducted alongside alleles from Canis species. The highly polymorphic nature of these genes resulted in insufficient resolution to determine specific relationships between groups of alleles; however they were used to indicate the positioning of African wild dog alleles relative to the alleles of other canids (Figure 2.4). African wild dog DLA-DRB1 alleles were clearly shown to cluster into two distinct and separate monophyletic clades, rather than being scattered across branches, as found with Grey wolf and Ethiopian wolf alleles. Furthermore, African wild dog alleles were clearly positioned within, rather than peripheral to, the canid DLA-DRB1 allele tree, indicating similarity to other canid alleles. In particular, comparison of amino acid sequences highlight that certain African wild dog DLA-DRB1 lineage B alleles and certain Ethiopian wolf alleles differ by just one amino acid at HVR 1 and are identical at HVR 2 (data not shown).

DLA-DRB1 alleles from both A and B lineages were found in all populations with more than three samples. Four of seven lineage A, and five of nine lineage B DLA-DRB1 alleles were detected in two or more sampling areas, which were often separated by large geographic distances (Table 2.2). For example, DLA-DRB1*90202 was found in countries across Eastern (Laikipia-Kenya, New Serengeti-Tanzania) and Southern Africa (Hwange-Zimbabwe, Okavango-Botswana, NW Namibia and South Africa).

African wild dog populations were found to differ from each other in DLA-DRB1 allelic composition, allelic diversity and heterozygosity. For non-managed populations, the number of alleles per population varied between 3 and 9, and average observed heterozygosity varied from 53.6 % - 92.9 % (Table 2.2). Despite being the most thoroughly sampled population, Laikipia had the smallest number of alleles (3 alleles, n=56), and correspondingly, also had the lowest observed heterozygosity (53.6%). However, nucleotide diversity was actually highest in this population (π=0.0758, θ=0.0716), suggesting that the three alleles are highly divergent; there were 29 variable sites amongst these three alleles. In contrast, nucleotide diversity was lower in the three

other populations, which had between 7 and 9 alleles (π=0.0509, 0.0595, 0.0613; θ=0.0435, 0.0435, 0.0484). Rarefaction was used to standardize population sample sizes to n=10, and showed Hwange to be most diverse in terms of numbers of alleles expected with that sample size (7.8 alleles), although New Serengeti and Okavango had only slightly lower levels of diversity (6.1 and 5.8 alleles respectively). All three of these populations had at least 50% more diversity than Laikipia (2.9 alleles). Although levels of observed heterozygosity were generally high, there was not an excess of heterozygosity relative to Hardy Weinberg expectations in any non-managed population. The South African sample set consisted almost entirely of heterozygotes (46/49). However, this is a managed group of animals derived from multiple sources, rather than a natural population. Together, the EU zoos were found to have 12 of the 14 DLA-DRB1 alleles detected in Southern African populations, and levels of heterozygosity (82%) comparable to non-managed wild populations (53.6-92.9%). One allele (DRB1*90101) was found at high frequency among the zoo samples (33.5%).

The two DLA-DQB1 alleles differed at eight sites within HVR 2, resulting in five amino acid differences. This included one new polymorphic amino acid site that is monomorphic in other canids tested to date, and four unique amino acid residues. DLA-DQB1*90101 was considerably more frequent (87.5%) than DLA-DQB1*90201 (12.5%) resulting in a predominance of DLA-DQB1*90101 homozygotes (81%). In fact, I found just six DLA-DQB1*90201 homozygotes in 234 samples. Both DLA-DQB1 alleles were found across Eastern and Southern Africa (Table 2.3), however, DLA-DQB1*90201 was noticeably absent from Hwange, Zimbabwe. Previous research has shown strong linkage disequilibrium between the canid DLA-DRB1, DQA1 and DQB1 loci (Kennedy et al. 2007a). There was insufficient variation at the DLA-DQA1 and DQB1 loci for haplotype designation in African wild dogs. However, I did detect an association between DLA-DQB1*90201 and DLA-DRB1 lineage A alleles. Six out of seven individuals homozygous for DLA-DQB1*90201, had only lineage A DLA-DRB1 alleles, (DRB1*90101, *90201, *90202 or *90204). Furthermore, all DLA-DQB1*90201 heterozygotes had at least one DLA-DRB1 lineage A allele, most commonly DRB1*90101, DRB1*90201, or DRB1* 90202.

2.5 Discussion

Research on MHC class II loci in *Canis* species has shown moderate to-high levels of diversity at the DLA-DRB1, DLA-DQA1 and DLA-DQB1 class II loci with frequent trans-specific polymorphism (allele sharing) among Canis species. In this study I conducted a geographically widespread survey of MHC class II variation in the highly endangered African wild dog to extend knowledge of the canid MHC to more distantly related canid species. African wild dogs belong to a monotypic genus that is phylogenetically and morphologically divergent from *Canis* species (Bardeleben et al. 2005; Wayne et al. 1997). In total, I found 17 alleles at the DLA-DRB1 locus, one allele at the DLA-DQA1 locus and two alleles at the DLA-DQB1 locus, all of which are currently unique to African wild dogs. At DLA-DRA, African wild dogs were monomorphic for the same allele found in other canids.

Balancing selection is a key mechanism in the maintenance of variation at MHC loci (reviewed in Garrigan & Hedrick 2003) and is indicated by an increased ratio of non-synonymous (dN) to synonymous (dS) substitutions at the amino acid residues of the functionally important PBR (Seddon and Ellegren 2002). Although dN/dS was elevated at putative PBR sites of DLA-DRB1 alleles in African wild dogs, there was not a significant excess of non-synonymous substitutions ($p=0.073$). This is not typical of canid DLA-DRB1 alleles; there was a significant excess of dN/dS at PBR sites in Grey wolves, Coyotes and domestic dogs (Seddon and Ellegren 2002). Whereas dN/dS ratios provide information on historical selection, excess heterozygosity can provide an indication of current selection at a locus (Aguilar et al. 2004; Garrigan & Hedrick 2003). Despite the high heterozygous frequencies found in non-managed free ranging populations, the observed heterozygosity did not exceed Hardy Weinberg expectations. This is not atypical for MHC studies (Garrigan & Hedrick 2003).

The distribution of alleles from polymorphic loci under balancing selection are predicted to show very different distributions from that of neutral loci. In particular, they are expected to show lower levels of differentiation in allele composition between populations (Schierup et al. 2000). Neutral genetic markers show strong structuring and differentiation between African wild dogs populations, in particular between Eastern and Southern Africa (Girman et al. 2001). At the MHC, I found 17 DLA-DRB1 alleles, which clustered into two highly distinct lineages. These two lineages showed no evidence of geographic structuring; all areas where more than three animals were sampled had alleles from both lineages. Similarly, individual DLA-DRB1 alleles were not geographically restricted, with many alleles detected in populations spanning Eastern and Southern Africa. The discordance between patterns of MHC and neutral variation could indicate that selective

forces are shaping patterns of MHC diversity across African wild dog populations; for example, selection for alleles which confer resistance to diseases common to most populations.

Two DLA-DQB1 alleles were detected in African wild dogs. However, allele DLA-DQB1*90201 was considerably rarer (12.5%). This rare allele was found across Eastern and Southern Africa but was absent from Hwange. This may be the result of the low frequency of DLA-DRB1 lineage A alleles in Hwange (20%), which appear to be associated with DLA-DQB1*90201. The stark differences in frequency of the two DLA-DQB1 alleles may be indicative of selection on adaptive differences between these alleles or haplotypes.

High MHC allelic diversity in a population and high heterozygosity in individuals is thought to be important because it theoretically expands the range of pathogens to which a population or individual can respond (Doherty & Zinkernagel 1975; Sommer et al. 2002). I found that the number of DLA-DRB1 alleles and levels of heterozygosity varied between populations (Table 2.2), even after population sample sizes were standardized using rarefaction. This may reflect differences in demographic history and connectivity. The highest allelic diversity in non-managed populations was found in Hwange (9 alleles, n=15), which is a long-standing stable population located within an admixture zone (Girman et al. 2001). In contrast, the lowest number of alleles was found in Laikipia (3 alleles, n=56), a recently recolonised population, that is also relatively isolated (Woodroffe et al. 2007b). Clearly, however, recolonisation does not always result in low numbers of alleles, since the recently recolonised New Serengeti population was considerably more diverse than Laikipia. However, the New Serengeti is linked to a number of other African wild dog populations and therefore may have been recolonised by a mixture of founders from multiple source populations. Despite lower allele numbers, nucleotide diversity among alleles was higher in the two recently re-colonized populations (Laikipia and New Serengeti) than in two long-standing populations (Hwange and Okavango) (Table 2.2). The lower nucleotide diversity measures in Hwange and Okavango likely reflect the presence of closely related or similar alleles, and may indicate that in these populations new diversity has been accumulating, whereas in the recently recolonised populations there has been insufficient time for the evolution of new variants. More research is required to explore whether differences in MHC diversity between populations reflect differences in disease characteristics of populations or neutral processes such as size of historical bottlenecks.

The use of fitness related genes, such as the MHC, in endangered species management remains a contentious issue. Nonetheless, it is valuable to evaluate the impact of management actions, such as translocations and captive breeding, on adaptive

genes. In 2006, 16 African wild dogs were translocated from South Africa to Hwange, Zimbabwe. Sampling of six of these South African translocated animals detected one allele (DLA-DRB1*90301) not present in the 15 resident Hwange samples. This may indicate that the translocation has introduced new MHC diversity into the Hwange population. Our results show that 12 of the 14 DLA-DRB1 alleles found in Southern African populations and both DLA-DQB1 alleles are represented in the European zoo African wild dog population. Nonetheless, allele DLA-DRB1*90101 clearly dominates this population (33.8%). High frequency of this allele does not appear typical to Southern Africa where the EU zoo founders originated; it has less than 10% representation in Hwange, Okavango and South Africa. Mapping DLA-DRB1 alleles onto the EU zoo pedigree (data not shown) shows that over-representation of this allele is the result of an extreme bias in founder contributions, and is a major cause of homozygosity in the EU zoos (21/35 homozygotes were homozygous for DRB1*90101). Management of this population is now focusing on equalizing founder representation.

The patterns of MHC variation detected in African wild dogs are best interpreted through comparison with other canids. Extensive research on the MHC in Canis species show frequent trans-specific polymorphism at DLA-DRB1, DQA1 and DQB1 loci (Kennedy et al. 2007a; Kennedy et al. 2001; Seddon & Ellegren 2002). By contrast, all alleles characterized at these three loci in African wild dogs were unique to this species and not yet found in any species of Canis. Furthermore, phylogenetic analyses of African wild dog DLA-DRB1 alleles showed clustering into two distinct branches (species-specific allelic clustering), rather than a scattered distribution throughout the DLA-DRB1 tree indicative of trans-specific polymorphism (as seen in Grey wolves and Ethiopian wolves). Such a distribution may suggest that the canid DLA-DRB1, DLA-DQA1 and DLA-DQB1 allele lineages diverged prior to speciation within the genus Canis 1-2 Myr ago (Seddon & Ellegren 2002), but after the divergence of the Lycaon and Canis genera approximately 4-5 Myr ago (Wayne et al. 1997). However, given that allele sharing is most common among species of Canis at DLA-DQA1 and DLA-DQB1 loci (Kennedy et al. 2007a; Seddon & Ellegren 2002), whereas there are one and two alleles respectively in African wild dogs, it is also possible that shared alleles have been lost.

Allelic diversity at DLA-DQA1 and DLA-DQB1, was much lower in African wild dogs than expected based on the pattern found in other canids (Table 2.4; (Hedrick et al. 2000; Kennedy 2007; Kennedy et al. 2007a; Kennedy et al. 2002b; Seddon & Ellegren 2002). This cannot be explained by the endangered status of African wild dogs or differences in sampling since they had lower levels of DLA-DQA1 and DLA-DQB1 variation than Ethiopian and Mexican wolves; two other endangered canids sampled from single populations (Table 2.4). It is particularly striking that five DLA-DQA1 alleles were found in fewer than seven Mexican wolves sampled from a single population, whereas in

this study I found just a single allele in 234 African wild dogs sampled across Eastern and Southern Africa. The lack of variation at these loci does not appear to be the result of non-matching primers as all samples amplified successfully; if a mutation had occurred in the primer site, homozygous individuals for these alleles should fail to amplify. African wild dogs showed the most variation at the DLA-DRB1 loci, where they had the same number of DLA-DRB1 alleles to Grey wolves sampled across a similar geographic range in both European and North American regions (Table 2.4), and slightly higher numbers of DLA-DRB1 alleles in single populations, than other endangered canids. However, because African wild dog DLA-DRB1 alleles are derived from just two allelic lineages, amino acid diversity amongst alleles was considerably lower than for other canids: 17 variable amino acid sites across 17 DLA-DRB1 alleles in African wild dogs, compared with 26 variable amino acid sites across 17 alleles in the North American Grey wolf (Seddon & Ellegren 2002) and 22 variable amino acids sites amongst just four alleles in a single Ethiopian wolf population (LJ Kennedy, submitted). Furthermore, there was less variation at the putative PBR site residues, which are thought to be primarily responsible for functional differences between alleles (Sommer 2005), in African wild dogs compared with Ethiopian wolves (LJ Kennedy, submitted) and North American Grey wolves (Kennedy et al. 2007a): total number of variable PBR sites, 9, 11 and 15 respectively; average number of residues/PBR site, 1.5, 1.7 and 2.2 respectively. Consequently, one might speculate that although African wild dogs have a large number of DLA-DRB1 alleles, they may have little functional diversity. Overall, our data suggest that African wild dogs are genetically depauperate at the MHC relative to other canids. They have uncharacteristically low amino acid diversity at the DLA-DRB1 locus and low numbers of alleles at the DLA-DQA1 and DQB1 loci, for a canid, even for an endangered one.

African wild dogs may have lost allelic diversity across all MHC class II genes due to historical bottlenecks, with strong disease pressures subsequently maintaining or generating MHC variation at the least conserved region, in this case the DLA-DRB1 locus. The presence of just two highly divergent monophyletic allelic lineages for both DLA-DRB1 and DLA-DQB1 is consistent with the hypothesis that this species suffered severe bottlenecks, resulting in the loss of alleles, and subsequent evolution of new diversity (van Den Bussche et al. 1999). However, both DLA-DQB1 alleles and both DLA-DRB1 lineages, were represented across African wild dog populations. A range-wide bottleneck would be unlikely to produce such a consistent pattern of diversity loss across populations because this would result in the random loss of variation. It is more likely that African wild dogs suffered local population extinctions across most of the African wild dog range, with remnant populations retaining both allelic lineages and subsequently expanding to re-colonise their former range.

It is clear from our study that African wild dogs are atypical in their patterns of MHC diversity amongst the canids which have been studied to date. However, all canids previously studied at the MHC (domestic dogs, Grey wolves, Coyotes, Ethiopian wolves, Red wolves, and Mexican wolves) have been closely related species of *Canis*, that have a long history of hybridization (Verginelli et al. 2005; Vilà et al. 1997). Consequently I cannot distinguish whether African wild dogs show different patterns of MHC polymorphism to *Canis* species because of factors related to African wild dog demographic history, rather than their distant phylogenetic relationship to the *Canis* genus, or the fact that they lack extensive hybridization in their recent evolutionary history. Future work is planned on other non-hybridizing species of the wolf like clade to investigate these alternative hypotheses.

Adaptive genetic variation is of primary interest in conservation genetics, therefore MHC data have particular application to endangered species programs (Aguilar et al. 2004). In this study I have shown that the highly endangered African wild dog has a reduced level of MHC variation compared with other canids, perhaps as a result of historical bottlenecks. Amongst African wild dog populations, levels of MHC diversity were found to vary, but more research is required to investigate the significance of this in relation to differences in disease incidence and exposure. Our data have shown that the distribution of MHC variation does not match the pattern of neutral genetic variation highlighted in previous studies (Girman et al. 2001), indicating that conservation plans based on neutral genetic data alone may not adequately conserve adaptive genetic variation. It is promising however, that such a high proportion of MHC diversity from free ranging populations has been successfully conserved within the European captive breeding programme. This high diversity is likely the result of the diverse origin of individuals that founded the European captive population.

Figure 2.1: Historic (light gray) and present (dark gray) range of African wild dogs according to McNutt (2008). Sampling locations are shown with circles. Carcass samples are depicted with smaller circles and italics. Country codes – Kenya KNY, Tanzania TNZ, Zimbabwe ZIM, Botswana BOT, Mozambique MOZ, South Africa SAF, Namibia NAM.

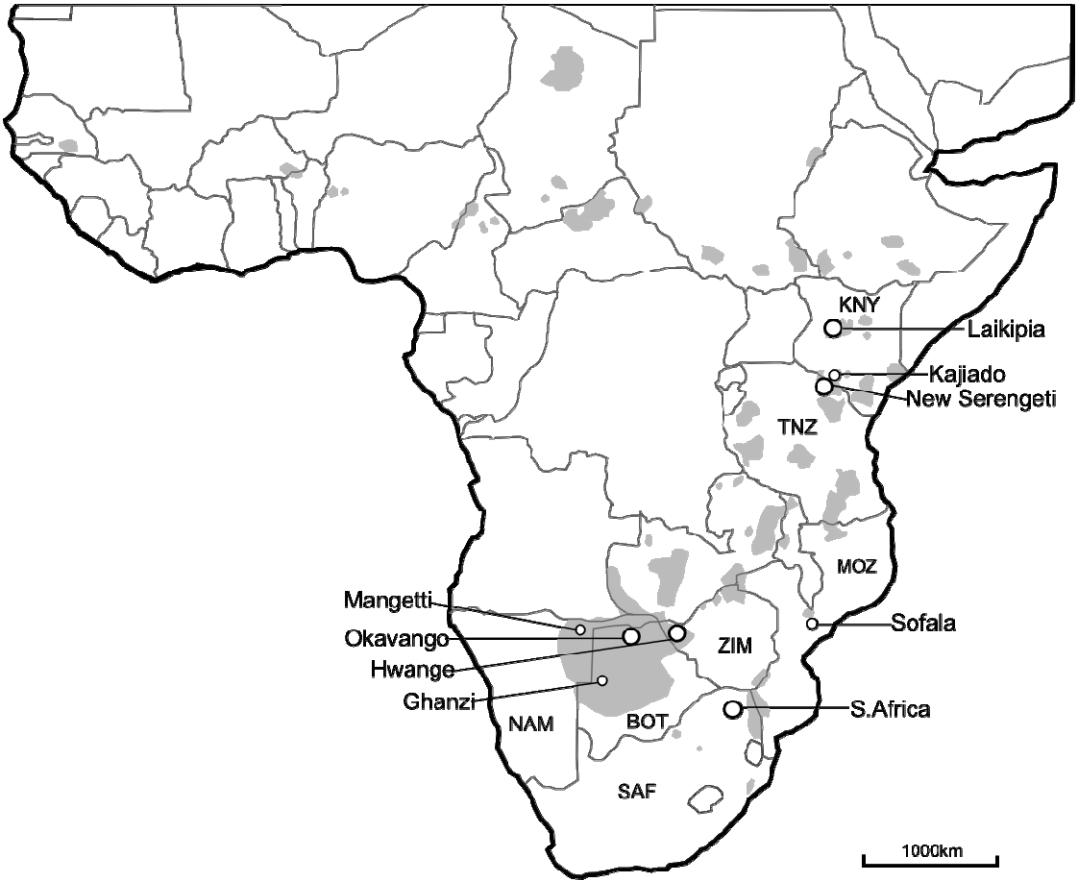


Figure 2.2: African wild dog DLA-DRB1 alleles aligned to domestic dog DLA-DRB1*00101 sequence. Matching amino acids are indicated with a dash, varying amino acids are indicated by single letter amino acid codes. HVR 1, 2 and 3 are shaded in grey. PBR are depicted with an asterisk. Alleles are grouped into two phylogenetically divergent allelic lineages, A (above the line) and B (below the line).

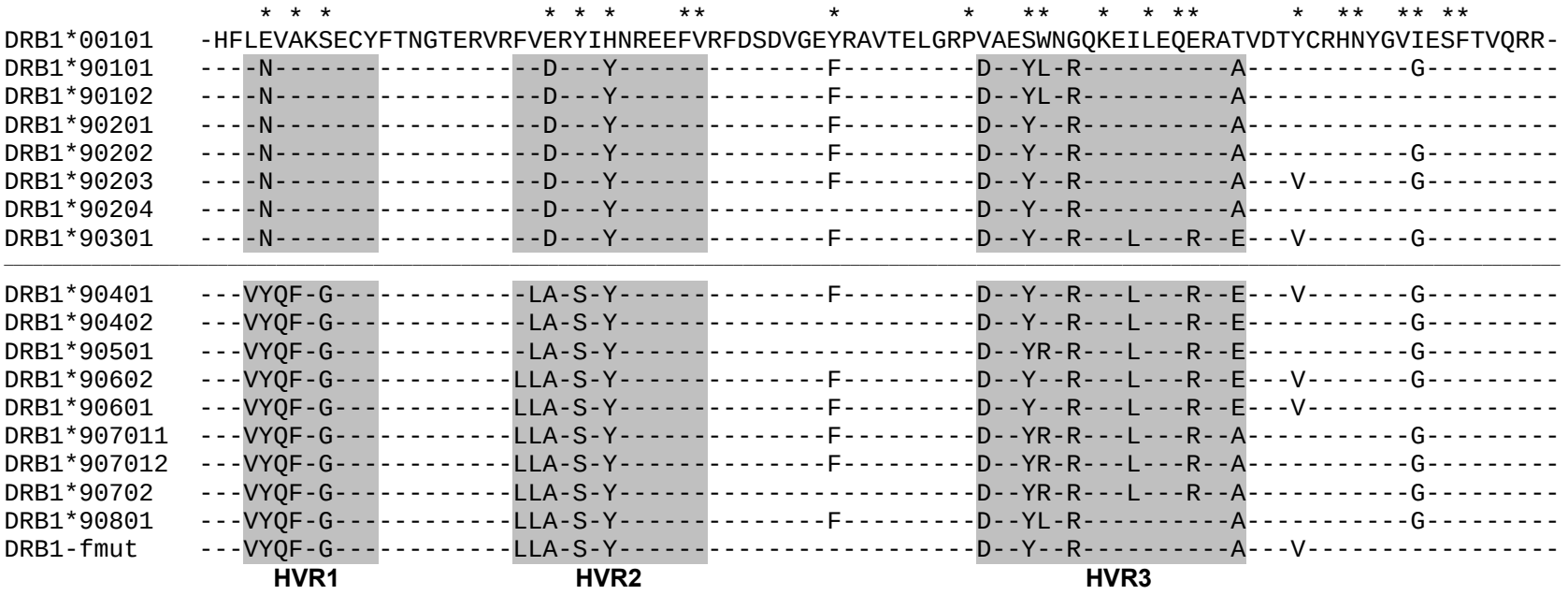


Figure 2.3: Segregation analysis of DLA-DRB1 alleles according to sequence based typing data of captive African wild dog samples from European zoos. African wild dog DLA-DRB1 alleles comprise two highly divergent allelic lineages, A and B. Lineage B alleles are underlined to demonstrate segregation of these allelic lineages. Family 1 represents an example where the mother has two lineage B alleles and the father two lineage A alleles. Each offspring is seen to inherit one lineage A allele from their mother, and one lineage B allele from their father. The two expected genotype classes (90401/90101 and 90401/90102) occur at a frequency of 8 and 7, respectively, which is not significantly different than expected for a single locus ($p>0.95$). Family 2 is an example of segregation where both parents have one lineage A and one lineage B allele. Although the expected frequency of each genotype class (1.75 for each of the four possible combinations of parental alleles) is too low to reliably apply a chi-square goodness of fit test, each expected genotype occurs at least once. Three out of seven offspring are shown to inherit a lineage A allele from both parents, two offspring inherit a lineage B allele from both parents, and two offspring inherit a lineage A allele from one parent, and a lineage B allele from the other parent.

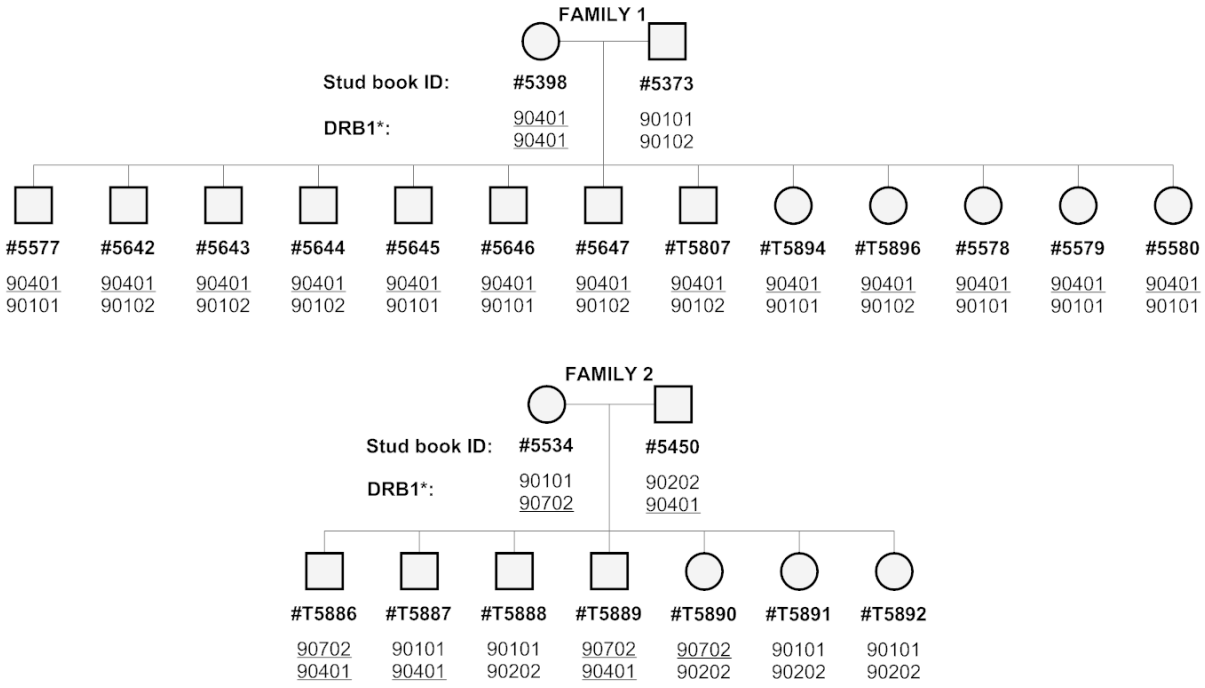


Figure 2.4: Neighbor-joining tree of DLA-DRB1 alleles using a Juke's Cantor model of nucleotide substitution. Bootstrapping was conducted with 5000 replicates; black circles depict nodes with >70% support. African wild dog sequences are highlighted in red, and cluster into two highly divergent lineages rather than being scattered throughout the tree as for Grey wolf (blue) and Ethiopian wolf (yellow) alleles.

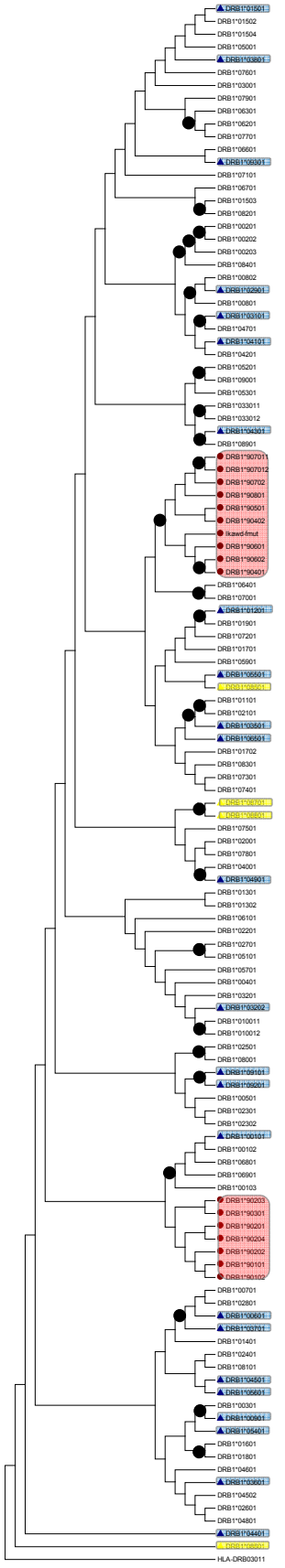


Table 2.1: List of captive African wild dog samples contributed to this study by European zoological institutions.

Contributing zoological institution	No. samples contributed
Aalborg zoo, Aalborg, Denmark	9
Artis zoo, Amsterdam, Netherlands	4
Attica Zoological Park, Spata, Greece	2
Beekse Bergen Safarai Park, Hilvarenbeek, Netherlands	11
Borås Djurpark Zoo, Alvsborg, Sweden	12
Centre d'Etude Rech Zool Augeron, Lisieux, France	4
City of Belfast Zoo, Belfast, UK	3
Colchester Zoo, Essex, UK	4
Duisberg zoo, Duisberg	3
Ebeltoft zoo, Ebeltoft, Denmark	4
Edinburgh Zoo, Edinburgh, UK	5
GaiaPark Kerkrade Zoo, Kerkrade, Netherlands	7
Kolmården Djurpark AB, Kolmården, Sweden	20
La Palmyre Zoo, Royan, France	3
Le Pal Parc Animalier, Dompierre-sur-Besbre, France	3
London zoo, London, UK	4
Munchener Tierpark Hellabrunn, Muenchen, Germany	7
Parken Zoo Eskilstuna AB, Sodermanland, Sweden	2
Port Lymne Wild Animal Park, Hythe, UK	23
Fondazione Bioparco di Roma, Rome, Italy	2
Rostock Zoologischer Garten, Rostock, Germany	5
Safari De Peaugres, Peaugres, France	7
West Midland Safari & Leisure Park, Worcester, UK	20
Zoo Basel, Basel, Switzerland	5
Zoo Dortmund, Dormund, Germany	3
Zoo D'Amneville, Amneville, France	4
Zoo Dvůr Králové, Dvůr Králové nad Labem, Czech Republic	4
Zoological Center Tel Aviv, Ramat Gan, Israel	1
Zoological Society of Ireland-Dublin, Dublin, Ireland	2
Zoo de Pont-Scorff, Pont-Scorff, France	13
Royal Museum of Scotland – deceased EU zoo animals	4

Table 2.2: Percentage of DLA-DRB1 alleles and lineages across sampling localities, subdivided into free-ranging nonmanaged populations, and samples from a managed population, and carcass and captive samples.

		Non-managed populations				Managed	Carcass samples				Captive
	Locus DRB1*	**Laikipia KNY (n=56)**	**N.Serengeti TNZ (n=14)**	**Hwange ZIM (n=15)**	**Okavango BOT (n=28)**	**S.Africa SAF (n=49)**	**Kajiado KNY (n=1)**	**Ghanzi BOT (n=1)**	**Sofala MOZ (n=3)**	**Mangetti NAM (n=1)**	**EU zoos (n=200)**
Lineage A	**90101**			6.7	8.9	5.1					33.5
	90102		3.6		32.1						12.3
	90201			3.3	32.1	24.5		100.0			9.5
	90202	57.1	14.3	10.0	7.1	25.5				100.0	5.8
	90203		42.9				50.0		100.0		
	90204		14.3								
	90301					4.1					3.5
Lineage B	**90401**			10.0							13.0
	90402			33.3	1.8	7.14					1.5
	90501	15.2		3.3							
	90602		3.6								6.5
	90601	27.7			12.5	4.1	50.0				
	907011			16.7		21.4					4.0
	907012			13.3							0.3
	90702			3.3							9.8
	90801		17.9		5.4	8.2					0.5
	fmut		3.6								
Total # alleles		3	7	9	7	8	2	1	1	1	12
% Lineage A		57.1	75.0	20.0	80.4	59.2	50.0	100.0	100.0	100.0	57.1
% Lineage B		42.9	25.0	80.0	19.6	40.8	50.0	0.0	0.0	0.0	42.9
H_o %		53.6	92.9	73.3	82.9	93.9*	NA	NA	NA	NA	82.0
H_e %		57.9	76.7	84.4	77.5	82.0	NA	NA	NA	NA	82.9
Standardised # alleles, n=10		2.9	6.1	7.8	5.8						

π	0.0758 (0.024)	0.0613 (0.009)	0.0509 (0.0001)	0.0595 (0.011)
π(syn)	0.0482	0.0450	0.0294	0.0391
π(nonsyn)	0.0810	0.0632	0.0552	0.0628
Θ	0.0716 (0.0019)	0.0435 (0.0001)	0.0484 (0.0001)	0.0435 (0.0004)

n=number of individuals typed.
H_o = observed heterozygosity (%). *denotes significant excess to Hardy Weinberg proportions (HWE), $p<0.001$.
H_e = expected heterozygosity under HWE (%).
The number of alleles was standardized for a population size of 10 using rarefraction, in HP-Rare. Nucleotide diversity was calculated as pairwise diversity π and segregating sites θ in DNAsp. π(syn) = nucleotide diversity at synonymous sites. π(nonsyn) = nucleotide diversity at nonsynonymous sites. Population diversity metrics were not calculated where less than four individuals were sampled in a population, or for the managed South African sample set and captive samples, which do not represent true populations. Country codes – Kenya KNY, Tanzania TNZ, Zimbabwe ZIM, Botswana BOT, Mozambique MOZ, South Africa SAF, Namibia NAM.

Table 2.3: Percentage of DLA-DQB1 alleles across sampling localities, subdivided into free-ranging nonmanaged populations, and samples from a managed population, and carcass and captive samples.

	Non-managed populations				**Managed**	**Carcass samples**				**Captive**	**Total (n=234)**
Locus DQB1*	**Laikipia KNY (n=56)**	**N.Serengeti TNZ (n=9)**	**Hwange ZIM (n=13)**	**Okavango BOT (n=53)**	**S.Africa SAF (n=6)**	**Kajiado KNY (n=1)**	**Ghanzi BOT (n=1)**	**Sofala MOZ (n=3)**	**Mangetti NAM (n=1)**	**EU zoos (n=92)**	
90101	89.3	72.2	100.0	70.8	100.0	100.0	100.0	100.0	100.0	97.8	87.5
90201	10.7	27.8		29.2						2.2	12.5

n=number of individuals typed. Country codes – Kenya KNY, Tanzania TNZ, Zimbabwe ZIM, Botswana BOT, Mozambique MOZ, South Africa SAF, Namibia NAM.

Table 2.4: Comparison of DLA alleles found in different canid populations.

Study species/population	**DRB1**	**DQA1**	**DQB1**	**Reference**
Species sampled in multiple populations				
African wild dog: n=368	17	1	2	
[1]Grey wolf: Canada, Alaska: n=194	17	12	15	(Kennedy et al. 2007a)
Grey wolf: Northern Europe: n=163	17	9	10	(Seddon & Ellegren 2002)
Grey wolf: Total n=407	26	18	21	(Kennedy et al. 2001)
Species sampled from single populations				
[2]African wild dog: Single pop^{n} n=14-56	3-10	1	2	
[3]Mexican wolf: Single pop^{n} n<7	4	5	3	(Hedrick et al. 2000)
[4]Ethiopian wolf: Bale Mountains pop^{n} n=99	4	2	5	L.J.Kennedy *submitted*

[1] Grey wolves are not an endangered species (Mech & Boitani 2008).
[2] Numbers of alleles detected in individual African wild dog populations where 14-56 animals were sampled per population.
[3] Mexican wolves, *Canis lupus baileyi*, are a critically endangered subspecies of the Grey wolf *Canis lupus*. They are thought to have gone extinct in the wild in the c.1970. All individuals extant today have been bred in captivity, and are derived from 7 founders (Hedrick et al. 2000).
[4] Ethiopian wolves are highly endangered, with just 500 individuals. MHC surveys were conducted on 99 samples from the largest (n=250) of the seven extant Ethiopian wolf populations (LJ Kennedy *submitted*).

Chapter 3: Puzzling Persistence of African Wild Dogs in Serengeti-Mara.

3.1 Abstract

An endangered population of African wild dogs (*Lycaon pictus*) disappeared from the Serengeti-Mara in 1991. The reasons for the extinction are not understood, but disease was implicated in the disappearance. In 2001, wild dogs naturally re-colonised the region. I conducted microsatellite genotyping on samples collected prior and subsequent to this event, as well as samples from three nearby populations. Contrary to expectations, clustering analyses and assisgnment tests demonstrate that the re-established animals are derived from the Serengeti-Mara. This result shows that wild dogs must have persisted in this region after 1991, either undetected in, or outside of, monitoring areas. Further, I did not detect a decline in genetic diversity at either neutral or major histocompatibility complex loci as predicted for a founder effect associated with a re-colonisation event.

3.2 Introduction

With most species in global decline, there seems little to celebrate in endangered species conservation. One recently heralded exception is the re-establishment of endangered African wild dogs (*Lycaon pictus,* hereafter referred to as wild dogs) in the Serengeti-Mara region (Figure 3.1A) following apparent extinction in 1991 (Woodroffe 2001). This extinction was a high-profile event (Morell 1995) due to an extensive debate spanning more than 20 articles across 9 years (reviewed in Woodroffe 2001) about the cause of a disease outbreak which was implicated in the extinction (Gascoyne et al. 1993; Kat et al. 1995). Concern that human handling may have caused latent rabies to develop and spread (Burrows et al. 1994) led to a ban on animal handling and reluctance to implement vaccination programs for both domesticated and wildlife species in some countries (Woodroffe 2001). In 2001, African wild dogs naturally re-established in the Serengeti-Mara region (Fyumagwa & Wiik 2001). Given the controversial history of this population, there is considerable interest about the population origin of these individuals. Since wild dogs are a highly mobile species, with dispersers groups capable of moving up to 250 km to establish new packs (Fuller et al. 1992) there are a number of potential source populations of the Serengeti-Mara wild dogs. In Eastern Africa, wild dog populations are no longer resident in Uganda, Rwanda or Burundi, but extant populations are known in Tanzania, Kenya, Ethiopia and Sudan. However, with few exceptions these wild dog populations are highly fragmented and isolated from each other by wide stretches of anthropogenically modified habitats, and thus recolonisation is more likely from the closest populations.

Wild dog monitoring in the Serengeti-Mara formerly concentrated on two areas within the Serengeti-Mara ecosystem: 1) the “Serengeti plains” in the south from 1964 (Burrows et al. 1994); and 2) the “Mara” just outside of the Masai Mara Nature Reserve in the north from 1987 (Scott 1991). The pre-extinction Serengeti-Mara monitoring area is hereby used as a collective term describing the area covered by the home ranges of wild dog packs from these two areas (Figure 3.1B). Between 1986 and 1991, 15 packs were observed in the pre-extinction monitoring area (Woodroffe 2001). However, by December 1991, all of these packs were recorded as locally extirpated (Woodroffe 2001) and the whole Serengeti-Mara wild dog population was subsequently reported (Creel & Creel 2002; Daszak et al. 2000; Woodroffe & Ginsberg 1999) and widely assumed to be, extinct (but see Burrows et al. 1994; Ginsberg et al. 1995). Outside of the main Serengeti-Mara monitoring area, no systematic data was collected and so we do not know where and how many wild dogs were resident, nor their demographic trends (Burrows 1995). Limited monitoring conducted in the Serengeti-Mara ecosystem between 1991 and 1998 reported only vagrant and single-sex dispersing groups (Woodroffe 2001). However, following

sightings of multiple groups of wild dogs in 1998, systematic monitoring was restarted and the Serengeti-Mara population officially deemed re-established in 2001 when the first denning (reproduction) was reported (Fyumagwa & Wiik 2001). Since then, the re-established population has grown rapidly, with a minimum estimate of 125 wild dogs reported in 2009. Whilst the home ranges of re-established and pre-extinction packs overlap, the re-established monitoring area does not include the entire pre-extinction monitoring area as wild dogs have not yet (January 2010) re-established as resident packs inside of the Serengeti National Park (SNP, Figure 3.1B), despite observations of sporadic incursions of wild dogs into SNP.

Rarely are samples available prior to extinction and following natural re-colonisation of an endangered species. However, through ongoing research programs, I obtained samples from individuals residing in the Serengeti-Mara region before and after the assumed extirpation, hereafter referred to as the pre-extinction and re-established Serengeti-Mara. This sample set was typed for variation in 10 microsatellite loci along with samples from the three nearby wild dog populations in eastern Africa (Selous, Masai-Steppe and Laikipia), to assess the source of the re-established Serengeti-Mara wild dogs. We also assessed whether the re-established Serengeti-Mara population exhibited reduced levels of genetic diversity at neutral microsatellite loci and at the major Histocompatibility complex (MHC), as predicted with founder effects associated with the recolonisation.

Here I demonstrate that, despite the observed disappearance of wild dogs in the monitoring area, the declaration of extinction was premature, as genetic evidence indicates that the re-established animals are derived from the Serengeti-Mara. Encouragingly, I show that there has not been a loss of genetic diversity in the Serengeti-Mara population. However, increased monitoring is now essential to elucidate and protect the metapopulation connections that are needed to allow demographic rescue.

3.3 Methods

Sampling and DNA extraction

I obtained samples from wild dogs residing in the Serengeti-Mara before (n=20 from ≥6 packs, S. Cleaveland, P. Kat) and after (n=13 from 4 packs, E. Masenga) the assumed extirpation. I also accessed samples from three other wild dog populations in eastern Africa: Selous in southern Tanzania (n=22 from 8 packs, S. Creel); Masai Steppe in northern Tanzania (n=32 from 3 packs, A. Visée); and Laikipia in northern Kenya (n=65 from ≥9 packs, R. Woodroffe; Figure 3.1A,B), which are the three geographically closest extant populations for which samples could be obtained. It was not possible to ascertain samples for all possible wild dog source populations, those that were not sampled are shown in Figure 3.1A.

All samples assessed in this study were derived from high quality samples, that is, blood, sera or tissue. Masai Steppe samples were provided as extracted DNA. Six Selous samples and four pre-extinction Serengeti samples had been extracted using phenol chloroform for another study DNA (Girman et al. 2001). All other samples were extracted using DNeasy tissue and blood extraction kits (Qiagen Inc) according to the manufacturer's instructions.

Microsatellite and MHC genotyping

DNA samples were genotyped at 10 microsatellite loci PEZ08, PEZ12, PEZ15 (J. Halverson in Neff *et al*. 1999); FHC2010, FHC2054, FHC2611, FHC2658, FHC2785, FHC3399, FHC3965 (Breen et al. 2001; Guyon et al. 2003; Neff et al. 1999) located on different chromosomes (Neff et al. 1999). The forward primer of each pair was labelled with ABI fluorescent dyes: NED (yellow), 6-FAM (blue) or HEX (green). Samples were amplified alongside negative controls by multiplex PCR using Qiagen Multiplex PCR mix. I followed default reagent concentrations recommended by the manual except in cases of DNA derived from serum or hair, where 0.4µl of 10mM Bovine Serum Albumin (Promega) was added per PCR reaction. PCR was performed on PTC-200 (MJ Research) theromcyclers with the following touchdown protocol: 15 min at 95°C; 12 touchdown cycles of 94°C for 30 s; 1 min at 30 s annealing, starting at 60°C and reducing at 0.5°C per cycle; and 72°C for 1 min. This was followed by 33 cycles of 89°C for 30 s, 55°C for 1 min, and 72°C for 1 min. The protocol ended with a final extension of 60°C for 30 minutes. Samples were run on an ABI 3730 with a ROX 500 size standard (by The Sequencing Service, University of Dundee) and analysed using GENEMAPPER 4.7 (Applied Biosystems). Low concentration or poorly amplifying DNA samples were amplified and genotyped three times. Samples with missing data for more than three loci

were excluded from analyses. For each population and locus, I tested for deviations from Hardy Weinberg Equilibrium using GENALEX6 (Peakall & Smouse 2006) and assessed for significance after Bonferroni correction for multiple tests.

Sequence-based typing of the Major Histocompatibility Complex (MHC) DLA-DRB1 locus (hereafter referred to as DRB) was done according to the methods detailed in Marsden et al. (2009) which is included in this thesis as Chapter 2. In brief, I amplified and sequenced exon 2 of the DRB locus. Sequences were visually analysed to correct automated base calling errors in Match Tools Navigator (Applied Biosystems). Homozygous samples were used to compile locus-specific allele libraries, which also included all other known canid alleles obtained from the DLA nomenclature committee (L. Kennedy *pers.comm*). Heterozygous samples were then analysed in Match Tools (Applied Biosystems), which compares heterozygous sequences with the allele library to predict which allele combinations could generate the observed pattern (Kennedy et al. 2002b).

Microsatellite clustering analyses

A neighbour-joining tree based on Nei's allele-sharing distance was reconstructed in POPULATIONS v 1.2.30 (Langella 1999). Bayesian clustering analysis was conducted using STRUCTURE 2.3 (Pritchard et al. 2000), assuming no prior population or location information, with correlated allele frequencies and admixture. I used 100,000 burn-in cycles and 500,000 Markov Chain Monte Carlo (MCMC) runs for K = 1-10, with ten replicates per K value. The most likely number of clusters (K) that best fit the data was selected based on the ΔK statistic (Evanno et al. 2005) and consistency amongst replicates (Pritchard et al. 2000). The ΔK statistic assesses the rate of change in the log probability of the data between successive K values (Evanno et al. 2005), and was calculated in the programme STRUCTURE HARVESTER v0.5 (Earl 2009). Following Evanno *et al.* (2005) I first ran STRUCTURE on our complete data set, and subsequently on data subgroups as indicated by ΔK.

Assignment tests

I conducted posterior probability assignment tests on the re-established Serengeti-Mara samples in STRUCTURE 2.3. I ran STRUCTURE at K=4, assuming correlated allele frequencies, admixture, a migration rate of 0.01 (Pritchard et al. 2000), 100,000 burn-in cycles and 500,000 MCMC runs. An assignment test was also conducted in Geneclass 2 (Piry et al. 2004) using the Rannala & Mountain criterion (Rannala & Mountain 1997) and Paetkau resampling algorithm (Paetkau et al. 2004), assessed over 10,000 simulations.

Microsatellite and MHC genetic diversity and heterozygosity

For the Serengeti-Mara samples I calculated observed (H_o) and expected (H_e) heterozygosity in GenALEX (Peakall & Smouse 2006) and the number of alleles (*A*) and allelic richness standardized for sample sizes (*Rs*) in FSTAT 2.9.3 (Goudet 1995) for both microsatellite and MHC alleles. I also calculated nucleotide diversity amongst MHC alleles as the average number of segregating sites θ and pairwise diversity π, using DnaSP 4.20 (Rozas and Rozas 1995).

3.4 Results

Clustering analyses

An allele sharing tree showed that the African wild dog samples clustered into four distinct geographic groups: 1) Laikipia; 2) Selous; 3) Masai-Steppe; and 4) pre-extinction and re-established Serengeti-Mara (Figure 3.1C). Similarly, STRUCTURE analyses were most consistent with four clusters within the data set (Figure 3.2). The ΔK statistic indicated the strongest signal of population subdivision to be K=2, where Laikipia was distinct from all other samples (Figure 3.2). However, there was a strong secondary peak at K=4 indicative of finer scale population structuring (Evanno et al. 2005), where the pre-extinction and re-established samples were identified as a single cluster, and Laikipia, Masai Steppe and Selous were each assigned to separate clusters. It is noteworthy that only when K≥7, did the re-established Serengeti-Mara sometimes appear to form a cluster that was distinct from the pre-extinction Serengeti, but this was not a consistent solution across replicates. As recommended by Evanno (2005) I re-ran STRUCTURE on each of these two clusters; 1) Selous, Serengeti-Mara and Masai-Steppe and 2) Laikipia. For cluster 1, the results showed a clear single peak at K=3 corresponding to the pre-extinction and re-established samples as a single cluster, with the other two clusters defined by Selous and Masai Steppe samples (Figure 3.2). Independent analysis of Laikipia subdivided the population into groups of related packs (data not shown).

Assignment tests

Based on microsatellite loci, posterior probability assignment tests implemented in STRUCTURE showed that individuals from the re-established Serengeti-Mara population had on average a 95% probability of belonging to the same population as the pre-extinction animals (range 85-97%), compared with 1.4-2.4% probability of belonging to any of the other putative source populations. The more conservative assignment test that allows for unsampled populations, implemented in Geneclass 2, assigned nine of the 13 re-established Serengeti-Mara animals to the same population as the pre-extinction individuals. The remaining four samples were assigned to unsampled populations. Given the limited geographic sampling of pre-extinction animals, this could suggest they are from unrepresented areas of the Serengeti-Mara and thus appear different despite being from the same genetic population. Conversely, they could be migrants from unsampled populations.

Genetic diversity

I observed no significant decline in H_o, H_e, or *Rs* at microsatellite loci between the pre-extinction and re-established populations (Table 3.1: paired T-test; H_o, t=0.72, p=0.486; H_e, t=1.90, p=0.09; *Rs*, t=1.70, p=0.123). In general, levels of genetic diversity were comparable to the three other wild dog populations (H_o, 0.61-0.68; H_e, 0.61-0.68; *Rs,* 4.31-4.9; Table 3.1) as well as other wolf-like canid populations (Aspi et al. 2009; Aspi et al. 2006; Randall et al. 2010). No significant deviations from Hardy Weinberg equilibrium at any microsatellite locus were observed (GENALEX).

I amplified nine different MHC class II DRB alleles (Table 3.2). I found no evidence of pseudogenes (stop codons or frameshift mutations), and a maximum of two alleles were amplified in each individual, suggesting a single functional gene was amplified. As with the microsatellite data, I found no evidence of decline in MHC diversity between the pre-extinction and re-colonised populations (Table 3.1). In fact, H_o, H_e and *Rs* were all higher in the re-colonised population. However, there was shift in allelic composition and frequency between pre-extinction and re-colonised samples (Table 3.2). The pre-extinction and re-colonised populations shared just 2 alleles, and allele DRB1*90601 that was found at high frequency (34%) in the pre-extinction population was absent in the re-colonised population. Further, allele DRB1*90203 which was common (43%) in the re-colonised population was not found in the pre-extinction population. African wild dog DRB alleles have been shown to be derived from just two allelic lineages, A and B (Marsden et al. 2009). I observed that the re-colonised population had a significantly higher (36%) proportion of lineage A alleles than the pre-extinction population (Chi Squared, d.f.=1, p<0.005, Chi sq value = 8.288). The re-established population also had a higher proportion (50%) of individuals with an allele from each lineage (i.e. lineage heterozygotes) than the pre-extinction population (22%), although the difference was not significant (Chi-Sq=0.90, d.f. =1).

3.5 Discussion

Contrary to expectations, both Bayesian and allele-sharing clustering analyses showed that re-established Serengeti-Mara wild dogs grouped with pre-extinction individuals rather than putative source populations (Figure 3.1C,D). Furthermore, assignment tests indicated that the majority of re-established Serengeti-Mara individuals were derived from the same population as the pre-extinction Serengeti-Mara. Consequently, these findings strongly imply that, although wild dogs were no longer observed to be resident in the well-monitored areas of Serengeti-Mara after 1991, they may have persisted there undetected. Alternatively, they may have existed in unmonitored areas nearby, providing a source of individuals for re-establishment in 2001. It has been widely assumed that all former residents (15 packs) in the pre-extinction monitoring area died; however, the fates of most are unknown (Woodroffe 2001). Passive monitoring by scientists, tour guides, and indigenous communities suggests an absence of breeding packs from the pre-extinction monitoring area until at least 1998. However, these monitoring efforts may not have been sufficient to detect a population at low density in Serengeti-Mara or nearby areas. Specifically, the re-established animals I sampled clustered most closely with pre-extinction animals from along the Kenyan-Tanzanian border, where monitoring had been minimal (Figure 3.1C). However, genetic data alone cannot provide information on the puzzling cause of the disappearance of so many packs in Serengeti-Mara, nor on reasons for the subsequent recovery. More extensive monitoring would have been required to evaluate the causes of their disappearance. Interestingly, although the population is recovering rapidly (minimum estimate of 125 dogs), to date no packs have re-established in the Serengeti National Park (SNP), where much of the pre-extinction monitoring was focused (Figure 3.1B). The reason for the absence from SNP remains to be determined but wild dogs are known to avoid lions (Mills & Gorman 1997), which have increased inside SNP since 1991 (Packer et al. 2005).

It is rare to be able to compare genetic samples before and after a local demographic decline in an endangered species and the availability of samples here emphasises the importance of continuous long-term field projects such as in the Serengeti-Mara (Thirgood 2007). Although our sample size is small, I found no evidence of a loss of genetic diversity in neutral microsatellites or genes that should be subject to selection (MHC-DRB) suggesting that the loss of packs in the Serengeti-Mara did not impact genetic variation of the regional metapopulation (Table 3.1). However, at the DRB locus I detected a large increase in H_o as well as a shift in the composition and frequency of alleles and lineages between pre-extinction and re-colonised samples. Due to the functional relevance of the MHC locus, these changes could represent adaptive differences between the pre-extinction and re-established study populations resulting from

changes in selective pressures. Indeed, allelic composition and heterozygosity were more consistent at neutral loci than at the DRB: 79% of microsatellite alleles detected in the re-established population were found with pre-extinction population compared with 33% at the MHC; H_o in the re-established population was 4% lower at microsatellite loci but 26% higher at MHC loci. However, further study is required as these changes may be a stochastic consequence of limited sample sizes.

Our study highlights the importance of geographic connectivity in small populations of highly mobile species and emphasizes the need for more, rather than less, monitoring so that source populations and dispersal corridors can be identified and adequately protected. This monitoring may require animal handing, as radiotelemetry devices offer the only direct method to follow the movements of individuals (Woodroffe 2001). In recent years, several other wild dog populations have likewise been re-established by natural re-colonisation (Woodroffe et al. 2005). Such demographic rescue events illustrate the critical importance of maintaining landscape connectivity even when the likelihood of re-colonisation appears small or when local extinctions seem to have occurred.

Figure 3.1: (A) Map of sampling locations. Green = Laikipia (LAI), blue = Serengeti-Mara (S-M), yellow = Masai Steppe (MST), red = Selous (SEL). Numbers 1-5 represent populations that were not sampled in the area surrounding Serengeti-Mara. (B) Approximate area of Serengeti-Mara Pre-extinction (dark blue) and Re-established (light blue) monitoring areas. (C) Population subdivision based on neighbour-joining tree of allele-sharing distance across 10 microsatellite loci. Re-established Serengeti samples (brown) clearly cluster with pre-extinction Serengeti samples rather than with the three other populations, which each appear as distinct clusters. Asterisks depict pre-extinction Serengeti-Mara samples from the northern Mara area. (D) Population structure based on STRUCTURE analyses of 10 microsatellite loci. Shown is the most likely level of clustering (K=4) (Figure 3.2). Columns are individuals, with the proportion of an individual's genotype assigned to each cluster (K) denoted by different colours. The re-established Serengeti-Mara samples (S-M, R) clearly cluster with the pre-extinction Serengeti-Mara samples (S-M, Ex) rather than the three other populations, which appear as discrete clusters.

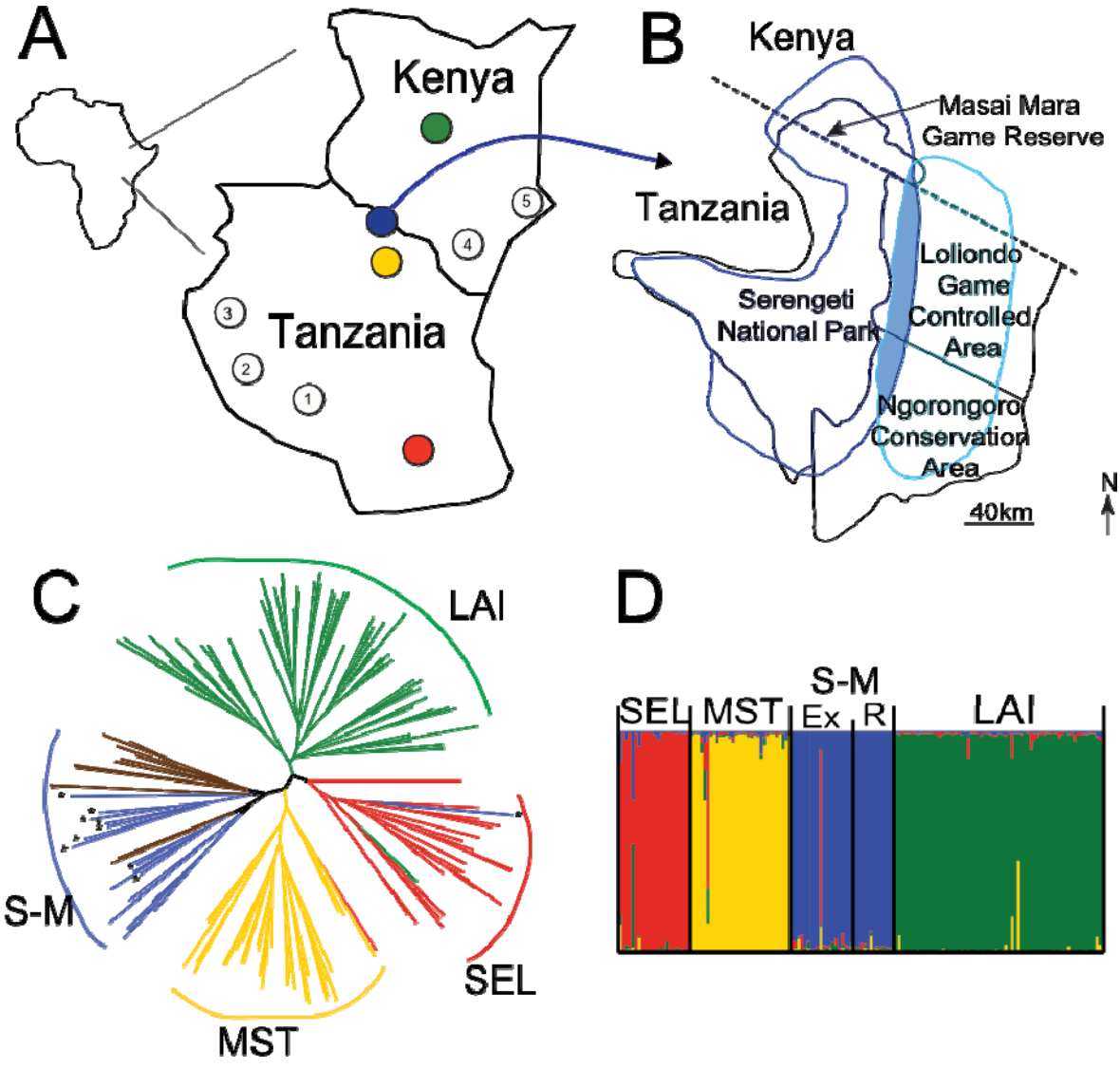


Figure 3.2: Magnitude of Δ*K* as a function of *K*. a) Where all samples are included the highest level of structure (*K*=2) identifies Laikipia to be distinct to all other samples. There was a prominent lower peak at *K*=4, where the pre-extinction and re-established samples are identified as a single cluster, and Laikipia, Masai Steppe and Selous as three separate clusters. b) Where the divergent Laikipia samples are excluded, a single peak at *K*=3 is identified, which corresponds to the pre-extinction and re-established samples being a single cluster, with the other two clusters being Selous and Masai Steppe.

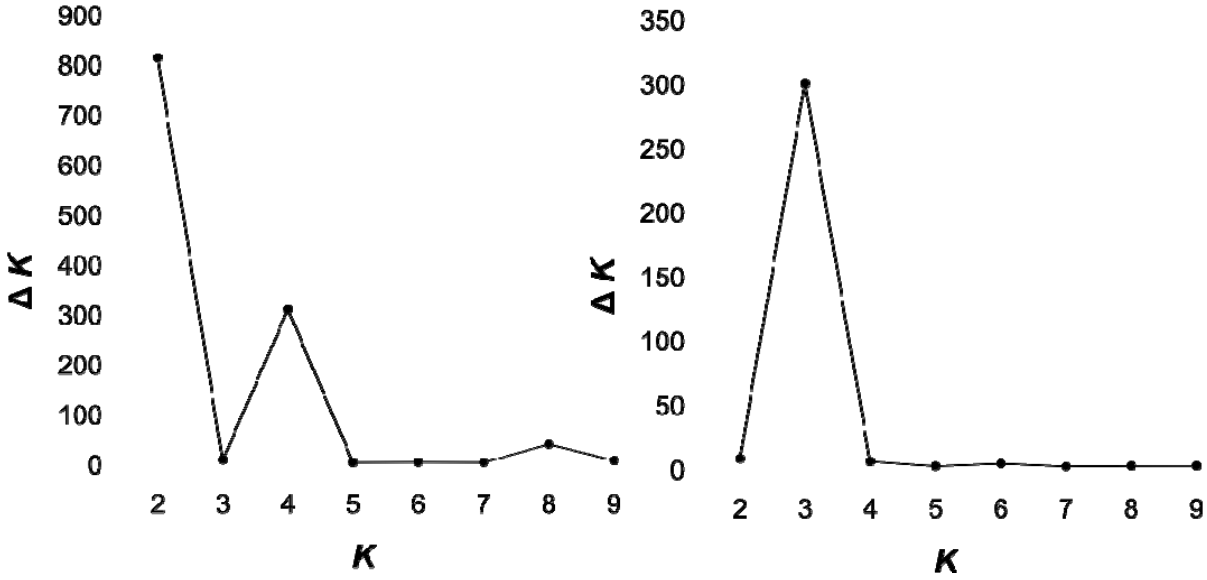


Table 3.1: Genetic diversity and heterozygosity estimates for samples based on microsatellite (msat) and MHC markers, for samples from the pre-extinction and re-established Serengeti-Mara.

		Microsatellites				MHC-DRB					
Sample source	N (msat/MHC)	A[1]	Rs[2]	Ho[3]	He[4]	A	Rs	Ho	He	π (SD)	θ (SD)
Pre-extinction	20/18[5]	6.10	4.97	0.69	0.72	5.00	4.8	0.67	0.74	0.064 (0.028)	0.053 (0.028)
Re-established	13/14[6]	4.60	4.31	0.65	0.67	6.00	6.0	0.93	0.76	0.060 (0.009)	0.049 (0.024)

[1] Mean number of alleles per locus.
[2] Allelic richness standardized for differences in sample sizes.
[3] Observed Heterozygosity.
[4] Expected Heterozygosity.
[5] Two pre-extinction samples would not amplify or produce readable sequence at the DRB locus.
[6] One re-established sample failed to amplify across microsatellite loci.

Table 3.2: Frequency of MHC-DRB alleles and lineages in the Pre-extinction and Re-established Serengeti-Mara.

	Allele	Allele frequency Pre-extinction	Re-established
Lineage A alleles	DRB1*90201	0.03	
	DRB1*90202		0.14
	DRB1*90203		0.43
	DRB1*90204	0.36	0.18
Lineage B alleles	DRB1*90401	0.11	
	DRB1*90601	0.33	
	DRB1*90602		0.04
	DRB1*90801	0.17	0.18
	DRB1*91101		0.04
Lineages %A:%B		39:61	75:25

Chapter 4: Demographic processes determine patterns of genetic diversity across African wild dog populations.

4.1 Abstract

African wild dogs *Lycaon pictus* are an endangered canid that has suffered extensive declines in both former geographic ranges and population numbers. Current populations, with few exceptions, are small and exist in locations isolated from each other by wide stretches of anthropogenically modified habitats. In this study I combined neutral (microsatellite, mtDNA) and adaptive (MHC) markers to elucidate demographic history, gene flow, evidence of selection and spatial and temporal patterns of genetic diversity across wild dog populations from Eastern and Southern African. All wild dog populations were found to be small (N_e<30) and showed evidence of bottlenecks. Coalescent models detected a genetic signature of a large and recent demographic decline in wild dogs, which correlates with human expansion, but contrasts with the demographic history of other African mammals. Habitat fragmentation and loss appears to have resulted in strong population structuring of wild dog populations, with limited gene flow between them. The spatial and temporal structure of microsatellite and MHC diversity were correlated, and appeared to be largely determined by demographic stability and size of populations. This suggests that selection may be unable to counter strong genetic drift in these small wild dog populations. Overall, I found that the predominant factor determining patterns of both neutral and adaptive genetic variation in wild dogs is demographic history.

4.2 Introduction

Knowledge of patterns of genetic variation within a species is essential to the understanding of population structure, local adaptation and differences in levels of diversity between populations (Bos et al. 2008). Furthermore, in conservation genetics, this information is critical for setting conservation priorities, identifying management units and guiding translocation strategies. 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 as a result of higher rates of genetic drift and inbreeding, and lower rates of geneflow (Frankham 1996; Templeton et al. 1990). Neutrally evolving genetic markers, such as mitochondrial DNA (mtDNA) and microsatellites, are the most appropriate for elucidating demographic change and genetic structure and therefore have been widely implemented in population genetic studies (Bos et al. 2008). However, selection is a dominant force in shaping adaptive genetic variation, which forms the basis of evolutionary change and local adaptation (Gebremedhin 2009; Hoglund 2009). Since neutral genetic markers are not subject to selection, adaptive markers must be assessed in order to understand patterns of adaptive variation and the forces that govern it.

As our knowledge of loci under selection improves, adaptive markers are increasingly being incorporated into studies, alongside neutral loci, to elucidate local adaptation and differences in evolutionary potential between populations e.g. (Campos et al. 2006). The major histocompatibility complex (MHC) includes multiple genes that code for a set of cell-surface molecules involved in the recognition of intra- (class I) and extra- (class II) cellular protein antigens as part of the immune response (Eggert et al. 1998; Klein 1980; Piertney & Oliver 2006). MHC loci have been shown to have extraordinarily high levels of variation and this diversity is thought to be maintained by balancing selection (reviewed in Garrigan & Hedrick 2003). There are a large number of studies showing associations between specific MHC alleles and susceptibility or resistance to specific diseases (reviewed in Hill 1998; Piertney & Oliver 2006), which supports the contention that pathogens are the main selective force that maintains variation at the MHC (reviewed in Jeffery & Bangham 2000; Spurgin & Richardson 2010). As one of the most well understood adaptive loci (Bernatchez & Landry 2003; Miller et al. 2001), and with clear relevance to population viability and evolutionary ability (Hoglund 2009; Piertney & Oliver 2006; Siddle et al. 2007), the MHC is currently one of the best markers available to use as a proxy for adaptive genetic variation.

When designing studies of genetic variation it is important to consider that selective and neutral forces vary both spatially and temporally, and thus genetic variation

is also predicted to vary (Oliver et al. 2009a; Seddon & Ellegren 2004). Under directional selection, for example, local selection pressures are expected to result in genetic divergence in MHC variation between populations subjected to different pathogen repertoires (Bryja 2007; Mäkinen et al. 2008; Miller et al. 2001; Vassilakos et al. 2009). Furthermore, changes in pathogen communities are predicted to result in changes in MHC variation, within a population, over time (Charbonnel & Pemberton 2005; Oliver et al. 2009a; Westerdahl et al. 2004). Despite spatial and temporal variability, many studies that have assessed neutral and MHC variation have been based on a set of samples from a small number of populations at a single point in time (e.g. Bos et al. 2008; Campos et al. 2006; Charbonnel & Pemberton 2005; van Oosterhout et al. 2006). 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 an ecologically appropriate spatial scale for the species in question (e.g. Crandall 2009; Koutsogiannouli et al. 2009) and from more than one temporal period (e.g. Demandt 2010; Oliver et al. 2009a; Westerdahl et al. 2004), but this has rarely been done in tandem.

Such detailed studies are particularly important in conservation genetics, where anthropogenic declines and habitat loss are altering the neutral and selective forces that shape patterns of genetic diversity of endangered species. Population declines and fragmentation are predicted to result in isolated populations with small effective population sizes; characteristics that reduce the efficacy of selection on maintaining adaptive diversity and increase the loss of genetic diversity by genetic drift and inbreeding (Charlesworth 2009; Crow & Kimura 1970; Frankham et al. 2002; Kimura 1983). Understanding the extent of demographic declines, their impact on selection, genetic diversity and population structuring, as well as determining which populations are adaptively different or suffering genetic threats, is critical for prioritising conservation efforts and reducing risk of species-wide extinctions. However, the level of sampling required for such an extensive study is especially challenging in endangered species, where sample sizes are inherently limited.

The African wild dog, (*Lycaon pictus*) is a highly mobile, social, wolf-like canid that hunts and breeds cooperatively in packs averaging 5-15 adults (Creel & Creel 2002). Historically, African wild dogs ranged across most of sub-Saharan Africa (Woodroffe et al. 2004b). However, dramatic range reductions resulting from extensive habitat loss and persecution mean that they now occupy just 7% of their former range (IUCN/SSC 2008, 2009; Woodroffe et al. 1997). In the wild, fewer than 8,000 individuals remain, scattered across a small number of fragmented populations which are largely isolated from each other by wide stretches of anthropogenically modified habitats. (IUCN/SSC 2008, 2009). The small sizes of remnant populations, only nine of which are known to constitute more than 200 animals, make them vulnerable to extinction (Woodroffe et al. 1997; Woodroffe

et al. 2004b). Disease is argued to represent a significant threat to African wild dogs, which share susceptibility to diseases of common sympatric canids such as jackals and domestic dogs (Alexander et al. 2010). Outbreaks of diseases have resulted in significant population declines in African wild dogs in the past (reviewed in Woodroffe et al. 2004a). Consequently, knowledge of the MHC is particularly pertinent to African wild dog conservation.

The aim of my study was to assess the forces that shape patterns of genetic variation in the endangered African wild dog, by assessing a spatially and temporally variable set of samples with a combination of both neutral (microsatellite and mitochondrial DNA) and adaptive (MHC) markers. Considering the endangered status of this species, my sample set is somewhat unusual in terms of: 1) the number of samples (>350); 2) their spatial scale (13 monitoring areas distributed throughout Eastern & Southern Africa); and 3) temporal separation (three populations were sampled at two time points). Specifically, I addressed the following questions: 1) Is there a genetic signature of demographic decline in African wild dogs? 2) How are African wild dog populations currently structured? 3) How are neutral and MHC diversity in African wild dogs structured temporally and spatially? 4) Is there evidence of selection at the MHC and/or local adaptation of African wild dog populations?

4.3 Methods

4.3.1 Sampling and DNA extraction

To access genetic samples, I contacted established free ranging Africa wild dog field projects in Eastern and Southern Africa. As far as I am aware, my sample set represents almost all free ranging wild dog samples available in this region (excluding faecal and museum samples). In three monitoring areas (Kruger, Okavango and Serengeti), samples were available from two temporal periods, which will be referred to as "Old" and "Recent" (Old = 1980-1995; Recent = post 2000; Table 4.1; Appendix 2). Blood, tissue, hair and serum samples were collated from thirteen monitoring areas (Figure 4.1; Table 4.1; Appendix 2): 1) Kruger; 2) Lowveld; 3) Okavango; 4) Hwange; 5) Selous; 6) Masai Steppe; 7) Serengeti-Mara, 8) Laikipia. Five or fewer samples were available from the remaining monitoring areas: 9) Ghanzi; 10) NE-Namibia; 11) Sofala; 12) Niassa; 13) Kajiado. For this reason, these samples were only used in descriptive analyses of the distribution of mtDNA haplotypes and MHC alleles, where these additional five monitoring areas improved geographic coverage. Finally, I sampled 16 wild dogs that were translocated into Hwange (six moved from Pilansberg South Africa and ten moved from within Zimbabwe) to assess the genetic impact of artificial translocations.

For samples collected prior to 1997, DNA had been extracted for another study using phenol chloroform extractions (Girman et al. 2001). Samples collected post 1997 were extracted using DNeasy tissue and blood extraction kits (Qiagen). Blood, serum and tissue samples were extracted according to manufacturer's instructions. Hair was extracted according to a user-developed protocol available from Qiagen (Qiagen 2006).

4.3.2 Genetic typing

Mitochondrial haplotyping

To enable comparison with previous work, I amplified a 327 bp segment of mitochondrial mtDNA D-Loop control region 1 overlapping the region assessed by Girman et al (2001). Mitochondrial DNA was amplified by polymerase chain reaction (PCR) using a modified version of the canid-specific primers, Thr-L and DLH, which were redesigned specifically for African wild dogs (Leigh 2005): forward 5' ACT ATT CCC TGA TCT CCC CC 3'; reverse <u>CAG GAA ACA GCT ATG ACC</u> CCT GAA GTA AGA ACC AGA TGC C. The underlined section of the reverse primer marks an M13 tag, which was used to permit sequencing in a single direction. These primers overlap the 381 bp mtDNA segment assessed by Girman, beginning at bp 93 (1st variable site is bp 171) and extending an

additional 22 bp at the 3' end. PCRs were performed in a 20-µl reaction volume containing: 1.25 x Q solution (Qiagen); 1.25 x PCR buffer (containing 15mM $MgCl_2$); 3.1 mM MgCl; 0.2 mM of each dNTP (Invitrogen); 0.19 µM of each primer; 1 unit of Hot Star *taq* (Qiagen); and approximately 10ng of template DNA (except for negative controls). PCR was conducted according to the following protocol: 5 min at 95°C, 30 cycles of 95°C for 30s, 55°C for 30s, and 72°C for 30s. The protocol ended with a final extension of 72°C for 10 minutes. The number of amplification cycles was increased from 30 to 37 cycles for weak DNA samples derived from hair, blood spots and serum. PCR products were cleaned with ExoSAP-IT (USB) according to the manufacturer's instructions and then sequenced using an ABI 3730 (using The Sequencing Service, University of Dundee or The Genepool, University of Edinburgh). Sequences were aligned and analysed using Geneious Pro v 4.5.5 (Biomatters Ltd).

Some samples included in my study had been manually sequenced at this same mtDNA region by Girman et al. (2001) but I re-sequenced all available samples where DNA and corresponding haplotype information were available (n=116). This was done to confirm the haplotype designations using automated fluorescent sequencing. However, I included all of the data (corrected for errors that I detected) presented in Girman et al. (2001), so the sample sizes for mtDNA are larger than those listed for microsatellite and MHC loci since DNA was not available from all of the samples they described. In total I sequenced 192 samples and combined my results with published data from a further 274 samples from previous studies (Girman et al. 2001).

Microsatellite genotyping

I screened 25 domestic dog microsatellite loci for amplification and polymorphism in wild dogs and selected ten loci based on signal quality, polymorphism, and allele range sizes for multiplexing: PEZ08, PEZ12, PEZ15 (J. Halverson in Neff *et al*. 1999); FHC2010, FHC2054, FHC2611, FHC2658, FHC2785, FHC3399, FHC3965 (Guyon et al. 2003; Neff et al. 1999); Breen *et al*. 2001; details in Appendix 3). All loci selected were located on different chromosomes (Neff et al. 1999). The forward primer of each pair was dye-labelled with ABI fluorescent dyes: NED (yellow), 6-FAM (blue) or HEX (green). Samples were amplified alongside negative controls by multiplex PCR using Qiagen Multiplex PCR mix. I followed default reagent concentrations recommended by the manual except in cases of DNA derived from serum, hair and blood spots, where 0.4ul of 10mM Bovine Serum Albumin (Promega) was added per PCR reaction. PCR was performed on PTC-200 (MJ Research) theromcyclers with the following touchdown protocol: 15 min at 95°C, 12 touchdown cycles of 94°C for 30 s, followed by 1 min 30 s annealing, starting at 60°C and reducing at 0.5°C per cycle, and 72°C for 1 min. This was followed by 33 cycles of 89°C for 30 s, 55°C for 1 min, and 72°C for 1 min. The protocol ended with a final

extension of 60°C for 30 minutes. Samples were run alongside a ROX 500 size standard on an ABI 3730 (by The Sequencing Service, University of Dundee) and analysed using GENEMAPPER 4.7 (Applied Biosystems). Weak DNA samples have a higher probability of allelic drop out; therefore, I amplified and genotyped DNA samples derived from hair, blood spots and serum three times. I also re-amplified and genotyped a further 20% of blood and tissue samples to verify results. Where a unique allele was found in a single animal, that animal was genotyped twice. Samples with missing data for more than three loci were excluded from analyses. For each population and locus, I tested for deviations from Hardy Weinberg Equilibrium using GENALEX6 (Peakall & Smouse 2006) and assessed for significance after Bonferroni correction for multiple tests.

MHC-DRB typing

Sequence-based typing was conducted on exon 2 of the DLA-DRB1 locus (hereafter referred to as DRB), which was previously shown to be highly variable in African wild dogs (Marsden et al. 2009). Three other MHC class II loci (DLA-DQB1, DLA-DQA1 and DLA-DRA), found to be monomorphic or biallelic in African wild dogs (Marsden et al. 2009), were not assessed here. The DRB locus was typed according to Marsden et al. (2009), which is included as Chapter 2. In brief, DRB sequence data were analysed using Match Tools and Match Tools Navigator (Applied Biosystems), as described in Kennedy et al (2002b). This method relies on an allele library built from homozygotes that is used to predict the most likely allelic combinations present in a heterozygous sequence. Twenty heterozygous individuals did not match any pair of known alleles, indicating the presence of new alleles. Therefore, I cloned products from seven animals using the TOPO TA cloning system and One Shot Competent cells (Invitrogen). In total I identified six new alleles, three from cloned individuals as well as three alleles in homozygotes, which resolved all twenty heterozygous sequences. All new alleles were submitted to the DLA nomenclature committee to be assigned official names, and to confirm that the alleles were unique to wild dogs. To check for allelic drop out I conducted a second round of PCR and sequencing for all homozygote samples derived from serum, blood spots, hair and degraded tissue, and for approximately 90% of homozygotes derived from high quality blood and tissue samples. There was a 20% allelic drop out rate in the weak samples (5/25), whereas no drop out was detected in the high quality samples (0/61).

It is common for closely related alleles at a MHC locus to be classified into lineage classes (Ditchkoff et al. 2005). African wild dog DRB sequences have previously been shown to belong to two lineages (A & B) according to amino acid sequence similarity and phylogenetic analyses (Marsden et al. 2009). Therefore, I refer to the DRB alleles as A1-9 and B1-13, and have conducted some analyses based on lineage rather than allelic data. For official allele names see Appendix 1.

4.3.3 Summary statistics

Population level mtDNA diversity was calculated as the number of haplotypes and nucleotide diversity (π) in Arlequin v3.11 (Excoffier 2006). Microsatellite and DRB diversity were measured as allelic richness (A_R), standardised allelic richness ($Std\text{-}A_R$; i.e. Mean number of alleles/locus standardised to the smallest sample size per locus, n=8) calculated using rarefaction in HP-RARE (Kalinowski 2005), and observed (H_o) and expected heterozygosity (H_e), fixation index (F_{IS}) as calculated in GENALEX6 (Peakall & Smouse 2006). Permutation tests were carried out in GENETIX (Belkhir et al. 2004) to test whether F_{IS} values deviated significantly from zero. To assess whether populations were differentiated in terms of their DRB alleles, I computed pairwise Fisher's exact tests in GENEPOP (Raymond & Rousset 1995) which tests for homogeneity in DRB allele frequencies between populations. Significance values were adjusted according to the Bonferroni correction for multiple comparisons.

4.3.4 Effective populations size and demographic history

Contemporary estimates of effective population size (N_e) were calculated in NeEstimator v1.3, using two single sampling point methods (Peel et al. 2004). The linkage disequilibrium (LD) method (Hill 1981) tests for evidence of linkage disequilibrium between alleles at different loci arising as a result of increased genetic drift at smaller effective population sizes. The heterozygote excess method (Pudovkin et al. 1996), is based on the presence of binomial sampling error in allele frequencies between the sexes when few individuals breed, which results in excess heterozygosity (Schwartz et al. 1998). I also used a temporal based method (moments based approach; (Waples & Yokota 2007) for Kruger, Okavango and Serengeti, where I had temporally separated samples. This method assesses changes in allele frequencies between generations as a result of genetic drift increasing as N_e decreases.

I tested for evidence of recent bottlenecks in microsatellite data using two methods implemented in BOTTLENECK 1.2.02 (Piry et al. 1999), as well as the M ratio test implemented in the programme M P Val (Garza & Williamson 2001). These tests are based on the assumption that, in small bottlenecked populations, higher rates of genetic drift increase the loss of rare alleles. Simulations have shown that bottleneck tests are sensitive to sampling (Garza & Williamson 2001; Luikart et al. 1998; Piry et al. 1999); therefore, I only ran these tests on populations with 30 or more samples (Kruger, Okavango, Masai Steppe, Serengeti and Laikipia). The first method in BOTTLENECK tested for a shift in allele frequency distributions that is predicted to occur in bottlenecked populations as rare alleles are lost quicker than common alleles (Luikart et al. 1998). Secondly, I tested for an excess of heterozygosity arising as a result of the loss of rare

alleles during a bottleneck using a Wilcoxon sign-rank test (Cornuet & Luikart 1996). For the excess heterozygosity test, I used the two-phase mutational model (Di Rienzo et al. 1994), a variance of 12 as recommended by Piry (1999), 1000 simulations, and varied the percentage of mutations that change in a step wise manner between 70-90%. Significance was adjusted using Bonferroni correction. I also tested for evidence of bottlenecks by assessing allele frequency distributions using the M ratio test (Garza & Williamson 2001). The M ratio is calculated as the weighted frequency of alleles (k), divided by the overall range of allele sizes (r). Bottlenecks result in a reduced M ratio because the loss of any allele reduces k, whereas r only decreases when alleles at the extremes of the range are lost; therefore, k decreases at a faster rate than r (Garza & Williamson 2001). I ran the program using the values suggested by the authors (Garza & Williamson 2001): proportion of one-step mutations, p_s = 90%; average size of multi-step mutations, Δ_g = 3.5. I tested a range of θ values (0.2, 0.6, 1.2 and 2.0), which correspond to pre-bottleneck effective populations sizes of 100, 300, 600 and 1,000, respectively. The mutation rate, μ, in the M ratio test is fixed in the programme to be 5 x 10^{-4} per locus per generation, which is lower than the estimated canid microsatellite mutation rate (Francisco 1996), and therefore results in a conservative test. I used 10,000 simulations for each run and assessed significance after Bonferroni correction. M ratio tests are able to detect older bottlenecks than the other tests as the M ratio takes longer to recover than heterozygosity excess. This is because new rare alleles may not increase the M ratio but will always increase the metrics in the other two methods (Garza & Williamson 2001). For mtDNA haplotypes, I tested for signatures of demographic expansion or bottlenecks using mismatch distributions, Tajima's D and Fu's F statistic, implemented in the programme ARLEQUIN (Excoffier 2006). I tested for departures from equilibrium using 10,000 simulations.

I further investigated evidence of demographic changes in African wild dog populations using the Bayesian coalescent hierarchical model-based approach implemented in the programme 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 parameters: current effective population size (N_0), ancestral population size (N_1), time since the change in demographic change (T), and allele size distribution for microsatellites assuming stepwise mutations at rate μ. All parameter prior distributions were log normal. I conducted this analysis on Kruger (old), Okavango (now) and Selous data sets; all other populations were excluded from this analysis due to small sample sizes or a known recolonisation history. All runs were conducted assuming an exponential demographic model. I used generation times of 6.2 (Kruger), 4.5 (Okavango), and 5.4 (Selous) reported in Creel (2004) and conducted runs with wide uninformative priors, as suggested by Goossens et al. (2006). For each population I ran 5 chains with different starting points, 50,000 updates and a thinning

interval of 50,000. MSVAR output was assessed in the BOA package in the programme 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 1998). The last 25,000 updates of each of the five chains were then combined to calculate the lower (5%), median (50%) and upper (95%) quantiles of the posterior distributions of the parameters N_0, N_1, T (Zhang et al. 2007).

4.3.5 Spatial patterns of genetic diversity

To compare with previous work (Girman et al. 2001) the geneology of African wild dog mtDNA haplotypes was reconstructed in MrBayes v3.1.2 (Huelsenbeck & Ronquist 2001) using the best-fit nucleotide substitution model as indicated by Mr Model Test 2.2 (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 burnin. Based on these settings, two independent runs were conducted to check for convergence.

To identify population structure, I conducted a hierarchical analysis of molecular variance (AMOVA) in Arlequin v 3.11 (Excoffier 2006). AMOVA uses a distance matrix approach to partition variance into individual components. For my data, alternative a priori hypotheses of population groupings (e.g., Eastern and Southern Africa) were tested to identify those groupings that resulted in more variation between groups than among populations within groups or among individuals within populations, as this is indicative of population structure (Holsinger & Weir 2009). AMOVA's were conducted on mtDNA haplotype data and DRB and microsatellite allele frequency data, based on F_{ST} values. Significance was assessed using 1,000 permutations. Other estimators, such as R_{ST}, that consider evolutionary distance, were not used as R_{ST} has high variance and because accurate estimates require many loci, which was not the case with my data (François & Nicolas 2002).

I further assessed population structure of microsatellite data through allele sharing analysis, Principle Coordinates Analysis (PCO) and Bayesian clustering analysis. A neighbour-joining tree based on Nei's allele-sharing distance (D_{AS}) was calculated in POPULATIONS v 1.2.30 (Langella). PCO was conducted in GenALEX using genetic distances between individual multi-locus genotypes (Smouse & Peakall 1999) and with a median value computed for each population (Novembre & Stephens 2008). Bayesian clustering analysis was conducted in STRUCTURE v 2.2 & 2.3 (Pritchard et al. 2000). STRUCTURE uses a Bayesian clustering model-based algorithm to elucidate the number of genetic clusters (K) within a sample set. The model is based on global allele

frequencies of genotypic data and the method attempts to find genetic clusters that are in Hardy-Weinberg equilibrium and not in linkage disequilibrium (Evanno et al. 2005). For each K value, the model generates an estimated log probability of the data, which is used to determine the most likely value among the range of K values tested, as well as a likelihood value for each individual being assigned to each cluster. I conducted runs in STRUCTURE 2.2 assuming no prior population information, with correlated allele frequencies and admixture, 200,000 burn-in cycles, 2,000,000 Markov chain Monte Carlo runs (MCMC) for K = 1-10, with ten replicates per K. I plotted likelihood values and variance amongst the 10 replicates at the different values of K, as well as the ΔK statistic which assesses the rate of change in the log probability of the data between successive K values (Evanno et al. 2005) using STRUCTURE HARVESTER v.05. (Earl 2009). The value of K that best fit the data was selected based on ΔK statistic and consistency amongst replicates (Pritchard et al. 2000). Since STRUCTURE uses global allele frequencies, there is little power for assignment of individuals derived from populations with small sample sizes. Therefore the monitoring areas with five or fewer samples were excluded from all STRUCTURE analyses.

STRUCTURE 2.3 has a new model (LOCPRIOR) specifically designed to increase the power to detect structure where data sets are limited by small sample sizes or low polymorphism (Hubisz 2009). Therefore, I used this model (keeping all other parameters the same) to detect any cryptic population structure. I used sampling location as the prior.

I assessed whether there was a pattern of isolation by distance at microsatellite loci using Mantel tests, which assess whether there is a correlation between genetic and geographic distance. Genetic distance was based on the estimator D_{est} (Jost 2008) as other estimators such as F_{ST} and G_{ST} have been shown to have a non-monotonic relationship with differentiation and may actually decrease with increasing differentiation (reviewed in Jost 2008). This was a problem in my study as some populations were completely differentiated at the DRB locus (shared no alleles) but F_{ST} recorded values of 0.32-0.41, whereas D_{est} gave the expected value of 1 (Appendices 4 & 5). Pairwise population differentiation values (D_{est}) were computed in the programme SMOGD (Crawford 2009), using 1000 bootstraps. Geographic distances (km) were calculated according to the straight line distance between population pairs. Mantel tests were computed in the vegan package of R (R core development team). The correlation between two matrices was computed using Pearson's product moment estimator to generate the observed mantel test statistic (*r*). Significance was tested by determining the position of *r* within a reference data set generated by 10,000 permutations (Sokal & Rohlf 1994) and adjusted according to Bonferroni correction.

I estimated mean long-term migration rates between African wild dog populations and long term effective population sizes in MIGRATE v.2.4 (Beerli & Felsenstein 1999). MIGRATE estimates migration rates and effective population size using coalescence theory. The parameters of $\theta = 4N_e\mu$ and $M = m_{xy}/\mu$ (where N_e = effective population size, assuming an average mutation rate, μ, of 10^{-2} for canid tetranucleotide microsatellites (Francisco 1996), and m = unscaled migration rate) are estimated from the data by exploring different genealogies using MCMC. Estimates of migration rates are based on assessing the rates of introduction of new variants into the population by migrants relative to mutation and calculated by multiplying θ and M. MIGRATE does not assume equal population sizes or symmetrical migration rates. However, it does assume constant population sizes, mutation and migration rates. I excluded poorly sampled populations from this analysis, and pooled temporal data for Kruger, Okavango and Serengeti. I assessed Eastern and Southern African populations separately as there was a sampling gap (~1500 km) between these two areas. I ran MIGRATE on microsatellite data within the likelihood framework using 10 short chains with 10,000 steps and 3 long chains of 100,000 steps and a burn-in of 10,000. I selected an adaptive heating scheme (temperatures 1.0, 1.2, 1.5, 3.0). θ and M were estimated from the data in the first run; subsequent to this, θ and M estimates of preceding runs were used as a prior in the following run. Runs were conducted until confidence intervals of the posterior probabilities of θ and M of neighbouring runs were overlapping. The number of migrants moving per generation between population x and y was calculated as $\theta * M$, where $\theta = 4\ N_e\mu$ and $M = m/\mu$ (Beerli & Felsenstein 1999).

I first tried to estimate contemporary migration rates using the programme BAYESASS 1.3 (Wilson & Rannala 2003), which detects genotypic disequilibrium in recent migrants within a Bayesian MCMC framework. However, there was not enough information in my data set for reliable estimates. I also used STRUCTURE 2.2 to test for the presence of migrants in populations which is implemented through the USEPOPINFO option (Pritchard et al. 2000). In this mode, STRUCTURE calculates posterior probabilities of membership of individuals to clusters that are defined *a priori*. In this way, individuals that are not members of assumed populations are indicated as migrants. *A prior* cluster memberships were based on previous STRUCTURE runs, which showed each study population to be a separate cluster, except for Hwange and Okavango, which belonged to a single cluster. Individuals with posterior probability values lower than 40% were viewed as putative immigrants to that population. I also used STRUCTURE posterior probability tests to identify the provenance of the animals that were translocated into Hwange. For every individual within the test population, STRUCTURE calculated the posterior probability of membership to a group of potential source populations. In this case Southern African samples were used as potential source populations.

4.3.6 Temporal patterns of genetic diversity

Samples were available from two time points for Kruger, Okavango and Serengeti, which enabled assessments of how genetic diversity changes over time. I tested for significant changes in microsatellite diversity (Gene diversity, H_e, H_o, A_R, F_{IS}) within populations over time using paired t-tests in MINITAB (Minitab Inc). Since I only had data for a single locus at the MHC, I examined whether the allelic repertoire within populations was consistent over time, as well as examining trends in diversity (Gene diversity, H_e, H_o, A_R, F_{IS}). Chi squared tests were performed to test for changes in the frequencies of the two DRB lineages. Fisher's exact tests were performed in GENEPOP (Raymond & Rousset 1995) to test for significant changes in the frequency of DRB and microsatellite alleles over time.

4.3.7 Tests for Selection

Synonymous and nonsynonymous genetic distances were calculated separately for putative peptide-binding region (PBR) sites and non-PBR sites using the Nei–Gojobori method with a Jukes–Cantor model of substitutions, as implemented in MEGA 4.0 (Tamura et al. 2007). Putative PBR sites were based on the human HLA-DRB1 (Bondinas et al. 2007). I tested for evidence of positive selection using a codon based Z test in MEGA. I tested for an excess of heterozygotes within populations based on both allelic and lineage data, using the programme GENEPOP (Raymond & Rousset 1995). Tajima's D and Ewen Watterson tests for selective neutrality were computed in Arlequin v 3.11 (Excoffier 2006). Significance of both tests was assessed using 1000 simulations and adjusted according to Bonferoni correction. Tajima's D is a test that compares two estimates of the parameter θ; θ_π which is based on nucleotide diversity under mutation drift equilibrium and θ_S which is based on the number of segregating sites under the infinite sites model. The D statistic is based on whether θ_π and θ_S are similar or different. Where $D=0$, neutral evolution is indicated, whereas purifying selection is indicated where $D<0$, and balancing selection is indicated where $D>0$ (Hartl & Clark 2007). However, demographic factors, such as expansions and contractions effect D in a similar way to selection, and therefore it is not possible to conclusively disentangle selective from demographic events. The Ewen-Watterson test assesses allele frequencies as evidence of selection. Under neutral expectations, it is expected that there will be one common allele, whilst all other alleles occur at low frequency. The Ewen-Watterson test compares the expected heterzygosity based on mutation-drift equilibrium (Ewen's formula) and compares that with the expected heterzygosity based on the allele frequencies of a population (Garrigan & Hedrick 2003).

Selection can be inferred from patterns of genetic differentiation. Specifically, selection is indicated where patterns of genetic differentiation at DRB and neutral markers are incongruent. Therefore an analysis of covariance (ANCOVA) was computed in the programme R (R core development team) to test whether there was a significant difference in the slopes describing the relationship between D_{est} (genetic differentiation) and geographic distance, when based on microsatellite and DRB data. For the DRB data, I computed values firstly as allelic data, and secondly as lineage data whereby an individual's alleles were coded according to the lineage from which they were derived (i.e. AA, AB, BB). Balancing selection at the MHC is expected to result in more even allele frequencies between populations, and thus lower genetic differentiation, than at neutral markers (Schmidt et al. 2008). Therefore, I used partial Mantel tests to test whether genetic differentiation at the DRB was larger than expected under neutrality.

4.4 Results

4.4.1 Summary statistics

mtDNA

Amongst 425 samples sequenced at the mtDNA D-loop, I identified 10 haplotypes with 23 polymorphic sites and between 1 and 18 nucleotide differences. This included eight previously identified haplotypes and two new haplotypes: S4 found in three immigrant males in Selous (S.Creel field observation), and S5 in five immigrant (3 males and 2 females; J.W. McNutt, field observation) in Okavango (Table 4.2). In each population I found between one and six haplotypes (Table 4.3) and nucleotide diversity ranged from π = 0 – 0.02. Monitoring populations in the central area of the sampling range had the highest diversity (Okavango, Hwange and Selous; $\pi > 10^{-3}$, ≥3 haplotypes, whereas Laikipia and Lowveld had the lowest diversity ≤2 haplotypes, π <0.001 (Table 4.3). Re-sequencing of samples to verify the eight haplotypes (AF335724-AF335731) previously identified by (Girman et al. 2001) showed some sequencing errors in three haplotypes primarily as a result of an indel that was not detected by manual sequencing but was apparent from fluorescent chromatographs. The modified haplotypes are provided in Appendix 6.

Microsatellites

I typed 321 samples at ten microsatellite loci and found between five and 30 alleles per locus (mean = 13). All populations were in Hardy Weinberg equilibrium at all loci except for a single locus in Kruger Old (FH2611) and Okavango Old (FH2658). Okavango, Hwange and Selous showed the highest levels of microsatellite diversity (Std-A_R=4.54-5.68, H_e=0.68-0.76; Table 4.3). Lowveld exhibited the lowest diversity (Std-A_R=4.09, H_e=0.59), and Laikipia had the lowest values of H_o (0.61, 0.62), which is indicative of relatively higher rates of inbreeding. Laikipia was the only monitoring population where F_{IS} (microsatellite) was significantly different from 0, which further suggests inbreeding in this population and also indicates that Laikipia is not in Hardy Weinberg equilibrium despite the results from tests based on individual loci.

DRB

Sequence-based typing of 341 samples across all 13 monitoring areas (i.e. including monitoring populations with small sample sizes, and Hwange translocates) at the DRB locus identified 21 alleles derived from two allelic lineages, A and B (Figure 4.2;

Table 4.4). This included 13 alleles described in Marsden et al (2009) and six new alleles (sequences are provided in Appendix 7). As for my previous analyses: all alleles were unique to African wild dogs; there was no evidence of pseudogenes (stop codons or frameshift mutations) to indicate amplification of non-functional alleles; and I am confident that a single locus was amplified due to lack of more than two sequences per individual, in conjunction with pedigree data from captive animals, which clearly showed biparental inheritance of alleles and cosegregation of allelic lineages (Chapter 2; Marsden et al. 2009). With the exception of two highly similar alleles from lineage B (B6 and B12), all DRB alleles varied at the amino acid level. Amongst alleles, 31/270 sites, 16/89 codons and 9/22 peptide binding region sites were variable. Within monitoring populations there was an average of 5.93 alleles (range 2 – 11), and an average heterozygosity of 75.3% (range 53.9-92.8%), with the highest diversity found in Selous, Okavango and Hwange (A_R=6.6-9.4; H_e=0.79-0.88) and lowest diversity found in Laikipia and Lowveld A_R=3, H_e=0.57-0.68; Table 4.3).

4.4.2 Effective population size and Demographic history

Contemporary estimates of effective population size (N_e) were consistently small in all monitoring populations (mean <30), regardless of the analysis method used (Table 4.5). This gave N_e/N estimates ranging from 0.02-0.21 (Table 4.5). Long term N_e estimates derived from MIGRATE were higher than those from other methods (Table 4.5). This may be the result of the presence of unsampled (ghost) populations which are known to elevate effectively population size estimates from MIGRATE because (unlike the other methods) MIGRATE assumes all populations were sampled (Beerli 2004). Alternatively, they may reflect that MIGRATE will overestimate N_e if population sizes are decreasing (Beerli 2009). There was insufficient power in my data to calculate N_e using the heterozygote excess method (results were infinity); therefore, these results are not shown.

I found significant evidence of bottlenecks in all populations. However, there was some variation amongst the different bottleneck tests. None of the five monitoring populations with 30 or more samples showed evidence of a mode shift in allele frequencies (allele frequency test), which would be indicative of a recent bottleneck (Table 4.6). Wheresas there was evidence of a heterozygosity excess in Kruger, Serengeti and Laikipia, although, only with some mutational models (Table 4.6). Furthermore, all monitoring populations showed significantly lower M ratios than those expected under mutation-drift equilibrium, for at least some values of θ, which is indicative of a bottleneck. The discrepancy between these results can be explained by differences in power between amongst the bottleneck tests under different demographic histories (Williamson-Natesan 2005). For example, the M ratio test is able to detect more distant bottlenecks than the other tests. The mtDNA data showed no significant values of Tajima's *D* or Fu's *F*,

indicating no evidence of a departure from neutrality, or unstable population dynamics. In all monitoring populations mtDNA mismatch distributions were significantly different from a model of population expansion (results not shown), indicating that populations have not increased in size.

Based on Bayesian coalescence simulations implemented in MSVAR, I found evidence of a large population decline from an ancestral N_e of approximately 600-900, to a current N_e of <10, within the last 100 years across all three monitoring populations assessed (Table 4.7). These estimates match expectations for this species based on their social breeding system (Creel & Creel 1998), and known history of extensive habitat loss and persecution within the last 30 years (Woodroffe & Ginsberg 1997). It is noteworthy that N_0, N_1 and T estimates were consistent across different priors, and the signal of decline was still apparent when a prior of population expansion was used (data not shown).

4.4.3 Spatial patterns of genetic diversity

mtDNA

As identified by Girman et al (2001), geneological analysis revealed that the mtDNA haplotypes were derived from two highly divergent clades (Figure 4.3): an Eastern clade (3 haplotypes) and a Southern clade (7 haplotypes). There was strong structuring of the two mtDNA clades. The three most easterly (Laikipia, Serengeti, Masai-Steppe) and two most southerly (Kruger, Lowveld) monitoring populations, had exclusively Eastern and Southern mtDNA clade haplotypes, respectively. However, Selous, Okavango and Hwange, which are located between these extremes, had haplotypes from both clades. Selous, in Southern Tanzania, shared no haplotypes with other East African monitoring populations, but did share haplotypes with Niassa (n=1) and Sofala (n=3) in Mozambique, as well as Okavango and Hwange. Unique haplotypes were found in Kruger (S1), Selous (S4), and Okavango (S5). The most abundant haplotypes were S2 (29%) which was found in all monitoring populations in Southern Africa, and E1 (26%), which was found in the most Eastern African monitoring populations, as well as Okavango and Hwange in Southern Africa. A hierarchical AMOVA of mtDNA haplotypes indicated significant population structure, with 46% of variation being apportioned between four regions: 1) Southern cluster (Kruger, Lowveld); 2) South Western cluster (Okavango, Hwange); 3) Central cluster (Selous); and 4) Eastern cluster (Masai-Steppe, Serengeti-Mara, Laikipia) (Figure 4.3b; Table 4.8), which correspond with geographic expectations.

Microsatellites

All monitoring populations had at least one private allele that was unique to that population (average 3.2, range 1-8) (Table 4.3). The majority of microsatellite variation (84%) was apportioned within populations (Table 4.8); therefore, although the hierarchical AMOVA indicated the presence of population subdivision, it was not informative in identifying groupings amongst samples. Nevertheless, clustering analysis of microsatellite data based on allele sharing distances (represented by a neighbour joining network) was indicative of considerable population subdivision. Individuals largely clustered as monitoring population specific groupings, although Okavango and Hwange clustered together (Figure 4.4). These results were largely concordant with the results from Bayesian clustering analysis in STRUCTURE. From the STRUCTURE analyses, the ΔK statistic identified two peaks, the height of which signifies the strength of the signal of population subdivision (Evanno et al. 2005). K=4 represented the deepest level of population subdivision (Evanno et al. 2005), whereas K=7 which represented finer scale population structuring (Figure 4.5). Only at K=4 & 7 were results consistent between replicates. The four clusters at K=4 were: 1 = Kruger, 2 = Okavango-Hwange, (Lowveld was admixed 1&2), 3 = Selous, Masai Steppe, Serengeti-Mara, 4 = Laikipia. At K=7, the clusters correspond to every monitoring population being distinct except for Okavango and Hwange, which appear to be a single genetic population (Figure 4.5). Subsequent analyses of the same data set using the LOCPRIOR model STRUCTURE 2.3 returned identical results (data not shown). Finally, I took a hierarchical approach and ran STRUCTURE 2.2 on subsets of the data independently; specifically each of the four groupings at K=4 and Eastern and Southern Africa. These additional analyses did not detect any further populations than shown at K=7 and thus the results were also congruent with K=7 for the overall dataset (results not shown).

Principle coordinates analysis also indicated strong structuring, but not complete isolation, of wild dog monitoring populations, as well as a pattern of isolation by distance (Figure 4.6). PC1 and PC2 accounted for a considerable amount variation, 28% and 18% respectively, and appear to correspond to strong structuring along north-south and east-west axes, resulting in a striking concordance between population medians and geographical positioning of sampling locations (Figure 4.6). I found significant evidence of isolation by distance at microsatellite loci based on mantel tests (r=0.4168, p=0.00023) and regression of microsatellite genetic distance (D_{est}) against geographic distance (R^2=0.17, p<0.05; Figure 4.8b).

Migration rates

Posterior probability tests indicated that the majority of individuals originated from the population from which they were sampled; 97% (312/321) of individuals were assigned to their monitoring population with more than 95% probability (Hwange and Okavango were pooled for this analysis – see methods). Overall, only one individual was indicated to be a migrant (posterior probability <40%): T04-187 from the Okavango was clearly identified as a migrant in both STRUCTURE (assignment probability 14%, highly significant mis-assignment) and PCO analyses (where it was an outlier: Figure 4.6).

Sixteen animals were recently artificially translocated into Hwange (North West Zimbabwe); six from Pilansberg, South Africa, and ten from unknown locations in Zimbabwe. Clustering analyses clearly detected all 16 animals to be of non-Hwange origin (data not shown). Posterior probability analyses assigned 8/10 of the unknown Zimbabwe animals to the Lowveld (South East Zimbabwe) with >75% confidence. PCO and STRUCTURE analyses also showed these individuals to cluster with the Lowveld. None of the animals derived from Pilansberg could be assigned to any monitoring population with >50% confidence. It is not clear whether this is because of mixed ancestry (Pilansberg is an artificially formed population of wild dogs) or because they are derived from an unsampled population.

Long-term migration rates were estimated in MIGRATE. However these results should be viewed with extreme caution and not interpreted literally because of the assumptions that underlie this method e.g. constant population sizes and migration rates (Abdo et al. 2004; Whitlock & McCaultey 1999). MIGRATE predicted high levels of geneflow between Okavango and Hwange (~5-7 migrants/generation), which is not unexpected given that these populations are located within the same habitat fragment. However, all other monitoring populations were spaced widely apart. Between these distant monitoring populations MIGRATE indicated low migration rates. Since I did not have samples from intermediate populations, these estimates are not very valid as it is highly unlikely that wild dogs have migrated directly from Okavango to Kruger (1.8/gen), or Selous to Laikipia (1.7/gen) (Figure 4.7). Rather, migration is likely occurring from an unsampled area between these two sampling points.

DRB

At the DRB, alleles from both DRB lineages were present in all monitoring populations (Table 4.4). Whether lineage A or B was most common varied between monitoring populations but not in any consistent geographic pattern. At the allelic level, AMOVA analyses showed that the majority of DRB variation was apportioned within

monitoring populations (79-80%), which corresponds with the pattern in microsatellites. Seven of the 19 DRB alleles found in monitoring populations were private alleles (Kruger, A7,B11; Okavango, A2, B9; Serengeti, A6; Selous, A8, B10; Table 4.4). Of the remaining 12 alleles, only 4 were found exclusively in Southern or Eastern African monitoring populations (assuming that Niassa and Sofala, are classed as eastern, as indicated by mtDNA and microsatellite data). Fisher's exact tests showed DRB allele frequencies to differ significantly between all monitoring populations (data not shown).

4.4.4 Temporal patterns of genetic diversity

Temporal samples were collected in Kruger, Okavango and Serengeti (Table 4.3). In Okavango, there was no significant change in genetic diversity (msats, Paired T test, H_e - T=-2.14, p=0.058; DRB, *Std-*A_R= 8.3 & 8.0; H_e - 0.85 & 0.86). However, in Kruger, there was a large and significant reduction in observed heterozygosity at microsatellite (15.4%; paired T test, T=2.24, p<0.05) and DRB loci (13.1%), and an increase in F_{IS} at both markers, which is indicative of inbreeding. Over time, within Kruger and Okavango, the old and recent sample sets had the same DRB alleles, and there was no significant change in the frequencies of DRB alleles (Fisher's exact test, Okavango; p=0.252, Kruger p=0.25163) or the frequencies of the two DRB lineages (Chi Sq, 1.d.f, p>0.05). In Serengeti, there was no significant change in microsatellite diversity (Paired T test, H_e T=0.72, p=0.486; Gene diversity T=1.60, p=0.145). At the DRB, H_o and *Std-*A_R increased, from 67% to 93% and 4.8 to 6 respectively. Furthermore, the alleles present changed considerably between these two time periods (Table 4.4), resulting in a significant change in DRB allele frequencies (Fisher's exact test p<0.0001) and in the frequencies of the two DRB lineages (Chi Squared, d.f.=1, p<0.005, Chisq value = 8.288).

4.4.5 Tests for selection

At the DRB, the ratio of nonsynonymous to synonymous substitutions at the putative peptide binding region (PBR) sites was greater than 1.0 and larger than in non-PBR. However, the ratio was not found to be significantly different between the PBR and non-PBR (PBR: d_N = 0.303, d_S = 0.183, d_N/d_S = 1.656, p = 0.093; non-PBR: d_N = 0.013, d_S = 0.011, d_N/d_S = 1.182, p = 0.460). There was no evidence of an excess of heterozygotes or lineage heterozygotes (individuals with an allele from both lineage A and B) relative to Hardy Weinberg equilibrium in any monitoring population. Tajima's D test for selective neutrality was not significant in any monitoring population. The Ewen's Watterson estimator detected a deviation in allele frequencies from neutrality in Okavango (p=0.001) but not other monitoring populations.

Overall, genetic differentiation (D_{est}) was high amongst monitoring populations at both microsatellite and DRB markers, but was greater at the DRB (DRB D_{est}, = 0.722, 95% CI = 0.691-0.751; Msats, D_{est} = 0.419). Nonetheless, genetic differentiation of DRB and microsatellite loci were significantly correlated (*r*=0.3784, *p*=<0.01; Figure 4.8a). Population pairwise differentiation estimates (D_{est}) were >0.2 between all monitoring populations except Okavango and Hwange and within monitoring populations over time (Appendix 4). The F_{ST} values were substantially lower than D_{est} values, but were high (>0.15, Wright's guidelines, (Frankham et al. 2002), by F_{ST} standards (mean F_{ST} msats = 0.15, mean F_{ST} DRB = 0.19: Appendix 5).

Both microsatellite and DRB showed evidence of significant isolation by distance based both on Mantel tests (DRB: *r*=0.506, *p*<0.01; msats: *r*=0.417, *p*<0.001) and regression of genetic distance against geographic distance (DRB: R^2=0.26, *p*<0.001; Msats: R^2=0.17, *p*<0.05). However, the increase in genetic divergence with geographic distance was significantly stronger at the DRB than microsatellites (ANCOVA, p<0.05; Figure 4.8b). Partial mantel tests showed that the relationship between DRB and genetic distance remained significant even after controlling for microsatellite differentiation (Partial Mantel Test; *r*=0.416, *p*=0.01), suggesting that genetic differentiation at the DRB is larger than expected under neutrality. By contrast, the correlation between microsatellites and geographic distance was not significant when DRB was controlled for (Partial Mantel Test, Bonferroni corrected *p* value=0.01; *r*=0.284). When genetic differentiation at the DRB was calculated using allelic lineages rather than alleles, no isolation by distance relationship was found (Regression; DRB lineages, R^2<0.0001, *p*=0.881) and there was no correlation with microsatellites (Regression; R^2 = 0.01, *p*=0.462; Figure 4.8c).

4.5 Discussion

4.5.1 Effective population size and Demographic history

Demographic history is a key factor influencing the patterns of genetic diversity within a species, as genetic drift and inbreeding are increased during periods of small effective population size, resulting in the loss of genetic diversity. Even though I assessed some of the largest extant wild dog populations, my estimates of current effective population sizes were consistently small across monitoring populations (contemporary estimates of N_e, <30; Table 4.5) and with the exception of Selous, N_e/N estimates were ≤0.1. Habitat loss, population declines, inbreeding and reduced geneflow are likely the primary cause of the small N_e of wild dog populations. However, the cooperative breeding system of wild dogs, whereby only a single pair of adults typically breed per pack (Girman et al. 1997), will also have contributed to a small N_e. Indeed, these estimates are similar to other endangered cooperatively breeding canid populations: Ethiopian wolves (*Canis simensis*, Bale Mountains N_e 10-25, (Randall et al. 2007); Grey wolf (Canis lupus, Finland, N_e 37-43 (Aspi et al. 2006).

A critical assessment in conservation biology is to determine whether population sizes have always been small or whether declines have occurred (Crandall 2009), as this has direct implications for conservation management. Therefore, I tested whether there was a genetic signature of decline in African wild dogs. All monitoring populations had significantly lower M ratios than expected under drift-mutation equilibrium, which is indicative of a bottleneck (M = 0.64-0.75, P<0.05). However, the allele frequency method detected no bottlenecks, and there was only limited evidence of bottlenecks in some monitoring populations using the heterozygosity method (Table 4.6). The discrepancy between these tests likely reflects differences in power of detection of bottlenecks under different demographic histories (Williamson-Natesan 2005). In particular, M ratio tests are able to detect bottlenecks over longer time scales than the heterozygosity excess method (Williamson-Natesan 2005), and can detect bottlenecks even when populations recover quickly, which African wild dogs have a propensity to do as a result of their large litter sizes (Woodroffe 2010).

I used Bayesian coalescent simulations of microsatellite data to investigate the scale and timing of demographic declines in African wild dogs. I found evidence for a large demographic reduction of approximately 2 orders of magnitude within the last 100 years (Table 4.7). The timing of this population decrease is congruent with the extensive and widespread declines in wild dogs associated with human related habitat loss and persecution that are known to have occurred throughout the 20th century (Woodroffe &

Ginsberg 1997). Interestingly, although this result matches the finding in another endangered species, the orang-utan (*Pongo pygmaeus*; (Goossens et al. 2006), it differs from other African mammals (African elephant, *Loxodonta Africana*, (Okello et al. 2008), African buffalo, *Syncerus caffer*, (Heller et al. 2008), Walia Ibex, *Capra walie*, (Gebremedhin et al. 2009) where declines have been suggested to correspond with mid-Holocene climatic change several thousands of years ago. The strong evidence of a human mediated demographic collapse in African wild dogs highlights the urgent need for conservation efforts to reverse this trend.

Although microsatellite data provided evidence of declines over the last 100 years, mtDNA and DRB data, which are informative at more distant time scales, also suggest further declines prior to this. Mitochondrial DNA haplotypes showed evidence of strong Eastern-Southern regional clustering. Similar structuring is common in other African mammals (reviewed in Hewitt 2004) such as Hartebeest (*Alcelaphus buselaphus*), Wildebeest (*Connochaetes taurinus*) and Topi (*Damaliscus korrigum*) (Arctander et al. 1999), and is thought to represent contraction to refugia during the climatic shifts that were experienced during climatic transition during the late Pleistocene-Holocene. The mtDNA pattern in wild dogs may indicate that this once widespread species declined to a few refugia where ancestral haplotypes were present. Later expansion to recolonise the former range, combined with subsequent habitat fragmentation and time, would then have resulted in the evolution of new mtDNA haplotypes and loss of haplotypes by drift. The high diversity of mtDNA haplotypes in Hwange and Okavango indicate this to be a potential refugium; however, high diversity is also expected to be retained more effectively in these larger populations. It is also possible that the rift valley, which separates Eastern and Southern Africa, has contributed to divergence of mtDNA between these regions (Figure 4.3). At the MHC, previous research showed that African wild dogs lack diversity at the DRB relative to other wolf-like canids (Marsden et al. 2009): wild dogs not only lack diversity at the DQA1 (1 allele) and DQB1 (2 alleles) loci that are strongly linked to the DRB locus assessed here, but they have just two allelic lineages at the DRB. This lack of diversity could be indicative of a dramatic bottleneck in the past. Given that both DRB allelic lineages were found in all populations across Eastern and Southern Africa, the DRB data are most congruent with dramatic range wide declines to refugia and subsequent recolonisation (Marsden et al. 2009), as indicated by the mtDNA haplotype distribution.

Overall, my results suggest that the consistently small N_e estimates I found across monitoring populations are not primarily a reflection of species ecology. Rather, they are a consequence of dramatic anthropogenic driven demographic declines (approximately two orders of magnitude) within the 20th century. The impact of these declines on population connectivity and genetic diversity is of considerable concern.

4.5.2 Spatial patterns of genetic diversity

Microsatellites were indicative of strong population structuring of wild dog populations. With the exception of Hwange and Okavango, the microsatellite data suggested each monitoring population to be a separate genetic cluster (Figure 4.5). Furthermore, estimates of migration rates between monitoring populations were low everywhere except for between these two monitoring populations (Figure 4.7). Hwange and Okavango are located in the largest continuous area supporting wild dogs (Figure 4.1) and this connectivity appears to permit gene flow to occur across quite distant locations (~400 km). Additional samples are required to determine whether NE Namibia wild dogs are contiguous with Okavango and Hwange. Elsewhere, the lack of gene flow between populations suggests that there has been genetic isolation associated with fragmentation of wild dog populations, resulting in a pattern of isolation with distance (Mantel tests; Figure 4.8b), which is particularly evident in the PCO (Figure 4.6). It is possible that the gaps in my sampling regime have resulted in an under-estimation of connectivity between wild dog populations. However, many of my sampling gaps reflect geographic areas that do not currently support wild dogs. Furthermore, STRUCTURE detected only a very small number of migrants from unsampled populations. This indicates that migration from ghost populations occurs at a low rate, which matches expectations given the high mortality of dispersers (Creel et al. 2004), in particular across unprotected areas (Woodroffe et al. 2007a).

There was good correspondence in the population groupings suggested from mtDNA (Table 4.8) and the uppermost level of structuring suggested from microsatellite data (K=4), except that the microsatellites identified Laikipia as distinct to other East African populations whereas mtDNA identified Selous as being distinct to other East African populations (Figure 4.3). Given that the E1 and E2 mtDNA haplotypes are so common in East Africa, these data likely reflect that Laikipia was recolonised by a distinct and unsampled East African population. However, it is unclear why Selous gives an isolated signal based on mtDNA whereas microsatellites suggest connectivity, although cytonuclear genomic dissociation (different evolutionary histories for nuclear and mtDNA genomes) (Roca et al. 2005) or sex-biased dispersal are possible explanations.

Where populations are small and fragmented, genetic diversity is likely to be lost as a result of genetic drift. I consistently found higher diversity in Okavango, Hwange and Selous, and lower diversity in Lowveld and Laikipia, across both neutral and DRB markers (Table 4.3). Higher diversity might be expected in Okavango, Hwange and Selous because these monitoring populations form part of some of the biggest most stable wild dog populations (n~2,500 and n~1,300 respectively, (UNEP 2008) and therefore they should be less affected by genetic drift. Laikipia and Lowveld were recently recolonised

following extirpation (Pole 2000; Woodroffe 2010), and likely suffered more severe genetic drift and loss of genetic diversity and increased homozygosity associated with founder effects, small population sizes and inbreeding. The low MHC variation I found in Lowveld and Laikipia contrasts with the expectation that selection can retain adaptive genetic variation despite strong genetic drift, as found in African buffalos (Wenink 1998), Island fox (*Urocyon littoralis* (Aguilar et al. 2004) and Guppies (*Poecilla reticulate,* (van Oosterhout et al. 2006). However, similar to African wild dogs, population bottlenecks have resulted in depleted MHC variation in a number of other species (Tasmanian devil (*Sarcophilus harrisii* (Siddle et al. 2007), Moose (*Alces alces* (Mikko & Andersson 1995), Bison (*Bison bonasus* (Radwan et al. 2007), (reviewed in Radwan et al. 2010), indicating that the impact of a bottlenecks on MHC diversity may be context specific.

Rarely do studies of endangered species have sampling at the spatial scale achieved in this study. However, with this, I was able to show that widespread habitat fragmentation has resulted in extensive genetic fragmentation of wild dog populations. Furthermore, I have been able to characterise levels of genetic diversity and inbreeding across multiple populations and thus highlight populations particularly at risk of genetic threats (Lowveld and Laikipia), as well as to identify risk factors of genetic decline (founder effects).

4.5.3 Temporal patterns of genetic diversity

In the absence of selective pressures and high levels of geneflow, genetic diversity is expected to be relatively constant over time in large stable demographic populations. Indeed, genetic diversity estimates were stable in Okavango (Table 4.3), which is part of the largest wild dog population (Woodroffe et al. 1997; Woodroffe et al. 2004b). By contrast, genetic diversity is expected to fluctuate as a result of genetic drift and/or inbreeding in populations that are small or declining. My data highlighted a considerable loss of H_o at both microsatellite and DRB loci in Kruger between 1991-5 and 2007 (Table 4.3). This was coincident with a large demographic decline of >70% from a peak of 434 animals in 36 packs in 1995 to a low of 120 animals in 17 packs in 2005 (EWT 2009) within Kruger National Park (~19,000 km2), and likely reflects increased inbreeding and genetic drift associated with reduced population size. Temporal changes in genetic diversity may also vary as a result of selection pressures. In the Serengeti-Mara, the disappearance of wild dogs from the main monitoring area in 1991 and subsequent local re-establishment in 2001 was not associated with significant changes in microsatellite diversity based on samples collected both pre-1991 and post-2001 (Table 4.3). However, at the DRB, there were notable changes (Table 4.3; Table 4.4). Observed heterozygosity increased by 26%, indicating a reduction in inbreeding or selection for heterozygotes. The DRB allelic composition also changed considerably, resulting in a significant change in the

proportion of lineage A and B alleles; the proportion of lineage A alleles increased from 39% to 75% (Table 4.4). By contrast, allelic composition and lineage ratios did not change in Kruger or Okavango over the same time period (Table 4.4). Due to the functional relevance of the DRB locus, the temporal changes in DRB alleles in the Serengeti-Mara could represent adaptive differences between the two sampling periods in response to changes in pathogen selective pressures. However, given my limited sampling, I cannot rule out the potential influences of stochastic processes.

Few studies have assessed temporal changes in allele frequencies and genetic diversity at both MHC and neutral loci. Those that have assessed multiple populations (Miller et al. 2001; Oliver et al. 2009a) have found like my study, that MHC and microsatellite allele frequencies are temporally stable in some populations, but not others. This indicates that selection is both spatially and temporally variable, which supports the hypothesis that fluctuating selection pressures are involved in balancing selection (Hill 1991). Similar to my findings in the Serengeti, temporal studies of individual populations of Great reed warblers (*Acrocephalus arundinaceus*; (Westerdahl et al. 2004) and Soay sheep (*Ovis aries* (Charbonnel & Pemberton 2005) found higher temporal divergence at MHC loci than at microsatellite loci, indicating temporal variation in local parasite selective pressures. In contrast to the Kruger population, Soay sheep on the St Kilda island were found to show no significant changes in microsatellite or MHC diversity over 12 years (Charbonnel & Pemberton 2005), despite population fluctuations from ~2000 to ~600 animals. However, Kruger may have been more vulnerable to genetic drift because it was reduced from and to smaller population sizes. A dramatic bottleneck from 100 to 11 birds in a small population of song sparrows (*Melospiza melodia*), was shown to result in a significant decline in microsatellite diversity (MHC diversity was not measured) (Keller et al. 2001), which corresponds more with my finding in Kruger. In this case, demographic recovery occurred within 3 years, and low levels of immigration resulted in recovery of genetic diversity to pre-bottleneck levels within 2 generations (2-3 years) (Keller et al. 2001). Similar genetic rescue of MHC and microsatellite diversity by migration has been reported elsewhere (Seddon & Ellegren 2004; Vilà et al. 2003; Wenink 1998). It is encouraging, therefore, that wild dogs are such a highly mobile species (Fuller et al. 1992). Nonetheless, genetic recovery in Kruger is likely to be delayed by both the absence of demographic recovery to date (which will accentuate the impact of the bottleneck), and the lack of connectivity to nearby populations (Figure 4.1).

Overall, my temporal data indicate that demographic declines, stochastic processes and selection contribute to changes in genetic diversity within populations over time. Consequently, I reiterate the contention of Oliver et al (2009a) that “snapshot” sampling could be misleading as to the genetic status of a population. This is particularly important in endangered species that suffer continuous demographic threats. The Kruger

population assessed here, for example, was regarded as 'big and secure' (Mills et al. 1998) and perceived as a 'stronghold' of African wild dogs (Woodroffe et al. 1997). However, the recent demographic declines and associated genetic impacts, highlight that even the largest wild dog populations are at risk. This indicates the need for continual monitoring and reiterates the enormous value that continuous long-term field projects can contribute to endangered species conservation (Thirgood 2007).

4.5.4 Tests for selection DRB

Detecting selection at the MHC can be difficult because selection events are transient and the signal of selection is not expected to be consistent through time (Garrigan & Hedrick 2003). Furthermore, in small populations, genetic drift may replace selection as the dominant force governing adaptive variation. Balancing selection is a key mechanism that maintains variation at the MHC (Garrigan & Hedrick 2003). It is indicated by: 1) an increased ratio of nonsynonymous (dN) to synonymous (dS) substitutions at the amino acid residues of the functionally important PBR (reviewed in Garrigan & Hedrick 2003); 2) an increased proportion of intermediate frequency mutations (Tajima's D); 3) excess heterozygosity relative to Hardy-Weinberg equilibrium; and 4) a more even allele distribution than expected under mutation-drift equilibrium (Ewen's Watterson tests). Based on these criteria, I did not find significant evidence of balancing selection at the DRB locus. There was not a significant excess of nonsynonymous substitutions, despite elevated dN:dS ratios at putative PBR sites of DRB alleles and 20/21 wild dog DRB alleles differing at the amino acid level (Radwan et al. 2010). Neither was there evidence of a deviance from selective neutrality in Tajima's D or an excess of heterozygotes relative to Hardy-Weinberg expectations in any population (despite observed heterozygosity between 0.54-0.93). I did, however, find a signal of balancing selection in Okavango as indicated by the Ewen's Watterson test, p=0.001). This may suggest selection operates in the largest wild dog population n~2500 (UNEP 2008), but this significant result may also be the product of a bottleneck, which results in rare alleles being lost more rapidly than heterozygosity (Garrigan & Hedrick 2003; Hartl & Clark 2007). The lack of a signal of selection based on these tests is not atypical for MHC studies and has been attributed to the low power of these tests as well as the transient nature of selective events (reviewed in Garrigan & Hedrick 2003; Zhai et al. 2009).

An alternative method used to indicate balancing selection is based on the expectation that selection should result in higher levels of diversity at adaptive loci and differences in the patterns of genetic diversity at adaptive and neutral genetic diversity (Figure 4.9) (Campos et al. 2006; Garrigan & Hedrick 2003). In this study, the populations with the highest microsatellite diversity had the highest DRB diversity, and vice versa. I also found a significant correlation in pairwise differentiation between populations at DRB

and microsatellite loci (Figure 4.8a; R-sq = 0.141, p = 0.004815). These correlations have been found in other studies (e.g. Campos et al. 2006), , and may suggest the role of neutral factors in shaping patterns of DRB variation, which is likely the result of the reduced power of selection in small populations; alleles are predicted to become effectively neutral when $s < 1/ (2N_e)$ (s is the selection coefficient and N_e is the effective population size), resulting in correlated patterns of neutral and adaptive diversity. However, although DRB and microsatellite loci were correlated, genetic differentiation and isolation by distance were more pronounced with DRB allelic data than with microsatellite data (Figure 4.8b). Furthermore, partial Mantel tests showed that patterns of genetic differentiation at the DRB remained significant even after controlling for neutral microsatellite variation (r=0.4156, p=0.00749), suggesting that genetic differentiation at the DRB is larger than expected under neutrality. Together, these results suggest that neutral forces alone cannot explain patterns of DRB variation. The higher divergence at DRB is also consistent with diversifying, rather than balancing, selection acting on DRB alleles (Figure 4.9), and may indicate local adaptation (Bryja 2007; Ekblom et al. 2007; Landry & Bernatchez 2001). However, it is also possible that this pattern has been generated by balancing selection through either fluctuating selection or negative frequency dependent selection for new rare alleles (Spurgin & Richardson 2010).

Results from other studies comparing patterns of divergence at MHC and microsatellites have been variable (see review in Bernatchez & Landry 2003). For example, higher divergence at the MHC relative to microsatellites, which is indicative of directional selection, was detected in Atlantic salmon (*Salmo salar*; (Landry & Bernatchez 2001), Great snipe (*Gallinago media*; (Ekblom et al. 2007) and Water voles (*Arvicola terrestris*; (Bryja 2007), whereas no difference was reported in brown trout (*Salmo trutta*; (Campos et al. 2006), Gila topminnow (*Poeciliopsis occidentalis*; (Hedrick et al. 2001b) and Bighorn sheep (*Ovis canadensis*; (Boyce et al. 1997). A reoccurring finding of most studies comparing neutral and MHC variation however, is that neutral forces and demographic processes are the predominant factors influencing patterns of MHC variation (Radwan et al. 2010), and that these may mask any current or historical signals of selection (Oliver et al. 2009a). A recent study of water voles by Oliver (2009a) showed that across six years, genetic differentiation at MHC loci varied between being higher than, lower than, and equal to, genetic differentiation at microsatellite loci. This highlights the influence of neutral forces and reiterates that selection at the MHC is both spatially and temporally variable (Oliver et al. 2009a).

DRB alleles in African wild dogs are derived from just two lineages, with alleles within lineages differing by few amino acids but large numbers of amino acid substitutions between lineages. If alleles within lineages are functionally similar, balancing selection may be expected to act on lineages rather than alleles. When I analysed the DRB data in

terms of lineages I found a pattern indicative of balancing selection; there was no evidence of isolation by distance or genetic differentiation (Figure 4.8c; DRB R-sq = <0.0001, p=0.881) and there was no correlation between DRB lineage and microsatellite data (Regression, R-sq = 0.01, p = 0.4618). Indeed, the presence of both lineages in all populations across Eastern and Southern Africa, is in itself suggestive of balancing selection (Table 4.4). It is highly unlikely that independent selection events consistently retained just these two allelic lineages across all populations. Rather, I suggest that wild dogs must have lost most DRB variation during a bottleneck, which restricted them to a few refugia. Subsequent expansion (as indicated by mtDNA data) and later selection resulted in both lineages being spread and retained in all populations. I found that in some populations lineage A was the most common, whereas in others lineage B was most common (Table 4.4). This may reflect adaptation to local pathogenic communities. Similar to the pattern in African wild dogs, a study of White tailed deer (*Odocoileus virginianus*) revealed just two DRB allelic lineages, where alleles within lineages differed by few amino acids (van Den Bussche et al. 1999). It was subsequently shown that alleles from each lineage were strongly associated with resistance to either ectoparasite ticks (lineage 1) or nematodes (lineage 2) (Ditchkoff et al. 2005) and therefore population DRB lineage composition may also reflect adaptation to local pathogenic communities.

In summary, my data suggest that balancing selection is operating on DRB lineages, and directional selection on DRB alleles within lineages. However, like many other MHC studies the overall footprint of selection in my data was low, which may reflect an erosion of historical signal due to strong genetic drift in small populations (Campos et al. 2006; Miller & Lambert 2004) and/or a lack of power of tests where sample sizes are small.

4.5.5 Conservation management implications

Adaptive genetic variation largely determines the ability and speed at which a species can adapt/evolve to changes in its environment. It is also the basis of local adaptation and evolutionary differentiation between populations (Gebremedhin 2009; Hoglund 2009). Consequently, knowledge of adaptive variation, such as the MHC is thought to be critical for endangered species conservation. At the species level, a major bottleneck has been suggested as the cause of the lack of MHC variation in African wild dogs relative to other wolf-like canids (Marsden et al. 2009). In addition, I found evidence of the loss of DRB diversity as a result of genetic drift within populations. Whether this lack of DRB variation will reduce the viability of wild dogs or wild dog populations is not clear, as examples from other wild animals present equivocal results (reviewed in Radwan et al. 2010). However, the catastrophic population crash in Tasmanian devils, which has been linked to an inability to mount an immune response against an emerging disease as

a direct result of a lack of MHC variation (Siddle et al. 2007), suggests that the lack of MHC variation in wild dogs should be a cause for concern.

My study suggests that extensive population declines, population subdivision and genetic isolation of wild dog populations are associated with habitat fragmentation and loss. In the Okavango and Hwange monitoring populations, where continuous habitat still exists, gene flow connects wild dogs over considerable distances (~400 km), resulting in high genetic diversity and few indicators of inbreeding (Table 4.3). This highlights the enormous value of maintaining habitat connectivity and suggests that conservation priorities should be directed towards reinstating connections between smaller and more isolated populations, which are more prone to genetic threats and extinction. However, it is vital that studies are conducted to ensure habitat connectivity results in genetic connectivity. Continuous habitat exists between the monitoring area in Southern Kruger and Lowveld (~500 km apart), yet I found these two monitoring populations were highly differentiated (D_{est} 0.28-0.37; Appendix 4) and shown to constitute distinct genetic populations in STRUCTURE (Figure 4.5). The drivers of this isolation remain to be determined as wild dogs are able to disperse long distances (>250km, (Fuller et al. 1992; McNutt 1996) and there is no clear topographic barrier to wild dog movements (wild dogs must have crossed several wide perennial rivers when undertaking the long dispersal distances recorded by Fuller (1992). However, it is possible that the low and declining wild dog densities in Northern Kruger, combined with fencing of Kruger National Park in the 1970's has contributed to separation. The establishment of the Greater Limpopo Transfrontier Park and the removal of fences on the north eastern border of Kruger may help to establish connectivity between populations.

Given the isolation and small size of many wild dog populations combined with continuing demographic threats, it is not surprising that some population extinctions have occurred in recent years, including the Laikipia and Lowveld populations assessed here (Woodroffe 2001). It has been suggested that the high fecundity and large dispersal abilities of wild dogs may enable resilience to population extinction (Pole 2000; Woodroffe 2010); for example, the Laikipia wild dog population recovered from local extinction to >300 animals in just over 10 years (Woodroffe 2001). However, in my assessments of Laikipia and Lowveld, the founder effects associated with population recolonisation were evident from the low genetic diversity and high indications of inbreeding I found in comparison to other populations. These clear genetic costs associated with extinction events indicate that although wild dogs may have demographic resilience to extinction, this is not matched by genetic resilience. This is not surprising given that wild dogs have likely suffered repeated and prolonged bottlenecks. In the Serengeti-Mara, despite the disappearance of wild dogs from the monitoring area, genetic data shows that a population extinction did not occur (Chapter 3). In this case, habitat connectivity appears

to have enabled local re-establishment and prevented the loss of genetic diversity, and thus again reiterates the importance of maintaining these connections (Chapter 3). Nonetheless, the natural recolonisation events from unknown source populations in Laikipia and Lowveld, and persistence of a population thought to have gone extinct in Serengeti, highlight gaps in my knowledge about the distribution of, and connectivity between, populations of wild dogs. Further monitoring, both genetic and otherwise, is required to better understand distribution and connectivity in this species.

Artificial translocations and reintroductions have been used to imitate gene flow where habitat connectivity cannot be reinstated. However, to prevent outbreeding depression (Templeton 1986), it is critical to source animals from genetically similar populations (Edmands 2007). In this study, I found some evidence of selection and local adaptation at the DRB, which indicates that artificial translocations should be approached with caution to prevent outbreeding depression. Translocations have a long history in African wild dog conservation (Gusset et al. 2008; Gusset et al. 2006), but they have not previously been assessed genetically. I examined the genetic impact of a recent translocation of 16 wild dogs into Hwange; eight animals shown here to be genetically derived from the Lowveld, and eight animals whose provenance was not known and could not be determined genetically. The translocated animals introduced many new alleles, increasing genetic diversity at both microsatellite and DRB loci (microsatellite A_R increased from 6.6 to 7.6; DRB A_R increased from 7 to 12). However, at the DRB this resulted in considerable changes to the allelic repertoire and also to the proportions of DRB alleles derived from lineages A and B, which may have implications for local adaptation. Furthermore, the translocations also resulted in a noticeable increase in F_{IS} from -0.05 to 0.78, and a decrease in observed heterozygosity at both microsatellites (Table 4.3, 80% to 68%, paired t-test, p=0.022) and DRB (79% to 71%, Table 4.3), indicating that the translocated animals were derived from inbred sources. Both inbreeding and outbreeding depression are thought to be important genetic threats but their impact on different species has been very variable and difficult to predict (reviewed in Edmands 2007). As such, I suggest future translocations should only be used as a management tool where populations are inbred, and in these cases genetic information should be used to select source animals from outbred and genetically similar populations.

4.6 Conclusion

To the best of my knowledge, my study represents one of the most extensive genetic studies of a red list endangered species conducted to date. I combined neutral (microsatellite, mtDNA) and adaptive (MHC) markers to elucidate demographic history, gene flow, evidence of selection, spatial and temporal patterns of genetic diversity across eight monitoring populations of African wild dogs. I found a genetic signature of a large and recent demographic decline in African wild dogs from microsatellite data. This contrasts with findings in other African mammals which have shown major declines associated with climatic change at the end of the Holocene. Habitat fragmentation and loss appears to have resulted in strong population structuring of African wild dog populations, with limited gene flow between them. The spatial and temporal structure of microsatellite and DRB diversity were correlated, and appeared to be largely determined by demographic stability and size of populations. The correspondence between neutral and adaptive markers was not expected; selection is predicted to result in differences in the distribution of adaptive variation relative to neutral markers, the latter of which are only affected by neutral processes (Garrigan & Hedrick 2003; Schierup et al. 2000). However, a similar correspondence has been found in other species (Campos et al. 2006). This result suggests that selection may be unable to counter strong genetic drift in these small African wild dog populations (N_e<30), which results in contemporary patterns of DRB variation being largely determined by genetic drift. This indicates that natural evolutionary processes may have become compromised by demographic declines in this endangered species. Nonetheless, the pattern of genetic differentiation against geographic distance showed some evidence of balancing selection on DRB lineages and diversifying selection on DRB alleles. This may indicate that selection still operates on DRB variation, although the extent and strength of selection is weaker than expected for the MHC. Overall, my study has shown that the predominant factor determining spatial and temporal patterns of both neutral and adaptive genetic variation in wild dogs is demographic history, as population size appears to influence the strength of selection, rate of genetic drift, inbreeding, extinction probability and related distribution of populations and gene flow.

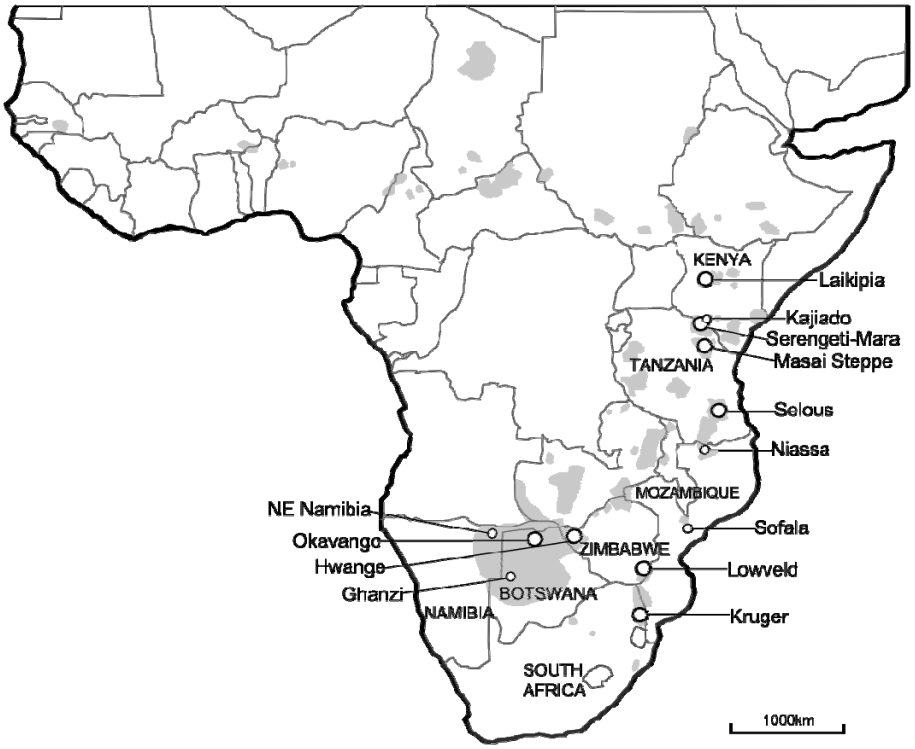
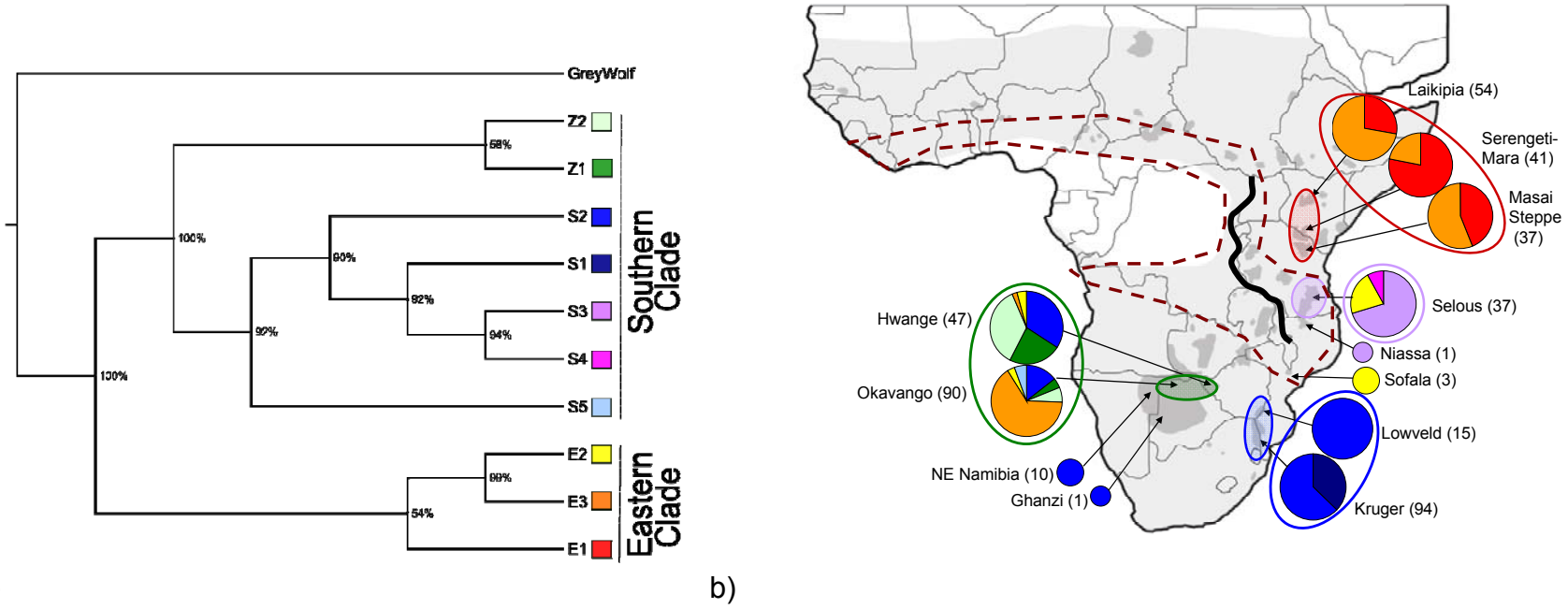


Figure 4.1: Historic (light grey) and present (dark grey) distribution of African wild dogs according to McNutt (2008). Locations of monitoring areas for which I have samples are depicted by circles. Sites represented by ≤5 samples, are shown as smaller circles.

```
       * * *            * * *      *         *         *   *     *  **  *   *      **  **
A1  HFLNVAKSECYFTNGTERVRFVDRYIYNREEFVRFDSDVGEFRAVTELGRPDAEYLNRQKEILEQERAAVDTYCRHNYGVGESFTVQRR
A2  --------------------------------------------------------------------------------I--------
A3  -------------------------------------------------------W------------------------I--------
A4  -------------------------------------------------------W---------------------------------
A5  -------------------------------------------------------W----------------V----------------
A6  -----------------------------------------Y-------------W------------------------I--------
A7  -------------------------------------------------------W-----L---R--E---V----------------
A8  -------------------------------------------------------W----------------V-------I--------
-----------------------------------------------------------------------------------------------
B1  --VYQF-G-------------LA-S------------------------------W-----L---R--E---V----------------
B2  --VYQF-G-------------LA-S----------------Y-------------W-----L---R--E--------------------
B3  --VYQF-G-------------LA-S----------------Y-------------R-----L---R--E--------------------
B4  --VYQF-G------------LLA-S------------------------------W-----L---R--E---V-------I--------
B5  --VYQF-G------------LLA-S------------------------------W-----L---R--E---V----------------
B6  --VYQF-G------------LLA-S------------------------------R-----L---R-----------------------
B7  --VYQF-G------------LLA-S----------------------------------------------------------------
B8  --VYQF-G------------LLA-S-----------------------------------------------V-------I--------
B9  --VYQF-G------------LLA-S------------------------------------L---R--E--------------------
B10 --VYQF-G-------------LA-S------------------------------W-----L---R--E---V-------I--------
B11 --VYQF-G------------LLA-S----------------Y-------------R-----L---R--E--------------------
B12 --VYQF-G------------LLA-S------------------------------R-----L---R-----------------------
B13 --VYQF-G------------LLA-S----------------Y-------------R-----L---R-----------------------
```

Figure 4.2: African wild dog DLA-DRB1 alleles aligned to allele A1. Matching amino acids are indicated with a dash, varying amino acids are indicated by single letter amino acid codes. Alleles are grouped into two phylogenetically divergent allelic lineages, A (above the line) and B (below the line). Amino acids of the putative peptide binding regions (Bondinas et al. 2007) are indicated by *.



a) b)

Figure 4.3: Mitochondrial DNA variation in African wild dogs. a) Bayesian geneology demonstrating the relationship between 10 African wild dog mtDNA haplotypes, rooted using the grey wolf as an outgroup. Node support is shown via Bayesian posterior probability values (%). Haplotype colours correspond with 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. Geographic clustering, as indicated by AMOVA, are shown with coloured ovals. The former and current distribution of wild dogs is depicted by light and dark grey shading, respectively, the distribution of miombo forest by dashed brown lines, and the western rift valley by a thick black line.

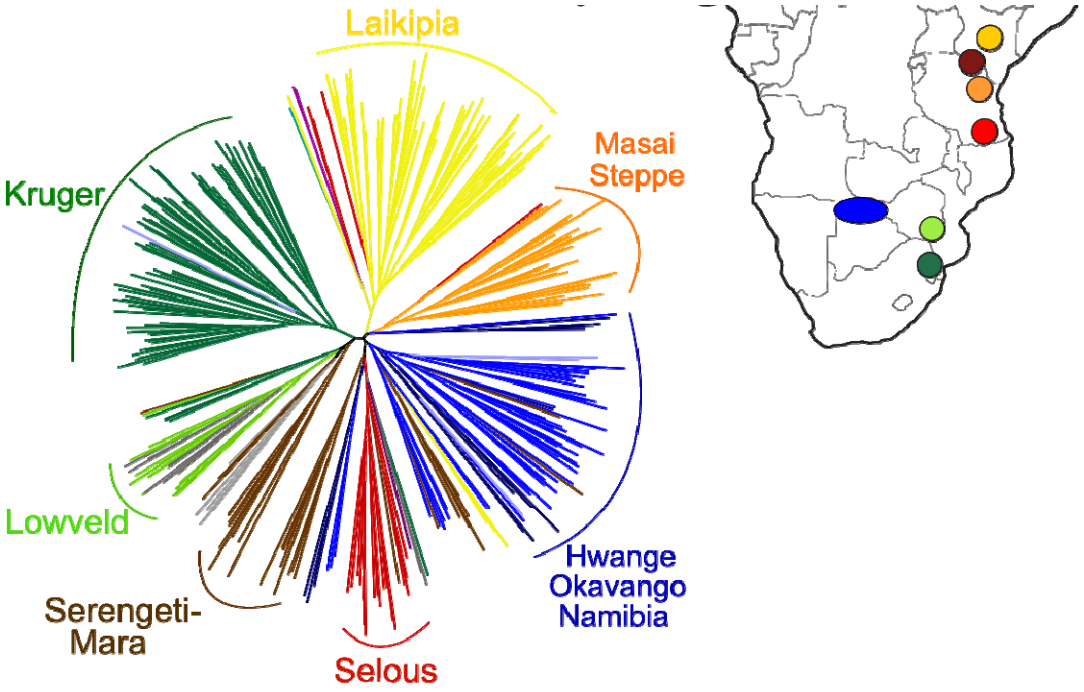


Figure 4.4: Unrooted neighbour-joining network based on Nei's allele sharing distance (DAS) across 10 microsatellite loci computed in Populations. Colours of branches indicate sampling location of the individual as reflected on the map: Kruger – dark green, Lowveld – light green, Namibia –turquoise, Okavango – blue, Hwange – dark blue, Sofala – pink, Selous – red, Masai Steppe – Orange, Serengeti-Mara –brown, Laikipia – yellow, 16 wild dogs of unknown provenance translocated into Hwange - grey.

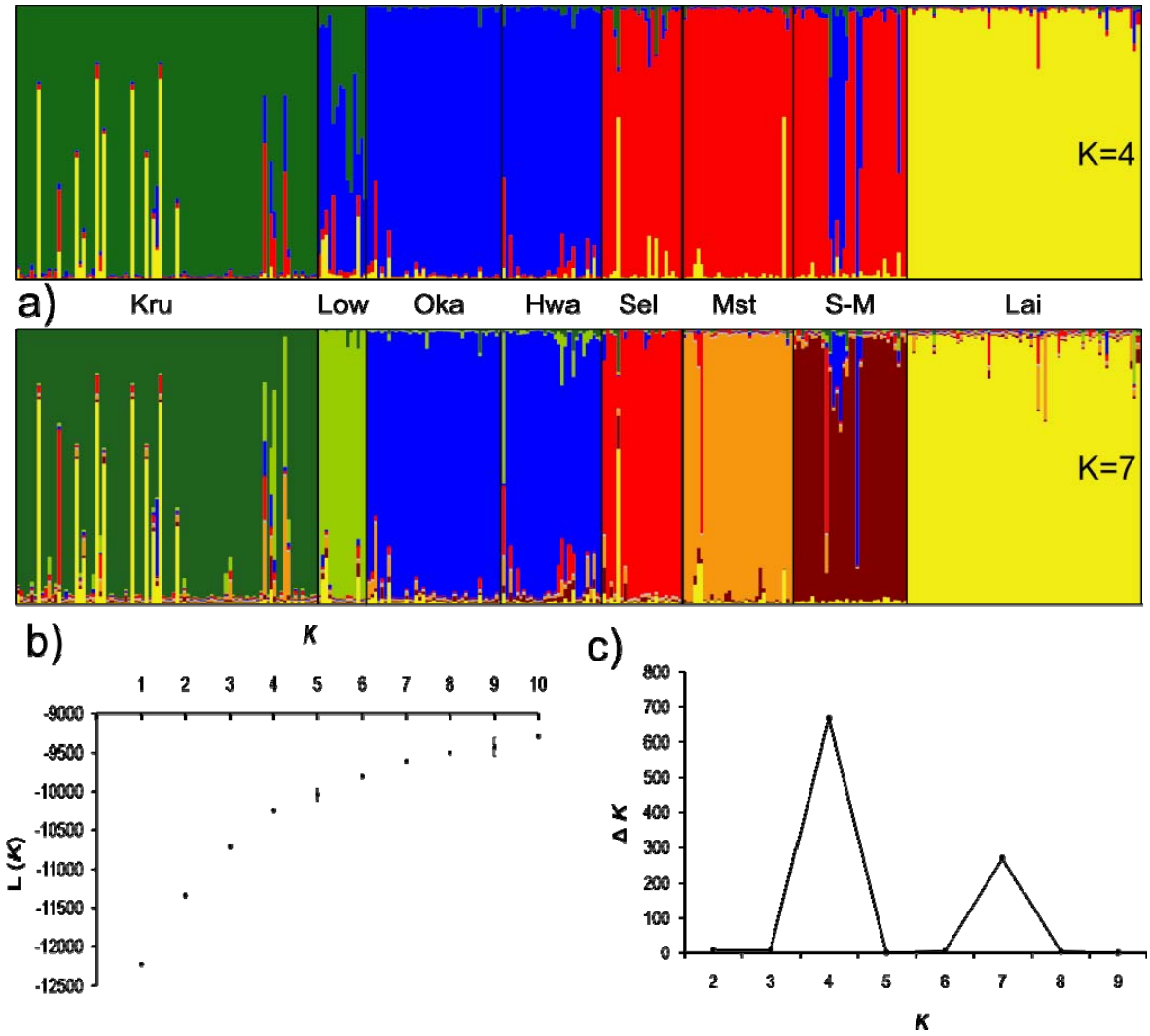


Figure 4.5: a) Genetic structure of African wild dog populations based on Bayesian clustering analyses (STRUCTURE) of samples at 10 microsatellite loci. Shown is the most likely level of population clustering (K=7) as indicated by the ΔK statistic (c). Columns are individuals, with the proportion of an individual's genotype assigned to each cluster (K) denoted by different colours. Colours correspond with location map in Figure 4.4. Populations are depicted by three letter codes; Kru-Kruger, Low-Lowveld, Oka-Okavango, Hwa-Hwange, Sel-Selous, Mst-Masai Steppe, S-M-Serengeti-Mara, Lai-Laikipia. b) Likelihood probability profile estimated from STRUCTURE 2.2 at K1-10 showing the mean and variance at each K. c) ΔK at each value of K, averaged across 10 replicates. The ΔK statistic identified two peaks where K=4 represented the uppermost level of structuring and K=7 which represents finer scale population structuring (Evanno et al. 2005).

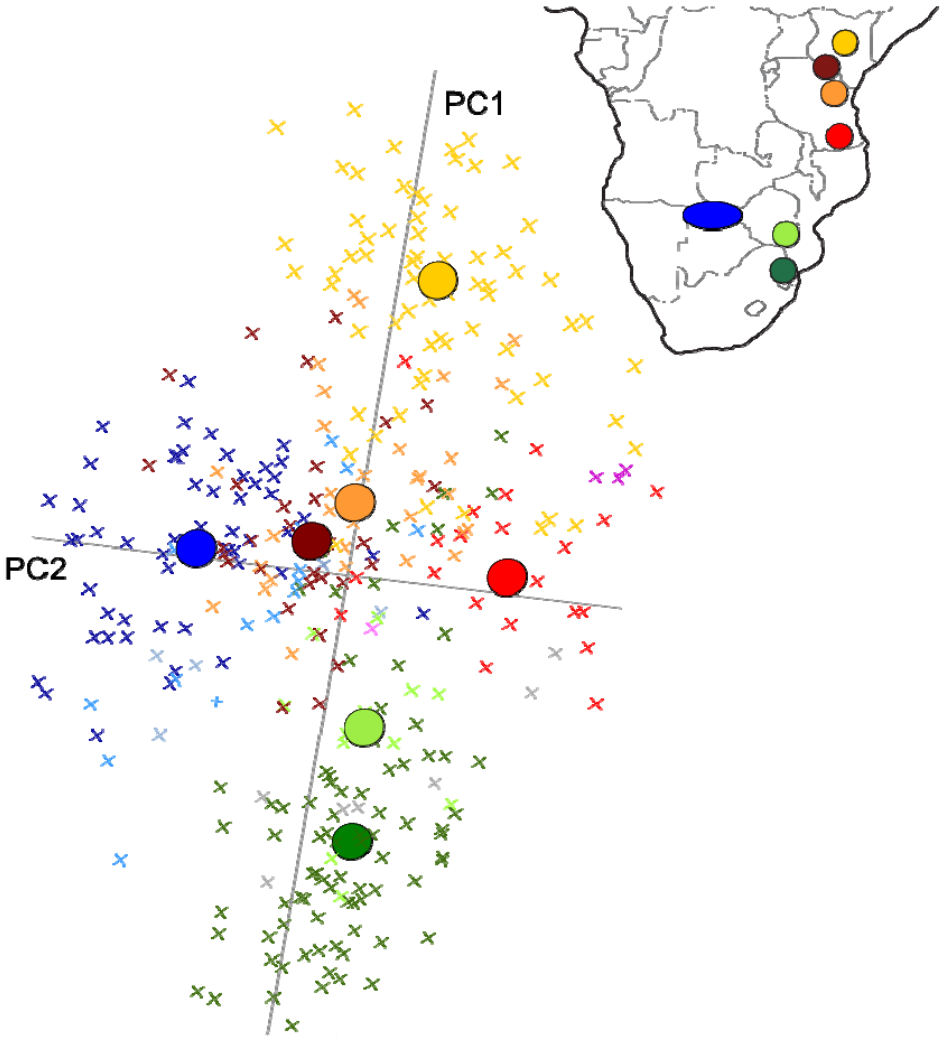


Figure 4.6: Principle coordinates analysis of African wild dog populations computed in GenALEX. Small + symbols represent individuals, and large circles represent population medians. Colours correspond to sampling location indicated on the map. The axes have been rotated to reflect the resemblance between the PCO and geographic sampling location. PC 1 (27.5 %) appears to represent a North South axis and PC 2 (18.4%) a East West axis.

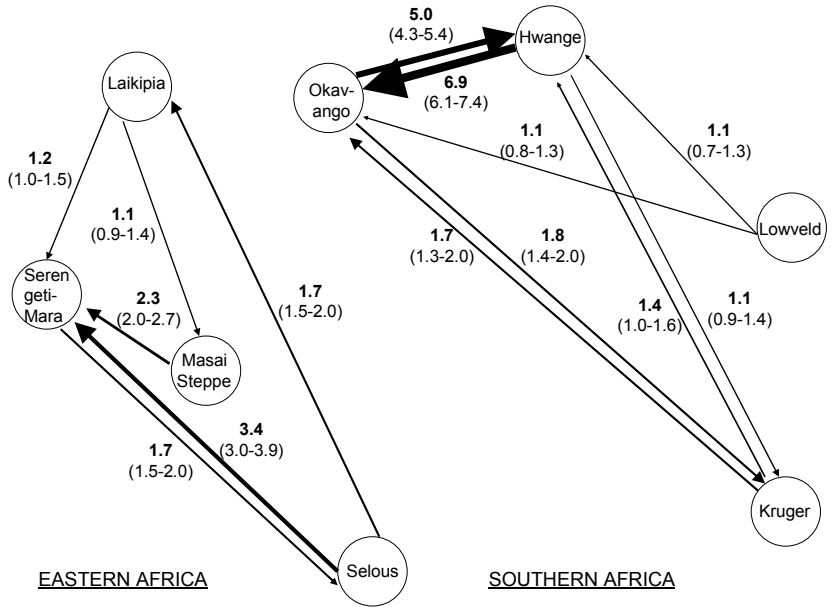


Figure 4.7: Mean long term migration rates between African wild dog populations in Eastern and Southern Africa. Estimates are based on number of animals/generation with 95% confidence intervals calculated as θ*M, however, these values should be viewed with caution and should not be interpreted literally (see text for more details). To improve clarity of the figure, migration rates were removed if the 95% confidence interval did not include 1. Arrow sizes are proportional to migration rate.

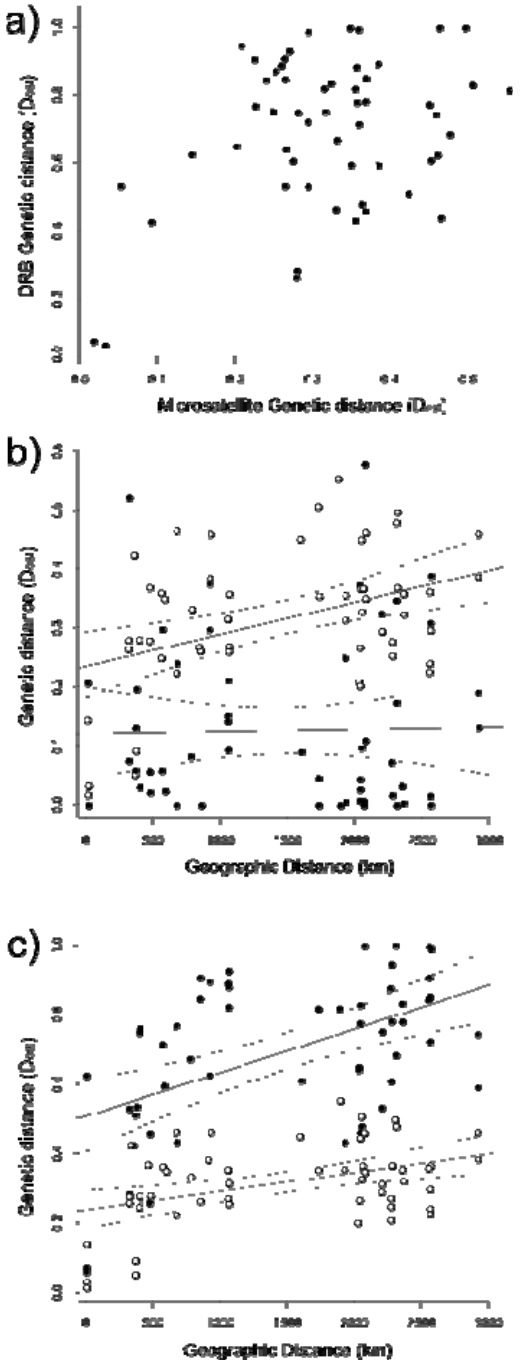


Figure 4.8: a) Correlation of pairwise genetic differentiation (Dest) between populations for DRB and microsatellite markers (R^2 = 0.141, p<0.01). b) Regression of pairwise genetic distance (Dest), against geographic distance (km) for DRB (white circles; R2=0.26, p<0.0001, intercept = 0.508, slope = 0.000125), and microsatellite (filled circles; R2=0.17, p<0.05, intercept = 0.238, slope = 0.000054) loci . c) Regression of pairwise genetic distance (Dest), against geographic distance for DRB lineages (white circles; R2<0.0001, p=0.881, intercept = 0.121, slope =0.0000034) and microsatellite (filled circles; statistics in b) loci. DRB lineages refers to the lineage or lineages an individual's alleles were derived from, i.e. AA, AB, BB.

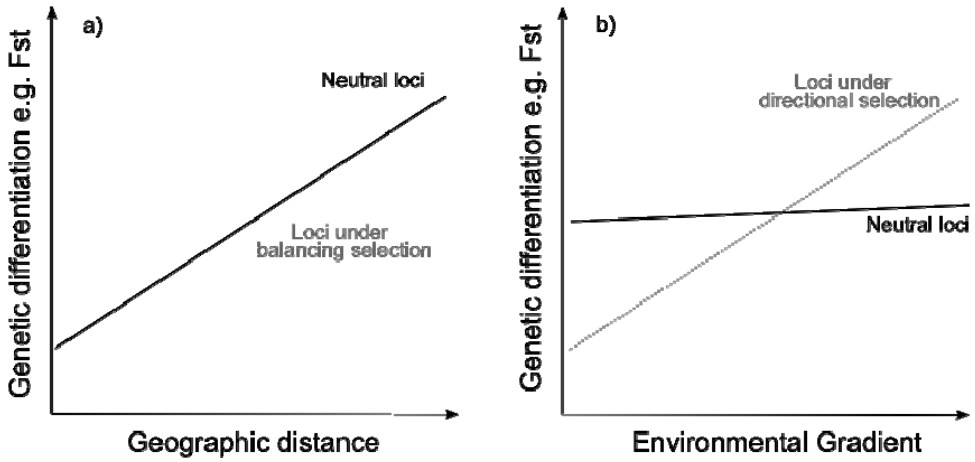


Figure 4.9: Theoretical expectations of how genetic differentiation changes with geographic distance and across an environmental gradient based on loci evolving neutrally and under selection changes (Redrawn from (Gebremedhin et al. 2009). a) Differentiation increases with geographic distance at neutral loci as a result of random differences accumulating between isolated populations. Balancing selection is expected to result in lower levels of differentiation between populations due to selection for high variation which results in selection for any new migrant alleles. b) Across an environmental gradient, local selection pressures result for differences accumulating more quickly at loci under directional selection than at neutral loci.

Table 4.1: Details of the origin of the African wild dog samples used in this study including the name and country of the monitoring area, the sampling years, number of animals and packs sampled, number of animals in monitoring area (Npop) and monitoring area three letter abbreviation.

Monitoring area	Country	Sampling years	n	packs	Npop	Abbreviation
Kruger	SE South Africa	Old:1990-95 Recent: 2007-08	67 24	≥9 6	400 100	Kru
Lowveld	SE Zimbabwe	2008-2009	15	6	130	Low
Okavango	N Botswana	Old: UNK Recent: 2000-06	19 42	* ≥15	200 200	Oka
Hwange	NW Zimbabwe	~1990-2007	19	*	250	Hwa
Selous	S Tanzania	1991-1995	23	8	100	Sel
Serengeti-Mara	Tanzania-Kenya border	Old:1981-90 Recent: 2005-07	20 14	≥5 5	50 160	S-M
Masai-Steppe	NE Tanzania	1995	32	3	UNK	Mst
Laikipia	Central Kenya	2003-2008	67	14	300	Lai
Monitoring areas where n≤5						
Ghanzi	W Botswana	2006	1	1		
NE Namibia	NE Namibia	1980-2007	5	4		
Sofala	Central Mozambique	2004	3	1		
Niassa	N Mozambique	2008	1	1		
Kajiado	S. Kenya	2008	1	1		

* pack information was not known for all monitoring areas.

Table 4.2: Sample sizes (n) and distribution of mtDNA haplotypes across African wild dog populations. Private haplotypes are underlined. Okavango and Serengeti-Mara represent frequencies based on both Old and Recent samples.

	n	S1	S2	S3	S4	S5	Z1	Z2	E1	E2	E3
Kruger Old	94	35	59								
Lowveld	15		15								
NE Namibia*	10		10								
Okavango	90		13			5	4	6	59	3	
Ghanzi	1		1								
Hwange	47		16				11	17	1	2	
Sofala	3									3	
Niassa	1			1							
Selous	37			26	3					8	
Serengeti-Mara*	41								32		9
Masai-Steppe	32								14		18
Laikipia	54								15		39
Total	**425**	**35**	**123**	**27**	**3**	**5**	**15**	**23**	**121**	**16**	**66**

*Four samples from this study were combined with data from six samples from (Girman et al. 2001).

Table 4.3: Summary of diversity statistics for mtDNA (mt), microsatellite (ms) and MHC markers. Listed are the number of animals (n) and packs (pk) typed at each marker in each population, number of haplotypes (*h*), number of private alleles or haplotypes (*PA*), nucleotide diversity (π), Gene diversity (Gdiv), Allelic richness (A_R), Allelic richness standardized for sample size ($Std\text{-}A_R$), Observed heterozygosity (H_o), Expected heterozygosity (H_e), Fixation index (F_{IS}). # indicates F_{IS} values significantly different from 0. Standard Error values are shown in brackets.

			mtDNA				Microsatellites						MHC					
Population	**n (mt/ms/MHC)**	**# pk**	**# *h***	**# *PA***	**π x10^3**	***Gdiv***	***A_R***	***Std-A_R***	**# *PA***	***H_o***	***H_e***	**F_{IS}**	**A_R**	***Std-A_R***	**# *PA***	**H_o**	**H_e**	**F_{IS}**
Kruger Old	94/67/67	≥9	2	1	1.45 (±1.4)	0.47 (±0.03)	6.3 (1.0)	4.60	2	0.78	0.70	-0.11	6	5.1	2	0.82	0.76	-0.08
Kruger Recent	0/20/24	6	na	na			5.4 (1.8)	4.53		0.66	0.67	0.02	6	5.5		0.67	0.66	-0.02
Lowveld	15/14/15	6	1		0	0	4.6 (0.7)	4.09	1	0.62	0.59	-0.05	3	3.0		0.73	0.60	-0.22
Okavango Old	90^/12/19	*	6	1	22.20 (±11.6)	0.54 (±0.06)	4.9 (0.4)	4.54	5	0.70	0.68	-0.03	9	8.3	2	0.84	0.86	0.02
Okavango Recent	-/42/42	15					7.4 (0.8)	5.05		0.79	0.74	-0.07	9	8.0		0.81	0.85	0.05
Hwange	47/14/19	*	5		11.20 (±6.4)	0.71 (±0.03)	6.6 (0.9)	5.68	2	0.80	0.76	-0.05	7	6.6		0.79	0.79	0.00
Selous	37/23/22	8	3	1	19.42 (±10.5)	0.47 (±0.08)	6.3 (0.8)	4.90	3	0.68	0.68	0.00	11	9.4	2	0.91	0.88	-0.03
Masai Steppe	32/32/17	3	2		6.18 (±4.0)	0.34 (±0.08)	5.5 (±0.7)	4.31	3	0.62	0.61	-0.02	5	5.0		0.83	0.71	-0.17
Serengeti-Mara Old	41^/20/18	≥5	2		9.32 (±5.5)	0.51 (±0.03)	6.1 (±0.7)	4.97	6	0.69	0.72	0.03	5	4.8	1	0.67	0.74	0.09
Serengeti-Mara Recent	-/13/14	5					4.6 (±0.5)	4.31		0.65	0.67	0.03	6	6		0.93	0.76	-0.22
Laikipia	54/67/63	≥14	2		7.50 (±4.6)	0.41 (±0.06)	6.0 (±0.8)	4.41	8	0.61	0.67	0.11#	3	3		0.54	0.57	0.05
Mean			2	0.4	7.45 (±2.3)	0.36 (±0.07)	5.8	4.67	3.8	0.69	0.68	-0.01	6.4	5.88	0.9	0.78	0.74	-0.05

^ This value includes old and recent samples for that population. * Information on the numbers of packs sampled was not available for all populations.

Table 4.4: Frequencies of MHC-DRB1 alleles (A1-A8 and B1-B11) and A and B allelic lineages in each population. The last three columns show the proportion of individuals that are homozygous for lineage A (AA) or B (BB), and those that have an allele from each lineage (AB). Private alleles and the most common allelic lineage within populations are underlined.

		Frequency of DRB alleles per population																		Frequency alleles/lineage		DRB lineage genotypes		
Population	**n**	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8**	**B1**	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**	**B8**	**B9**	**B10**	**B11**	**A**	**B**	**AA**	**AB**	**BB**
Kruger-Old	**67**			0.08	0.23			0.01						0.11	0.33					0.23	0.30	0.70	0.11	0.39	0.50
Kruger-Recent	**24**			0.06	0.17			0.02						0.08	0.54					0.13	0.25	0.75	0.04	0.42	0.54
Lowveld	**15**	0.1									0.4				0.5						0.10	0.90		0.20	0.80
Okavango-Old	**19**	0.16	0.24	0.13	0.03						0.21		0.11	0.05		0.03		0.05			0.55	0.45	0.42	0.26	0.32
Okavango-Recent	**42**	0.1	0.21	0.24	0.06						0.05		0.18	0.02		0.08		0.06			0.61	0.39	0.45	0.31	0.24
Hwange	**19**	0.08		0.05	0.13						0.29	0.03	0.08			0.34					0.26	0.74	0.05	0.42	0.53
Selous	**22**			0.02	0.11	0.18			0.05	0.11	0.02		0.09	0.02	0.18		0.18		0.02		0.36	0.64	0.14	0.45	0.41
Masai Steppe	**18**					0.14				0.06			0.36		0.14	0.31					0.15	0.85	0.06	0.18	0.76
Serengeti-Mara Old	**18**			0.03			0.36			0.11			0.33			0.17					0.39	0.61	0.22	0.33	0.44
Serengeti-Mara Recent	**14**				0.14	0.43	0.18							0.04		0.18	0.04				0.75	0.25	0.50	0.50	0.00
Laikipia	**63**				0.59							0.16	0.25								0.59	0.41	0.35	0.48	0.17
Areas with ≤5 samples																									
NE Namibia	**4**	0.33		0.17	0.33						0.17														
Ghanzi	**1**			1																					
Niassa	**1**												0.5				0.5								
Sofala	**3**					1																			
Kajiado	**1**					0.5							0.5												

Table 4.5: Estimates and 95% confidences intervals of effective population size based on three methods. Long term effective population size derived from MIGRATE based on $\theta = 4N_e\mu$ (Beerli 2006), moments based temporal method for populations with temporal sampling, and the linkage disequilibrium method. N_e/Npop estimates are provided for the linkage disequilibrium method; N pop taken from Table 4.1.

Population	n	Ne MIGRATE (long term)	Ne Temporal Method	Ne Linkage Disequilibrium	Ne/Npop Linkage Disequilibrium
Kruger All	87	26.4 (24.8-27.7)	21.8 (11.6-43.5)		
Old	67			19.6 (17.3-22.2)	0.05
Recent	20			10.2 (8.4-12.7)	0.10
Lowveld	14	27.5 (23.2-30.5)		7.8 (6.1-10.5)	0.06
Okavango All	54	34.9 (31.9-36.9)	28.3 (13.1-92.7)		
Old	12			10.4 (7.7-15.1)	0.05
Recent	42			16.6 (14.7-18.8)	0.09
Hwange	14	47.7 (40.2-52.4)		12.4 (9.9-15.9)	0.05
Selous	23	108.0 (100.9-117.3)		20.5 (16.2-27.0)	0.21
Serengeti-Mara All	33	36.2 (33.6-38.9)	11.9 (6.8-20.8)		
Old	20			12.6 (10.4-15.7)	
Recent	13			3.0 (2.6-3.6)	0.02
Masai Steppe	32	30.5 (28.6-32.7)		11.5 (9.8-13.6)	NA
Laikipia	67	29.1 (27.7-30.5)		21.5 (18.7-24.8)	0.07

Table 4.6: Findings from bottleneck tests conducted on 10 microsatelllite loci from five populations. Results are based on heterozygosity excess and M ratio test. All allele frequency distributions were normal.

		^Heterozygosity excess p values where SMM =				˜M ratio p values where θ =			
Population*	**n**	70%	80%	90%	M ratio	0.2	0.6	1.2	2
Kruger	87	0.002	0.007	0.065	0.677	0.017	0.005	0.003	0.004
Botswana	54	0.188	0.461	0.813	0.754	0.001	0.000	0.000	0.000
Masai Steppe	32	0.138	0.313	0.500	0.636	0.070	0.032	0.016	0.009
Serengeti	33	0.001	0.003	0.065	0.674	0.019	0.007	0.002	0.001
Laikipia	67	0.007	0.065	0.188	0.741	0.002	0.000	0.000	0.000

*Kruger, Botswana and Serengeti are the combined temporal sample set.
^ Heterozygosity excess assumed the two phase mutation model, varying propotions of step wise mutation model (SMM). Significance was tested with wilcoxon tests.
˜M ratios tests assumed the proportion of one-step mutations as 90% and the size of multi-step mutations as 3.5. p values show the probability of the M ratio being significantly lower than values expected under mutation-drift equilibrium at a range of θ values corresponding to pre-bottleneck population sizes of 100, 300, 600 and 1000 were tested.

Table 4.7: Estimates of ancestral (N1) and current (N0) effective population sizes and time since decline (T), for three populations, derived from coalescent simulations in MSVAR. Results show the median value, with 5% - 95% quantiles shown in brackets.

Population	**N_1** (5 - 95%)	**N_0** (5 - 95%)	**T** (5 - 95%)
Kruger	865 (225-3278)	1 (0.1 - 7)	10 (1 - 82)
Okavango	938 (284-3031)	2 (0.1 - 24)	14 (1 - 154)
Selous	670 (181-2439)	1 (0.1 - 20)	9 (1-123)

Table 4.8: Results from hierarchical analysis of molecular variance (AMOVA) computed in Arlequin v 3.11. Alternative a priori hypotheses of population groupings were tested to identify groupings that explain more variation (%var) between groupings (Φ_{CT}) than within, as is indicative of population structure. Significance was assessed using 1,000 permutations and is indicated by bold type. Populations are described by three letter codes (Table 4.1).

	mtDNA			*Msats*			*MHC*		
Grouping tested	df	%var	p	df	%var	p	df	%var	p
Southern-Eastern									
[Kru,Low,Oka,Hwa] [Sel,S-M,Mst,Lai]									
Among Groups [Φ_{ct}]	1	26.88	0.059	1	1.99	**0.019**	1	1.84	0.213
Among Populations [Φ_{sc}]	6	43.65	**<0.001**	6	13.54	**<0.001**	6	18.31	**<0.001**
Within Populations [Φ_{st}]	395	29.47	**<0.001**	640	84.47	**<0.001**	632	79.85	**<0.001**
Southern-Central-Eastern									
[Kru,Low,Oka,Hwa] [Sel] [S-M,Mst,Lai]									
Among Groups [Φ_{ct}]	2	40.97	0.053	2	2.35	0.052	2	0.40	0.47
Among Populations [Φ_{sc}]	5	30.71	**<0.001**	5	13.38	**<0.001**	5	19.26	**<0.001**
Within Populations [Φ_{st}]	395	28.32	**<0.001**	640	84.26	**<0.001**	632	80.34	**<0.001**
SouthEast-SouthWest-Central-Eastern									
[Kru,Low] [Oka,Hwa] [Sel] [S-M, Mst, Lai]									
Among Groups [Φ_{ct}]	3	46.34	**0.048**	3	2.72	**0.028**	3	2.87	0.182
Among Populations [Φ_{sc}]	4	22.62	**<0.001**	4	12.75	**<0.001**	4	17.02	**<0.001**
Within Populations [Φ_{st}]	395	31.03	**<0.001**	640	84.53	**<0.001**	632	80.12	**<0.001**
SouthEast-SouthWest-Eastern-North East									
[Kru,Low] [Oka,Hwa] [Sel,S-M, Mst] [Lai]									
Among Groups [Φ_{ct}]	3	17.87	0.282	3	2.05	0.068	3	10.41	**0.010**
Among Populations [Φ_{sc}]	4	49.57	**<0.001**	4	13.21	**<0.001**	4	9.96	**<0.001**
Within Populations [Φ_{st}]	395	32.56	**<0.001**	640	84.73	**<0.001**	632	79.63	**<0.001**
SouthEast-SouthWest-Central-Eastern-North East									
[Kru,Low] [Oka,Hwa] [Sel] [S-M, Mst] [Lai]									
Among Groups [Φ_{ct}]	4	38.82	0.133	4	3.81	**0.008**	3	10.13	**0.005**
Among Populations [Φ_{sc}]	3	29.09	**<0.001**	3	11.50	**<0.001**	4	10.02	**<0.001**
Within Populations [Φ_{st}]	395	32.09	**<0.001**	640	84.69	**<0.001**	632	79.86	**<0.001**

Chapter 5: The importance of considering phylogenetic relatedness when interpreting patterns of MHC polymorphism in endangered species.

5.1 Abstract

The major histocompatibility complex (MHC) is renowned for its high levels of variation, which results from balancing selection. This diversity is thought to be important to enable diverse immune capabilities. There is concern, though, that reduced population sizes in endangered species will result in increased levels of genetic drift and decreased levels of selection, resulting in an overall loss of MHC diversity. As a result, a range of studies have examined MHC variation in endangered species and assessed the level of diversity through comparisons with non-endangered reference taxa. However, the reference taxa used are commonly distant relatives or domesticated species. Therefore, any differences in levels of diversity may be the result of differences in evolutionary history rather than endangered status. Here, I assessed MHC variation across eight species from the wolf-like canid clade to investigate the potential factors that influence patterns of MHC diversity. As expected, demographic history and selection were important factors influencing MHC diversity. However, I also found evidence that phylogeny influences levels of MHC diversity and trans-specific polymorphism which suggests phylogenetic relatedness can confound patterns of MHC diversity. Furthermore, it was shown that comparative MHC studies that fail to control phylogeny may lead to inaccurate interpretation of MHC data. This has important implications for the choice of reference taxa used in comparative studies aiming to classify levels of diversity in (endangered) species. Clearly, unless very closely related species are compared, it will be difficult to disentangle whether any differences in diversity between species are the result of population declines rather than evolutionary history.

5.2 Introduction

The major histocompatibility complex (MHC) includes multiple genes that code for a set of cell-surface molecules involved in the recognition of foreign antigens as part of the immune response (Klein 1980; Piertney & Oliver 2006). A defining feature of MHC genes is their high level of polymorphism (Garrigan & Hedrick 2003), which is thought to confer immunity against a larger range of diseases (Doherty & Zinkernagel 1975; Sommer et al. 2002). This diversity is largely attributable to two characteristics, both of which are thought to result from pathogen-driven balancing selection: 1) large numbers of alleles (hundreds of alleles per locus have been reported in some species (HLA-Database 2010); and 2) high sequence divergence amongst alleles arising from retention of ancestral polymorphism (reviewed in Garrigan & Hedrick 2003). Evidence that balancing selection operates on the MHC has been demonstrated by an excess of non-synonymous variation ($d_N > d_S$) amongst MHC alleles, as well as the concentration of non-synonymous changes to the residues intricately involved with the recognition of foreign antigens (the peptide binding region, PBR, (Furlong & Yang 2008), i.e. the sites where amino acid changes are most likely to result in functional changes. Balancing selection is also evidenced by the retention of ancestral alleles/allelic lineages, that is, trans-specific polymorphism (Klein 1980).

Whilst the influence of balancing selection has been repeatedly demonstrated on the MHC, the ability of selection to maintain diversity is very sensitive to effective population size (N_e) (Richman 2000). When N_e is small, genetic drift replaces selection as the dominant evolutionary force, resulting in the random loss of variation (reviewed in Sommer 2005). As such, endangered species that experience population declines or bottlenecks are predicted to lose diversity at the MHC and therefore potentially have increased susceptibility to disease (O'Brien & Evermann 1988). The best example of this comes from Tasmanian devils, where a catastrophic population crash has been linked to an inability to mount an immune response against an emerging disease as a direct result of a lack of MHC variation due to previous bottlenecks (Siddle et al. 2007). Overall, therefore, MHC diversity is thought to be important for both population and species viability (Hoglund 2009; Piertney & Oliver 2006). As such, assessments of MHC variation are increasingly conducted in endangered taxa, with the aim to: 1) identify whether population declines have resulted in the loss of MHC diversity; and 2) assess whether selection still operates on the MHC in endangered species. Since pre-decline samples are rarely available, declines in diversity in endangered species are typically inferred through comparison with reference taxa (e.g. Hedrick et al. 1999; Marsden et al. 2009). However, unless sister taxa are available (e.g. Miller & Lambert 2004), using species comparisons to determine whether a species has lost MHC variation because of

demographic declines is confounded by the possibility that differences in diversity between species are the result of evolutionary history; particularly, if reference taxa are distantly related (e.g. Smith et al. 2010; Zhu et al. 2007) or domesticated species (e.g. Hedrick et al. 1999; Mikko & Andersson 1995). Investigations into the potential influence of evolutionary history on MHC diversity are currently lacking in the literature, and thus typically not considered in MHC comparisons. As such, there is a need for research into the contribution of both evolutionary history and endangered status to MHC diversity through studies of multiple closely related taxa.

The Canidae is a relatively young family that diverged very rapidly into 16 genera and 36 species within the last 8-10 million years (Lindblad-Toh et al. 2005; Perini et al. 2009). From a phylogenetic perspective, the Canidae creates an immense challenge because there has been insufficient evolutionary time for the accumulation of many genetic differences between species, which means very large amounts of sequence data are required to resolve evolutionary relationships (Lindblad-Toh et al. 2005). However, for an evolutionary study of the MHC, this is a great advantage. A diverse but recently evolved family with a number of closely related species is ideal for a comparative MHC study investigating the factors that affect patterns of diversity and selection in different species.

Recent phylogenetic analyses of the Canidae based on ~15kb of intron and exon sequence found strong support for four clades: 1) the red-fox like canid clade; 2) the South American canid clade; 3) the wolf-like canid clade; and 4) the grey and island fox clade (Lindblad-Toh et al. 2005). Furthermore, these analyses resolved the phylogenetic relationships amongst the wolf-like canids (excluding the Red wolf, *Canis rufus*): Grey wolves (*Canis lupus*); Coyotes (*Canis latrans*); Ethiopian wolves (*Canis simensis*); Golden jackals (*Canis aureus*), Dholes (*Cuon alpinus*); African wild dogs (hereafter wild dog, *Lycaon pictus*); Black-backed jackals (*Canis mesomelas*), and Side-striped jackals (*Canis adustrus*). Importantly, both Lindblad-Toh (2005) and a subsequent study by Perini (2009) found the genus *Canis* to be non-monophyletic. Specifically, the Black-backed and Side-striped jackals were found to be highly divergent to the other *Canis* species, and basal to the wolf-like canid clade (Figure 5.1), although these species have yet to be renamed.

The wolf-like canids consist of a combination of both abundant (IUCN category, least concern: Grey wolves; Coyotes; Golden jackals; Black-backed jackals; Side-striped jackals) and endangered taxa (Ethiopian wolves; Dholes; African wild dogs) (IUCN 2010). Based on a comparison with highly abundant Grey wolves and two other endangered taxa (Ethiopian wolves and Mexican wolves, which currently have subspecies status, *Canis lupus baileyi*), previous research found that patterns of MHC in wild dogs, were atypical

for a wolf-like canid, even for an endangered species (Figure 5.1; Chapter 2; (Marsden et al. 2009). Firstly, wild dogs lack variation at the DLA-DQA1 (1 allele, hereafter DQA) and DLA-DQB1 (2 alleles, hereafter DQB) loci, which have been shown to be variable in other wolf-like canids. Secondly, wild dogs have just two allelic lineages at the DLA-DRB1 locus (hereafter DRB), whereas DRB alleles in other wolf-like canids, are derived from a number of very divergent allelic lineages. Thirdly, wild dogs share no alleles with the other wolf-like canids, whereas allele sharing is common amongst the other species in that clade. These differences led to the suggestion that wild dogs may have lost MHC variation as a result of widespread demographic declines which have reduced the global wild population to less than 8,000 (IUCN/SSC 2008, 2009). However, these differences may be the result of wild dogs being more distantly related to the other wolf-like canids surveyed, which were all from the genus *Canis*. In addition, wild dogs are not known to hybridise with other species, whereas hybridisation within the genus *Canis* is extensive (Gottelli et al. 1994; Lehman et al. 1991; Verginelli et al. 2005; Vilà et al. 1997).

To distinguish between the relative influence of evolutionary and demographic history on patterns of MHC variation, I surveyed diversity of DRB, DQB and DQA in Coyotes, Golden jackals, Dholes, Black-backed jackals, and Side-striped jackals for comparison with my previous study (Marsden et al. 2009). The expanded dataset includes nearly all extant wild species (except the red wolf) in the wolf-like canid clade (Figure 5.1), to provide a finer scale picture of MHC diversity in abundant and endangered species in relation to phylogenetic history. Specifically, I addressed the following four questions: 1) Do endangered species show different patterns of diversity at MHC loci than abundant species? 2) Is there evidence for differences in selection pressures on endangered versus abundant species? 3) Does hybridisation affect interpretations of selection and diversity? 4) Are patterns of selection and diversity confounded by phylogenetic history?

5.3 Methods

5.3.1 Samples

To compare DRB, DQB and DQA variation within the wolf-like canid clade (Figure 5.1), I accessed samples or published MHC data from all wolf-like canid species (Table 5.1) with the following exceptions: Domestic dogs (*Canis familiaris*) were specifically excluded as they have been under intense artificial selection; Red wolves and Mexican wolves were excluded due to their absence in the most recent phylogeny (Lindblad-Toh et al. 2005), small data sets available (Red wolf) and in the case of the Mexican wolf, due to their sub-species status. Extracted DNA was sourced from the R. K. Wayne canid sample bank for 25 Coyotes, 13 Golden Jackals, four Dholes, eight Side-striped Jackals and 83 Black-backed jackals. A further 7 Dholes samples were provided by A. Ivengar and A. Kitchener. Wild dog DRB data, were taken from Chapters 2 (Marsden et al. 2009) and 4; wild dog DQB and DQA data were based on Chapters 2 as well as some additional sequencing of 164 individuals at the DQB and 37 individuals at DQA. All data for Grey wolves, Ethiopian wolves, and data for four additional Coyotes were taken from published sources (Table 5.1): Grey wolves (Kennedy et al. 2007a; Kennedy et al. 2001; Seddon & Ellegren 2002; Seddon & Ellegren 2004); Ethiopian wolves (L.J.Kennedy *Pers comm.*; GenBank; (Marsden et al. 2009); Coyotes (Seddon & Ellegren 2002). Sequences were downloaded directly from GenBANK or provided by the DLA nomenclature committee (L.J. Kennedy). Table 5.1 details sample sizes used in this study. In some cases, the numbers of individuals typed per locus varied across the three loci. This reflects the fact that data was excluded where it was not possible to make confident and unambiguous allele assignments.

5.3.2 Sequence Based Typing

Sequence-based typing was conducted on exon 2 of the DLA-DRB1, DLA-DQA1, and DLA-DQB1 loci using locus-specific intronic domestic dog primers that gave products of 303 bp (DLA-DRB1), 345 bp (DLA-DQA1), and 300 bp (DLA-DQB1). Primers were as follows (M13 and T7 tails are underlined): DRBln1: ccg tcc cca cag cac att tc (Wagner et al. 1996); DRBln2M13r: <u>cag gaa aca gct atg acc</u> tgt gtc aca cac ctc agc acc a (Wagner et al. 1996); DQAln1: taa ggt tct ttt ctc cct ct (Wagner et al. 1996); DQAln2: gga cag att cag tga aga ga (Wagner et al. 1996). DQB1BT7 <u>taa tac gac tca cta tag gg</u> ctc act ggc ccg gct gtc tc (Wagner et al. 1996); DQBR2: cac ctc gcc gct gca acg tg (Kennedy et al. 2002a).

Polymerase chain reactions (PCR) were performed in 25 µl reactions containing 1 x Q solution (Qiagen), 1 x PCR buffer containing 15 mM MgCl2 (Qiagen), 1 mM MgCl

(Qiagen), 0.4 mM of each DNTP (Invitrogen, San Diego, CA), 0.04 uM of each primer, 0.1 µg/µl BSA (Promega) OR 25mM of Dimethyl sulfoxide, 1 unit of Hot Star taq (Qiagen), and approximately 25 ng of template DNA. To detect contamination, each PCR was run with both the DNA extraction negative and a PCR-negative control containing no template DNA. Reactions were run on PTC-200 DNA engine machines (MJ Research Inc.). PCR amplifications were conducted with a touchdown protocol: 15 min at 95 ºC; 14 touchdown cycles of 95 ºC for 30 s; followed by 1 min annealing, starting at 62 ºC (DLA-DRB1), 52 ºC (DLA-DQA1), 68 ºC (DLA-DQB1), and reducing at 0.5 ºC per cycle; and 72 ºC for 1 min. This was followed by 20 cycles of 95 ºC for 30 s, 60 ºC (DLA-DRB1), 50 ºC (DLA-DQA1), 65 ºC (DLADQB1) for 1 min, and 72 ºC for 1 min. The protocol ended with a final extension of 72 ºC for 10 min. A generalised PCR touchdown protocol was used for any samples that amplified poorly at the DRB, DQB or DQA loci: 10 min at 95 ºC; 16 touchdown cycles of 94 ºC for 1 min; followed by 1 min annealing, starting at 63 ºC and reducing at 0.8 ºC per cycle; and 72 ºC for 1 min 30 s. This was followed by 30 cycles of 94 ºC for 30 s, 50 ºC and 72 ºC for 1 min. The protocol ended with a final extension of 72 ºC for 5 min. PCR products were visualised on 2% TBE agarose gels. As sequence-based typing relies on being able to read heterozygous sequences, it was vital that sequences had no background noise. Therefore, samples with a thin clear single band were cleaned using ExoSAP-IT (USB) according to the manufacturer's instructions. Otherwise, bands to be sequenced were excised and subsequently purified using QiaQuick gel extraction kits (Qiagen Inc.). Purified PCR products were directly sequenced on ABI 3730 sequencers (Genepool, University of Edinburgh; The Core, University of California, Los Angeles; The Sequencing Service, University of Dundee).

Sequence data were cleaned in Geneious v.4.5.5 (BioMatters) and then analyzed in Match Tools Navigator (Applied Biosystems), as described in Kennedy et al. (2002a). This method relies on an allele library built from homozygotes (or clones from heterozygotes) that is used to predict the most likely allelic combinations present in a heterozygous sequence (described in more detail in Chapter 2 and (Marsden et al. 2009). Since I was assessing MHC variation in previously uncharacterised species, I cloned 17 heterozygous individuals that did not match any pair of known alleles, using the TOPO TA cloning system and One Shot Competent cells (Invitrogen). At the DRB, I cloned one Dholes, three Golden jackals, one Side-striped jackal, and four Black-backed jackals, which identified 13 new alleles. At the DQB, I cloned five Black-backed jackals, which identified 7 new alleles. At the DQA, I cloned one Golden jackal and two Black-backed jackals, which identified three new alleles. Following standard protocol (Kennedy *pers.comm*), new alleles were given a temporary name corresponding to the name of the sample they were first recorded in. To maintain traceability of data, these temporary names are used here. However, sequences will be supplied to the DLA nomenclature

committee to be assigned official names, as well as deposited to GenBANK, prior to publication.

5.3.3 Diversity at the MHC

I calculated nucleotide diversity within species as the average number of segregating sites (θ) and pairwise diversity (π), using DnaSP 4.20 (Rozas and Rozas 1995). The average number of amino acid differences between alleles was calculated in MEGA 4.0 (Tamura et al. 2007), with standard errors calculated based on 5000 bootstrap replications. Observed heterozygosity was calculated as the percentage of heterozygous individuals (H_o). Potentially new alleles that were only found in heterozygous individuals and that could not be resolved were excluded from diversity estimates. However, since unresolved new alleles only occurred in the newly characterised species, this resulted in a bias towards lower H_o estimates in those species. As such, heterozygous individuals with unresolved new alleles were included in heterozygosity estimates. H_o estimates were not available from published data for Ethiopian wolves, and for Grey wolves, H_o estimates were only available from Seddon et al. (2004) (but not the other published data on wolves (Kennedy et al. 2007a; Kennedy et al. 2001; Seddon & Ellegren 2002).

5.3.4 Tests for Selection

A commonly used method to infer selection on a gene is to calculate the ratio (ω) of non-synonymous (d_N) : synonymous (d_S) variation, with positive selection indicated when $d_N : d_S$ is significantly greater than 1. I chose to assess $d_N : d_S$ as a method to test for selection at these MHC genes because, unlike standard neutrality tests such as Tajima's D (Hartl & Clark 2007), methods evaluating this ratio do not assume demographic stability and they are appropriate for species rather than population data. The McDonald Kreitman test, which compares levels of variation within a species to that between species, was not appropriate due to a lack of fixed differences between species expected with balancing-selection induced trans-specific polymorphism (Garrigan & Hedrick 2003; Hartl & Clark 2007).

I tested for positive selection at the DRB, DQB and DQA genes in each species using three methods based on $d_N : d_S$. I first used the PARRIS method (Scheffler et al. 2006), implemented in HYPHY (Pond et al. 2005) using the datamonkey webserver (www.datamonkey.org). PARRIS requires three or more alleles. Therefore, species that did not meet this requirement were excluded from this analyses. PARRIS is a global test based on estimating average $d_N : d_S$ ratios across a complete gene sequence to infer selection acting on the whole gene, whilst accounting for recombination and synonymous rate variation (Nozawa et al. 2009). PARRIS compares a model of no selection (M1) to a

model of selection (M2). In both models, d_N and d_S can vary, but in M1, d_N is restrained to be ≤ d_S (i.e. there is no evidence of positive selection) whereas in M2, d_N is able to be greater than d_S (i.e. indicating positive selection) (Michel et al. 2009). A likelihood ratio test is used to determine the model that best fits the data. This global test approach has the advantage of low rates of false positives and suitability for small data sets. However, it is also associated with low power because typically a small number of sites are under selection and therefore, by averaging across the sequence, signals of selection from a few sites are likely to be lost (Kosakovsky Pond & Frost 2005). The other disadvantage of this test is that it does not indicate where in the sequence selection is operating. At the MHC, selection is predicted to occur at codons involved in peptide binding. Therefore, I used a test implemented in MEGA 4.0 (Tamura et al. 2007) that partitions codons as being PBR or non-PBR, and tests whether d_N : d_S is significantly greater than 1 in the PBR and non-PBR separately. Following standard practice in mammalian studies, putative PBR sites were based on the human allele HLA-DRB1 (Bondinas et al. 2007). Synonymous and nonsynonymous genetic distances were calculated for putative PBR sites and non-PBR sites using the Nei–Gojobori method, with a Jukes–Cantor model of substitution, as implemented in MEGA 4.0 (Tamura et al. 2007). I tested for evidence of positive selection using a codon based Z test in MEGA. Monomorphic species were excluded from MEGA analyses. Since this is a global test, it suffers from a lack of power but, like PARRIS, has a low rate of false positives. Whilst MEGA can identify selection between two classes of codon (PBR or non-PBR) it cannot identify selection at individual codons. To this end, some methods have been developed to calculate d_N : d_S at individual codons, thus enabling one to identify specific sites under selection. I implemented one such method, fixed effects likelihood (FEL), in HYPHY using the datamonkey webserver. Due to the low power of this method with small data sets, FEL was only implemented on species with 5 or more alleles. FEL is a method that uses maximum likelihood to directly estimate d_N and d_S ratios at each codon independently (Kosakovsky Pond & Frost 2005). FEL does not assume a distribution of substitution rates for d_N and d_S, and takes error in estimation of d_N and d_S ratios into account. A likelihood ratio test (LRT) is then used to compare two models (neutral model where d_N=d_S and selection model where d_N and d_S are estimated separately) and thus determine if a particular site is under selection (Kosakovsky Pond & Frost 2005). FEL is a relatively conservative test, and has low rates of false positives. However, because sites are assessed individually, power to detect selection where there are less than 20 sequences is low (Kosakovsky Pond & Frost 2005).

For PARRIS and FEL, the most likely model of nucleotide substitution was selected based on results from Model Test, as implemented within HYPHY (Posada & Crandall 1998). These methods assume a phylogenetic tree and constant branch lengths across sites (Scheffler et al. 2006). However, recombination can result in branch lengths and tree topologies being different between sites and thus violates these critical

assumptions, which can cause false positives (Scheffler et al. 2006). Therefore, prior to any test for positive selection I used the GARD method (Kosakovsky Pond et al. 2006) implemented in HYPHY to test for the occurrence of recombination within sequence sets. When recombination is present, GARD identifies the break points and partitions the sequence accordingly, thus enabling each partition to have its own phylogenetic tree (Poon et al. 2009). GARD was run assuming a beta-gamma distribution of site to site variation, and four rates.

5.3.5 Trans-specific polymorphism

Due to the recombining nature of MHC genes, phylogenetic trees are not strictly appropriate for analysis of the MHC and there is too much variation to allow a network approach. In addition, due to the large number of alleles relative to the number of variable characters (DRB, 92:81; DQB, 50:67; DQA 32:21) trees were not expected to have a high degree of resolution. However, MHC allele trees are a useful tool for displaying relationships among alleles, particularly to evaluate patterns of trans-specific polymorphism expected under balancing selection. Phylogenetic trees for DRB, DQB and DQA were reconstructed in MrBayes v3.1.2 (Huelsenbeck & Ronquist 2001), using the best-fit nucleotide substitution model as indicated by Mr Model Test 2.2 (Hasegawa et al. 1985). The following human (HLA) sequences were used as outgroups: HLA-DQA*01, AY585236; HLA-DQB1*06, GQ422610; and HLA-DRB1*03011, AF352294. Four chains were run for 4,000,000 (DQA), 14,000,000 (DQB) or 100,000,000 (DRB) generations, with trees sampled every 100 generations. The burn-in was adjusted relating to the number of generations used for the different genes: 10,000 (DQA), 50,000 (DQB) and 200,000 (DRB). Trans-specific polymorphism was inferred where alleles were shared between species or where the closest relative of an allele in one species was an allele in a different species (based on phylogenetic trees).

5.3.6 Comparative methods

I used a generalized linear model (GLM) framework to test whether levels of MHC diversity and trans-specific polymorphism were significantly different between endangered and abundant species, and between species that can hybridise and those that cannot (detailed below). However, I also used a phylogenetic comparative method to control for phylogeny. It was predicted that if phylogeny does influence MHC diversity and trans-specific polymorphism, explanatory factors that were significant when not accounting for phylogeny, would become non significant when phylogeny was accounted for (Harvey & Pagel 1991). Specifically I used generalised estimating equations (GEE) which are an extension of standard GLM's that enable one to incorporate a phylogenetic tree as a correlation matrix, and thus account for phylogenetic relationships between species within

the model framework (Paradis 2006). A key advantage of GEE is that they are appropriate for categorical variables, unlike other phylogenetic comparative methods (e.g. phylogenetic independent contrasts, phylogenetic autoregression, autocorrelative methods, multivariate decomposition, generalised least squares), which require continuous traits (Paradis 2006). Furthermore, the results can be directly compared with a standard GLM as the model framework was the same except for the correction for phylogeny. The GLM and GEE analyses were both computed in the programme R (R core development team), with GEE implemented within the APE package (Paradis 2006). Population estimates are not known for many wolf-like species assessed here, therefore I classed species according to IUCN red list Endangered status (IUCN 2010). All of the species fell into just two categories; 'endangered' which are defined as species with a high risk of extinction (Ethiopian wolves, Dholes and wild dogs); or 'least concern', hereafter referred to as abundant, which are defined as species with the lowest risk of extinction, such as widespread and abundant taxa (Grey wolves, Coyotes, Golden jackals, Black-backed jackals). Grey wolves, Coyotes and Ethiopian wolves were classed as able to hybridise based on published data of hybridisation observed in nature (Gottelli et al. 1994; Lehman et al. 1991; Verginelli et al. 2005; Vilà et al. 1997). There is no documented evidence of Golden jackals, Dholes, wild dogs, Black-backed jackals or Side-Striped jackals hybridising with any other species in nature; therefore, they were classed as not able to hybridise. Phylogenetic comparative methods have been criticised for data trawling (Freckleton 2009). Therefore I *a priori* selected one trans-specific polymorphism metric and one diversity metric to evaluate. Following Lu (2001), trans-species polymorphism was quantified as the number of alleles out of the total in a species for which the closest relative is an allele found in another species (based on phylogenetic trees) rather than another allele from the same species, in addition to shared alleles. To reflect amino acid differences, diversity was calculated as d_N, which was computed in MEGA 4.0 (Tamura et al. 2007). The wolf-like canid phylogeny was taken from Lindblad-Toh (2005) which was based on ~15kbp of intron and exon sequence (Lindblad-Toh et al. 2005); branch lengths for this tree were provided by K-P Koepfli. For the GLM, error was modelled using a binomial distribution for the trans-specific polymorphism metric because the data was proportional, and using a normal distribution for d_N diversity metric. The explanatory variables were catergorical with two levels.

5.4 Results

5.4.1 Diversity at the MHC

Across the wolf like canids, a total of 92 DRB (n=1870 individuals typed), 50 DQB (n=398) and 31 DQA (n=271) alleles were recorded, including 27 new DRB alleles, 17 new DQB alleles and 10 new DQA alleles. DRB was found to be the most diverse locus both in terms of numbers of alleles, nucleotide and amino acid diversity, whilst DQA was the least diverse (Table 5.1; Table 5.6; Figure 5.2). There were some heterozygous individuals with new alleles which were unresolved in newly characterised species, and these were excluded from the study: Black-backed jackals (7 at DRB, 3 DQB, 1 DQA), Side-striped jackals (2 DRB); Golden jackal (4 DRB, 1 DQB, 1 DQA), Coyote (9 DRB, 5 DQB, 6 DQA).

Patterns of MHC diversity in the Coyote were complex and unusual for the wolf-like canids. Specifically, diversity at DQB and DRB was exceptionally high (17 alleles from 19 samples at the DRB, and 11 alleles from 15 individuals at the DQB). Furthermore, heterozygosity patterns across loci were atypical for the wolf-like canids and may indicate null alleles. For example, some Coyotes were heterozygous at the DRB and DQA, but not the DQB, or at the DQB, DQA and not the DRB. As such, estimates of the numbers of alleles at DRB and DQB in the Coyote should be viewed as underestimates, but I am confident in the alleles that I did detect and thus these data are useful for allele sharing comparisons. However, it is noteworthy a more extensive assessment study of Coyotes is currently being conducted to characterise MHC variation in this species (L.Kennedy *Pers.comm*).

I found large differences in the number of alleles per locus in different species (DRB, 4-27; DQB, 1-21; DQA, 1-16; Table 5.1). It was not possible to apply rarefaction to this sample set, as it was not only the number of samples that varied between species; the number of populations varied also, and for many samples population was unknown. As a result, sample sizes must be considered when comparing diversity between species. Nonetheless, it can be seen that differences between species cannot be explained by sampling intensity alone. For example, African wild dogs (n=271) and Dholes (n=11) had no variation at the DQA locus, and just two and one alleles respectively at the DQB locus, whereas fewer samples from Side-striped jackals (n=6-8) had 2-5 alleles at DQA, and 3 alleles at DQB. Further samples of Dholes are required to confirm the absence of polymorphism at the DQA and DQB. There was also considerable discrepancy in diversity between species in terms of sequence diversity of alleles (π, θ, amino acid diversity, number of variable PBR; Table 5.6). Sequence diversity within species was often

unrelated to the number of alleles. For example, at the DRB, sequence diversity metrics were similar between Grey wolves, Coyotes, Golden jackals and Ethiopian wolves, despite the number of alleles ranging from 4 – 27. However, this was not always the case; at the DRB, Dholes and Side-striped jackals (n=4 alleles) had considerably lower sequence diversity than Ethiopian wolves which also had just four alleles (Table 5.6). In contrast to the DRB and DQB, sequence diversity was similar between most species at the DQA (excluding species with no variation, Dholes and wild dog). Levels of heterozygosity within species were generally high (>50%) across all three loci, with the exception of the DRB locus in Dholes, DQB locus in wild dogs, and all three loci in Side-Striped jackals (Table 5.6).

Overall, Grey wolves and Coyotes consistently showed the highest diversity in terms of both numbers of alleles per locus and highest sequence diversity of alleles across all three loci, whereas Dholes, wild dogs, Side-striped and Ethiopian wolves showed the least diversity for at least one locus (Table 5.1; Table 5.6; Figure 5.2).

5.4.2 Tests for selection

There was a high level of non-synonymous variation amongst DRB, DQB and DQA alleles. Within species, all alleles differed non-synonymously except for two DRB alleles in Side-striped jackals (CAD6 & CAD19), two DRB alleles in African wild dogs (T5920 and lk5237v), two DQA alleles in Side-striped jackals (L051 and CAD2), and two DQA alleles in Grey wolves (DQA*014011 & *014012). A global selection test, implemented in PARRIS, found significant evidence that $d_N/d_S>1$ amongst DQA alleles of Grey wolves and Coyotes, and DQB and DRB alleles of Grey wolves, Coyotes and Black-backed jackals. Across species, the number of PBR codons that were variable ranged from 7-16/19 at the DRB, 0-15/19 at the DQB, and 0-6/15 at the DQA (Table 5.6). In general, though, Grey wolves and Coyotes typically exhibited the largest number of variable PBR codons per locus (Table 5.1; Table 5.6; Figure 5.2). At the DQA, positive selection was indicated in the PBR (ω significantly > 1) in all species, except the Ethiopian wolf (and wild dog and Dholes, which had just one allele each), whereas there was no evidence of selection at non-PBR (Table 5.5). The same pattern was found at the DQB and DRB, although in fewer species and in Ethiopian wolves (DRB) and Black-backed jackals (DQB), positive selection was indicated in both PBR and non-PBR data partitions (Table 5.3; Table 5.4).

Amino-acid variability profiles showed that variation was highly correlated with PBR (Figure 5.2). Codon-based selection tests based on FEL identified a small number of codons at each locus showing evidence of positive selection in some species (Table 5.3; Table 5.4; Table 5.5). In all cases, the codons under selection were putative PBR, or

codons directly adjacent to putative PBR. Although positive selection was only detected in some species, the specific codons showing evidence of selection were often the same in different species, even where species shared no alleles. For example, for DRB, codon 11 was found to be under positive selection in both Black-backed jackals, and Grey wolves despite the fact they shared no alleles (Table 5.3; Table 5.4; Table 5.5). Overall, there was a striking similarity in which codons were variable in different species (Figure 5.2), which likely reflects the combination of trans-specific polymorphism and common disease pressures.

5.4.3 Trans-specific polymorphism

All DRB, DQB and DQA alleles detected in African wild dogs were unique to wild dogs. The same was true for the Dholes. By contrast, allele sharing (trans-specific polymorphism) was detected in all other species (Figure 5.4). The number of alleles shared between divergent taxa was striking (Table 5.2); e.g., 33% of DQA alleles (3/9) were shared between Black-backed jackals and Grey wolves, which diverged ~ 4.3 Myrs ago (Lindblad-Toh et al. 2005). Overall, allele sharing was most common at the DQA locus (17/32, 50%, alleles shared) but also frequently observed at the DQB locus (10/50, 20%). At these two loci, alleles were shared between the most distantly related species (e.g. Black-backed jackals and Grey wolves; Side-striped jackals and Coyotes) and in some cases alleles were shared across 4/8 wolf-like canid species: DQA*01101; DQA*L051; DQB*01303; DQB*02303 (Figure 5.5; Figure 5.6). By contrast, at the DRB locus, allele sharing was much rarer (7/92, 8% of alleles), and only found between the most closely related species; Coyotes and Grey wolves (Figure 5.4). In addition to allele sharing, trans-specific polymorphism was also indicated by the retention of allelic lineages, as shown by alleles between species being more similar than alleles within species. This can be demonstrated using a phylogenetic tree when alleles within species show a scattered distribution throughout the tree rather than species-specific allelic clustering. There was some evidence of this pattern in most species, although species-specific allele clustering was clearly evident in both the wild dog and Black-backed jackals at the DRB (Figure 5.4; Figure 5.5; Figure 5.6). A scattered distribution of alleles was more common at the DQB and DQA across species, except for wild dogs, Dholes and Ethiopian wolves at DQA and wild dogs and Dholes at DQB. Overall, trans-specific polymorphism was found to be a common feature at DQA and DQB loci in all of the wolf-like canids, except the endangered Dholes and wild dogs.

5.4.4 Comparative methods

Endangered species were found to have consistently lower average levels of MHC diversity (d_N) and trans-specific polymorphism than abundant species at all three loci

(Figure 5.3). On the other hand, species that are able to hybridise had consistently higher average MHC diversity and trans-specific polymorphism than species that are not able to hybridise, at all three loci (Figure 5.3). However, in many cases there was considerable variation within, and overlap between, the different categories.

The GLM analyses showed that levels of diversity were significantly higher in abundant than endangered taxa at the DQA (GLM, $p<0.05$) and DQB loci (GLM, $p<0.05$), but this was not significant after phylogenetic correction (GEE, DQA, $p=0.066$; DQB $p=0.064$). Levels of trans-specific polymorphism were also significantly higher in abundant than endangered taxa at the DQA locus (GLM, $p<0.05$), but again this was not significant after phylogenetic correction (GEE, 0.130). These results suggest that evolutionary history does influence both levels of trans-specific polymorphism at the DQA, and MHC diversity at the DQA and DQB. Specifically, since endangered species were phylogenetically clustered, interpretation of whether risk of extinction is due to their shared ancestor having lower diversity or whether bottlenecks in population size in each species caused reduction in diversity at the MHC is confounded (Figure 5.1). Moreover, the ladder-like structure of the current canid phylogeny (i.e. the three endangered species are basal to the more abundant species, except for the side-striped and black-backed jackals), makes it more difficult to interpret results in terms of ancestral states.

There was no significant difference in levels of diversity between endangered and abundant taxa at the DRB (GLM, $p=0.393$, GEE, $p=0.463$) or in levels of trans-specific polymorphism between endangered and abundant taxa at the DQB and DRB (GLM, DQB, $p=0.721$, DRB, $p= 0.980$; GEE, DQB, $p=0.2547$, DRB, $p=0.519$). Similarly, levels of MHC diversity and trans-specific polymorphism were not significantly different between species that can and cannot hybridise at any locus (Figure 5.3). In these cases, it was therefore not possible to assess the influence of phylogenetic correction.

5.5 Discussion

Balancing selection on the MHC is predicted to result in a large number of medium frequency alleles, which increases heterozygosity, and reduces the likelihood that alleles will become fixed (Garrigan & Hedrick 2003; Muirhead 2001). Furthermore, balancing selection is predicted to retain ancestral polymorphism, resulting in high sequence divergence amongst alleles within a species. However, like other components of genetic variation, the numbers of alleles and retention of ancestral polymorphism at the MHC is expected to be very sensitive to effective population size (N_e) (Richman 2000) due to selection being weaker, and strength of genetic drift and loss of alleles being greater, when N_e is small (e.g. Sommer 2005). As such, the numbers of alleles and extent of trans-specific polymorphism is expected to be lower in endangered species (which have small N_e) compared to abundant species (Richman 2000). However, differences in the numbers of alleles and extent of trans-specific polymorphism between species may also be explained at least in part by phylogenetic distance between the species being compared. As such, using species comparisons to infer whether endangered species have lost diversity may be confounded by evolutionary history. In this study, I assessed MHC variation across species of the wolf-like canid clade (Lindblad-Toh et al. 2005). These data were used to assess how endangered status, selection, hybridisation and phylogeny impact patterns of MHC diversity.

5.5.1 Do endangered species show different patterns of MHC diversity to abundant species?

There was a large discrepancy in MHC diversity (# alleles, π, θ, PBR, amino acid) between the wolf-like canid species at all three loci (Table 5.6). Low sampling intensity may explain the small numbers of alleles found in some species (e.g. Side-striped jackal). However, sampling intensity alone cannot explain all differences between species. For example, despite being one of the best sampled species, wild dogs had the lowest DQA and DQB diversity (except Dholes, n=11). Similarly, the discrepancy in sequence diversity between species cannot be explained by just differences in numbers of alleles. For example, Golden Jackals and Ethiopian wolves had similar numbers of DQB alleles (6 and 5 respectively), but sequence diversity was much greater in Golden jackals (0.047) than Ethiopian wolves (0.019; Table 5.6). In general, endangered status was well correlated with MHC diversity; endangered species had consistently lower average levels of MHC diversity than abundant species at all three loci, although the differences were only significant at the DQB and DQA loci (Figure 5.3). However, diversity across the endangered species was quite uneven as indicated by variance shown in Figure 5.3. For example, sequence diversity amongst the four Ethiopian wolf DRB alleles (π=0.073) was

far greater than amongst the four Dhole DRB alleles (π=0.038), and comparable to abundant taxa with between 13 and 27 DRB alleles (π=0.071-0.075; Table 5.6). This suggests that a species may lose alleles but retain high sequence diversity if alleles from different lineages are retained by selection, as suggested by Hedrick (2002). In contrast to Ethiopian wolves, sequence diversity amongst wild dog DRB alleles was low, despite the presence of 21 alleles. This likely reflects that all wild dog alleles are derived from just two allelic lineages, with little variation within lineages. However, the low number of lineages alone cannot account for low sequence diversity in wild dogs, as Black-backed jackals also had just a small number of DRB lineages, but the highest DRB sequence diversity overall (Table 5.6; Figure 5.4). A very similar pattern to that seen at the DRB in wild dogs has been detected at the S locus (a plant self/non-self recognition system similar to the MHC and under balancing selection) of *Physalis crassifolia,* where there has been: 1) a large reduction in diversity compared to related species, resulting in the loss of most allelic lineages; 2) extensive 'allelic re-diversification' of the remaining allelic lineages, resulting in a large number of similar alleles (Richman 2000). As such, it appears that there has been a loss of DRB polymorphism during a bottleneck in wild dogs and subsequent re-generation of allelic diversity post bottleneck.

Overall, endangered species were found to have consistently lower average levels of MHC diversity than abundant species. However, the large variation in diversity amongst endangered species and lack of significance at the DRB, indicates that the relationship between endangered status and diversity is complex and that factors aside from endangered status influence differences in diversity between species.

5.5.2 Is there evidence for differences in selection pressures on endangered versus abundant species diversity?

Selection is thought to be the driving force behind diversity at the MHC, and therefore may account for some of the differences in diversity between species. I found evidence of selection at the DRB, DQB and DQA loci, but not in all species. Overall, Grey wolves, Coyotes and Black-backed jackals most consistently showed evidence of selection, but these species also had the largest numbers of alleles and therefore the highest power in tests. One of the strongest indicators of balancing selection is evidence for trans-specific polymorphism. In the wolf-like canids, up to 50% of alleles at a locus were alleles shared with at least one other species, with allele sharing most common at the DQA and DQB loci. I found shared alleles in all species except Dholes and wild dogs and between the most basal and recently diverged species (Table 5.2). In addition to allele sharing, trans-specific polymorphism was also evidenced by alleles showing a scattered rather than clumped distribution in phylogenetic trees in most species (Figure 5.4; Figure 5.5; Figure 5.6). The strong evidence of trans-specific polymorphism indicates

the influence of balancing selection despite the lack of statistical evidence for selection for many species.

Overall, all of the abundant taxa showed evidence of trans-specific polymorphism and significant results in selection tests for at least one MHC locus (Table 5.3; Table 5.4; Table 5.5; Figure 5.4). Amongst the endangered canids, such strong evidence of selection was only found in Ethiopian wolves. There was no evidence for selection or trans-specific polymorphism at any locus in wild dogs or Dholes. In these two endangered taxa, the loss of alleles appears to have resulted in loci becoming fixed, or, with so few alleles that there was insufficient power to detect selection. The exception to this is the DRB locus in wild dogs, which showed no evidence of selection despite being highly polymorphic (21 alleles). Given that 20/21 alleles differed non-synonymously, and most variation was restricted to the PBR, it is likely that the lack of significance reflects either low selection intensities or recent re-diversification at the wild dog DRB (which results in few differences between alleles), rather than an absence of selection *per se* (Garrigan & Hedrick 2003). At the DQB, one allele was detected at very high frequency in wild dogs (91%) and Dholes were fixed for a single allele. At the DRB, which has predominantly mid-frequency alleles in the other species, Dholes also showed one high frequency allele (77%). These patterns are not expected under balancing selection, but could indicate either directional selection, random drift or be a consequence of linkage.

These data suggest that lower selection pressures may explain some of the variation in diversity between endangered and abundant canids. However, variation in selection likely also explains differences amongst the endangered species. In comparison to Dholes and wild dogs, levels of DRB diversity in Ethiopian wolves were considerably higher and comparable to abundant taxa. This was unexpected for an endangered species with a global population of less than 500 individuals, sampled from one small and isolated population (Bale Mountains, n~250) (Gottelli et al. 2004). However, the Bale Mountains Ethiopian wolf population has suffered a number of severe disease epidemics (~75% mortality, (Randall et al. 2006). The results here suggest that these strong selective pressures have maintained DRB sequence diversity in Ethiopian wolves despite strong genetic drift, as suggested in the case of MHC diversity being found in the otherwise genetically monomorphic San Nicolas Island fox (*Urocyon littoralis dickeyi*) (Aguilar et al. 2004). Fixation of alleles at the DQA and DQB loci has been commonly documented in endangered species and/or small isolated populations (e.g. *Urocyon littoralis*, for a review of examples see (Munguia-Vega et al. 2007). This is thought to reflect that the lower selection intensities at the DQB and DQA relative to the DRB (Satta et al. 1994) have been less able to retain diversity (Munguia-Vega et al. 2007), as indicated by the low diversity in Dholes and wild dogs. As such, the numbers of DQB and DQA alleles found in Ethiopian wolves seems surprising (Table 5.1). However, the canid

MHC class II loci are tightly linked and inherited as a haplotype (Kennedy et al. 2007a). As such, strong selection at the DRB in Ethiopian wolves may have resulted in hitch-hiking at the DQB and DQA loci. In contrast, wild dogs appear to have undergone a dramatic bottleneck which reduced the species to just two DRB alleles, representing each of the two lineages, and so the number of linked DQB and DQA alleles was small (as discussed above). It is noteworthy that due to tight-linkage at the MHC, some studies have assessed diversity at the DRB under the assumption that this should be a good indicator of variation at the other MHC class II loci due to tight linkage (Gutierrez-Espeleta et al. 2001). However, as the examples from wild dogs and Dhole show, this may not always be the case in endangered species.

5.5.3 Does hybridisation affect interpretations of selection and diversity?

Evidence of trans-specific polymorphism is generally interpreted as evidence that alleles have been retained by balancing selection (Klein et al. 2007). However, trans-specific polymorphism could also be the result of convergent evolution or hybridisation. Convergent evolution is conceivable given that: 1) canids share susceptibility to a number of diseases (Hofmeyr et al. 2000); 2) DRB amino-acid profiles were strikingly similar between species sharing no alleles (Figure 5.2); 3) codons showing positive selection were identical in species sharing no alleles (Table 5.3). To assess this, I compared whether phylogenetic trees based on just non-PBR had the same topology as those based on the whole sequence for the DRB (NJ trees; data not shown). In agreement with balancing selection, both sets of trees resolved the same relationships. Yeager (1999) suggested that convergent evolution is indicated where the same amino acid at a residue is encoded by different codons. Grey wolves and Black-backed jackals share no DRB alleles, but there were 127 instances where Grey wolves and Black-backed jackals shared identical amino acid residues. Contrary to the expectation for convergent evolution, all 127 of these amino acids were encoded by the identical codon (data not shown). Together, these data indicate that sequence similarity is due to ancestry rather than convergent evolution.

Overall, there was a trend for higher levels of diversity and trans-specific polymorphism in species that can hybridise, but this was not significant at any locus. Direct sharing of DQB and DQA alleles between species that do not hybridise provides clear evidence of trans-specific polymorphism. At the DRB locus, it is possible that hybridisation has increased trans-specific polymorphism, as alleles were only shared between Grey wolves and Coyotes, which can hybridise. However, Grey wolves and Coyotes are also closely related (divergence ~ 1 million years ago (Anderson et al. 2009), and as such, allele sharing is also predicted to be more common between them, which would also increase diversity. As such, the data on whether hybridisation has contributed

to diversity and trans-specific polymorphism is inconclusive. Further research on this topic is required.

5.5.4 Are patterns of selection and diversity confounded by phylogenetic history?

This study has shown that patterns of MHC diversity vary greatly between the wolf-like canids. These differences were expected to be largely the result of differences in demographic history, that is, endangered status. However, although levels of diversity and trans-specific polymorphism differed between endangered and abundant taxa, there was much variation within these categories, indicating that other factors must be involved.

The results here suggest that some of this variation may be explained by evolutionary history, as comparative phylogenetic analyses showed that phylogenetic relatedness influences both diversity and trans-specific polymorphism at some MHC loci (Figure 5.3). It may not be surprising therefore, that Dholes and wild dogs, which are both basal to the more derived widespread species in the genus *Canis* (and thus potentially older) and the only extant representatives of their respective genera, should show different patterns of MHC diversity. Consequently, this means that studies must distinguish whether differences in patterns of MHC diversity in an endangered species are the result of ancestral losses in diversity rather than bottlenecks due to their endangered status, that is, distinguish between evolutionary and demographic history. The high species coverage of the wolf-like canids made it possible to investigate this. For example, sharing of DRB alleles and lineages was common to the closely related Coyotes and Grey wolves. By contrast, wild dogs and Black-backed jackals, which are basal to Coyotes and Grey wolves, both showed species-specific clustering of alleles into a small number of lineages and an absence of allele sharing, indicating that this difference could be due to evolutionary history. The DQA and DQB loci were found to be moderately polymorphic in all species except two of the endangered species (Dholes and wild dogs). The most phylogenetically parsimonious explanation of this pattern is that DQA and DQB diversity was lost in the common ancestor to wild dogs and Dholes, and then regained afterwards in the widespread species (see Figure 5.1). However, at DQA and DQB alleles were shared across Coyotes, Grey wolves and Black-backed or Side-striped jackals, which are phylogenetically separated by wild dogs and Dholes. This suggests instead that DQB and DQA variation was lost independently in both wild dogs and Dholes. The ladder-like structure of the current phylogeny likely explains the discrepancy between the statistical results and the most realistic biological explanation. Although the current molecular-based phylogeny is based on a large number of characters (~15kb, (Lindblad-Toh et al. 2005), it is possible that extinction of sibling species to the monotypic endangered species (e.g Lycaon in, Hartstone-Rose et al. 2010) has resulted in the ladder like structure and lack of

resolution, which would reduce the power of comparative analyses. As such, the lack of diversity in this instance appears to be the result of demographic reductions associated with their endangered status rather than evolutionary history.

Consideration of evolutionary history is also important if domesticated species are used in comparisons with endangered taxa, because the domestication process may have resulted in atypical patterns of diversity. For the wolf-like canids, domestic dogs which originated from middle eastern Grey wolves (vonHoldt et al. 2010) and have been extensively studied (e.g. (Kennedy et al. 2002b), and could have been used as reference taxa. However, domestic dogs show exceptionally high MHC diversity: 90 DRB alleles, 54 DQB alleles and 22 DQA alleles (L.Kennedy *Pers comm*). As such, all of the wolf-like canids appear depauperate of diversity in comparison.

Comparisons of MHC diversity between highly divergent taxa are common in the MHC literature (e.g. marsupials and placental mammals), but evolutionary history likely confounds these results. Direct allele sharing has only been observed in primates (Otting et al. 2002; Suárez et al. 2006), ungulates (Radwan et al. 2007), rodents (Cutrera & Lacey 2007) and Cetaceans (Xu et al. 2009), and rarely outside of genera (excluding DRA locus). In contrast, I found that allele sharing was very common in the wolf-like canids, even between genera. This might be interpreted as evidence that this is the result of unusually strong high balancing selection. However consideration of evolutionary history suggests it is most likely that it reflects the recent and rapid divergence of the wolf-like canids (Lindblad-Toh et al. 2005). This suggests comparisons of patterns of MHC diversity between highly divergent groups should be conducted with extreme caution.

5.6 Conclusion

Assessments of MHC diversity in endangered species are increasingly common due to concerns that 'low' MHC diversity may reduce immune competence. However, the results presented here for the wolf-like canids suggest that the relationship between MHC diversity and endangered status is complex. Patterns of MHC diversity in endangered species were shown not only to be affected by their demographic history, but also by selection and evolutionary history. The importance of evolutionary history as a factor influencing MHC diversity is intuitive but has rarely been considered when interpreting differences in MHC diversity between species, even if they are distantly related. Phylogenetic comparative methods are widely and routinely implemented in other fields to control for the influence of common ancestry on trait similarity between species. However, they are not typically employed in comparative analyses of diversity between endangered and non-endangered taxa. Unfortunately, the power of this method was reduced with this dataset because of the small sample size (8 species), ladder-like structure of the wolf-like canid phylogeny and phylogenetic clustering of both endangered status and hybridisation ability. Nonetheless, I found evidence that phylogeny does appear to influence levels of MHC diversity and trans-specific polymorphism. This has important implications for the choice of reference taxa used in comparative studies aiming to classify levels of diversity in (endangered) species. Clearly, unless very closely related species are compared, it will be difficult to disentangle whether any differences in diversity between species are the result of population declines rather than evolutionary history. As such, comparisons should be conducted with sister species wherever possible (e.g. Miller & Lambert 2004). If the reference taxa/on is/are a distant relative, it is vital that intermediate species are assessed, as was done here.

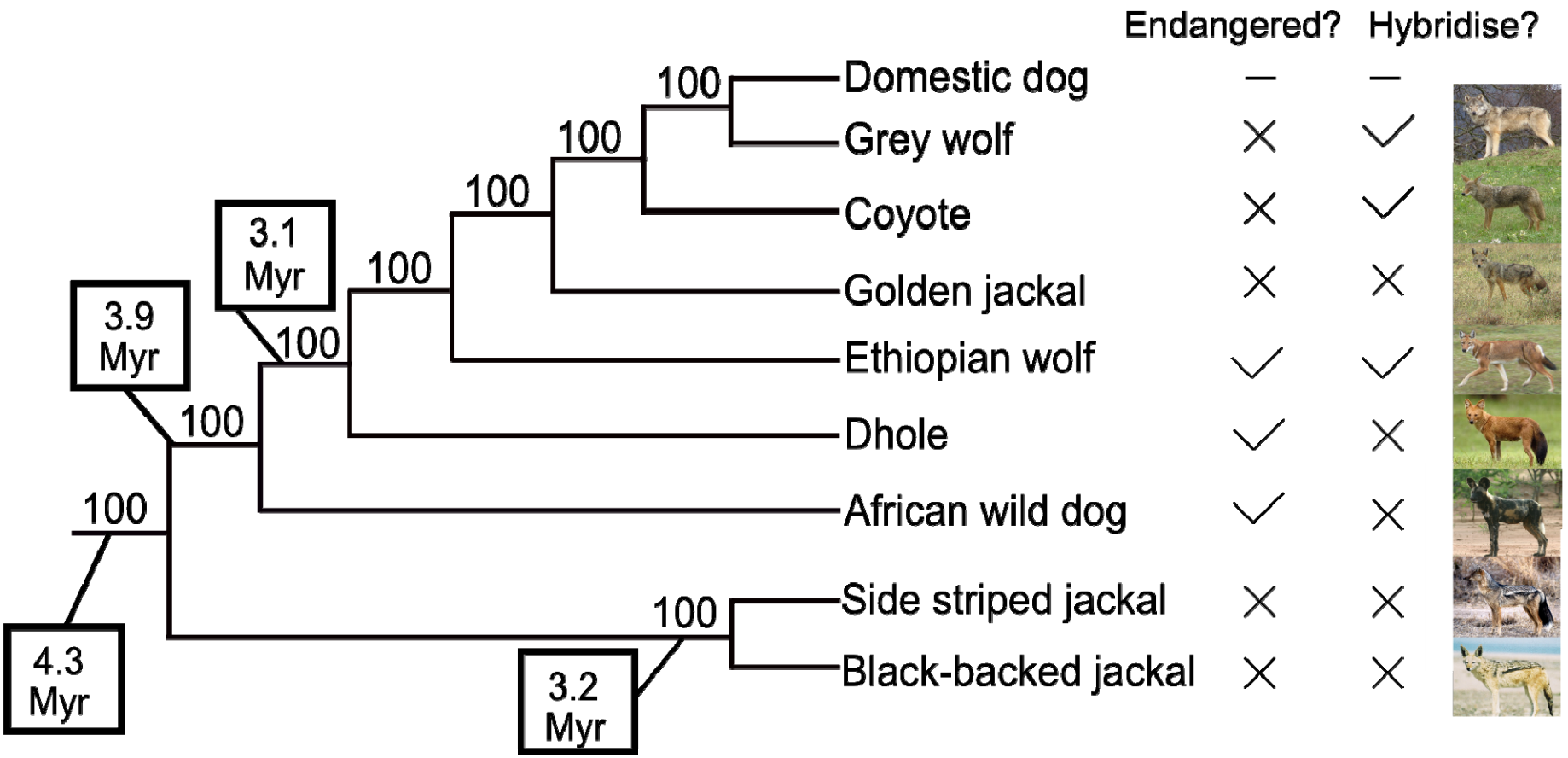


Figure 5.1: Phylogenetic tree of the wolf-like canids based on 15kb of intron and exon sequence, with Bayesian posterior probabilties. Re-drawn from Lindblah-Toh (2005) with divergence dates from Perini (2009). Also indicated is whether the species is endangered and whether it is able to hybridise. Domestic dogs, which were domesticated from Grey wolves, were not included in this study. Photographs correspond to the species in the tree. Credits: Grey wolf, Wild dog, Black-backed jackal - C.Marsden; Ethiopian wolf – F.Vial; Coyote, Dhole, Golden Jackal, Side striped jackal from http://commons.wikimedia.org/File:[Cuon.alpinus-cut.jpg:Coyote2008.jpg:Side-striped_Jackal.jpg]

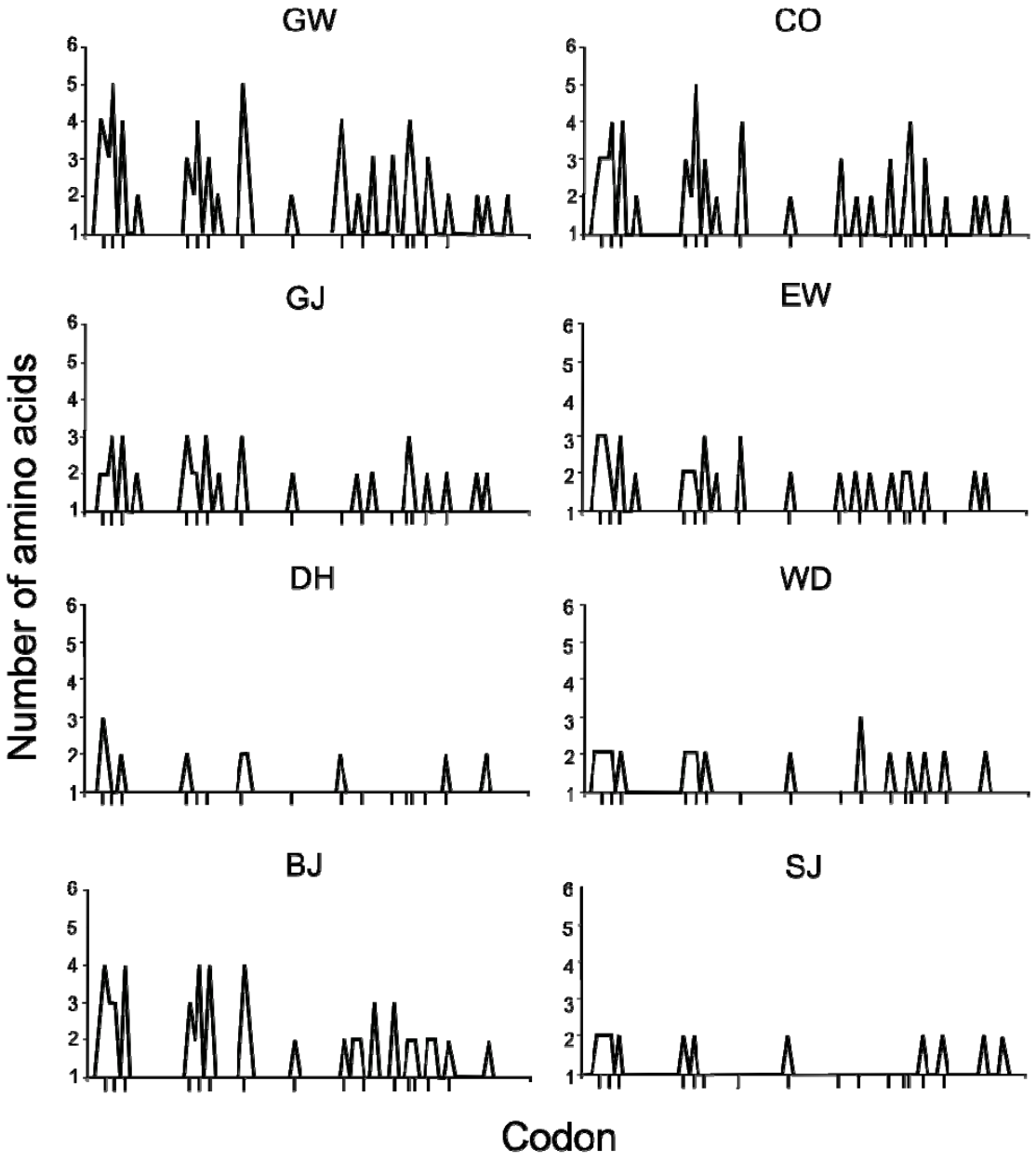


Figure 5.2: Amino acid variability at codons of the DRB across eight wolf-like canid species. The similarity in profiles between species is remarkable given that alleles are only shared between GW and CO. It can be seen that variability is concentrated in sites involved in PBR, indicated by tick marks on the X axis. PBR positions from Bondinas (2007). Species codes used are from Table 5.1.

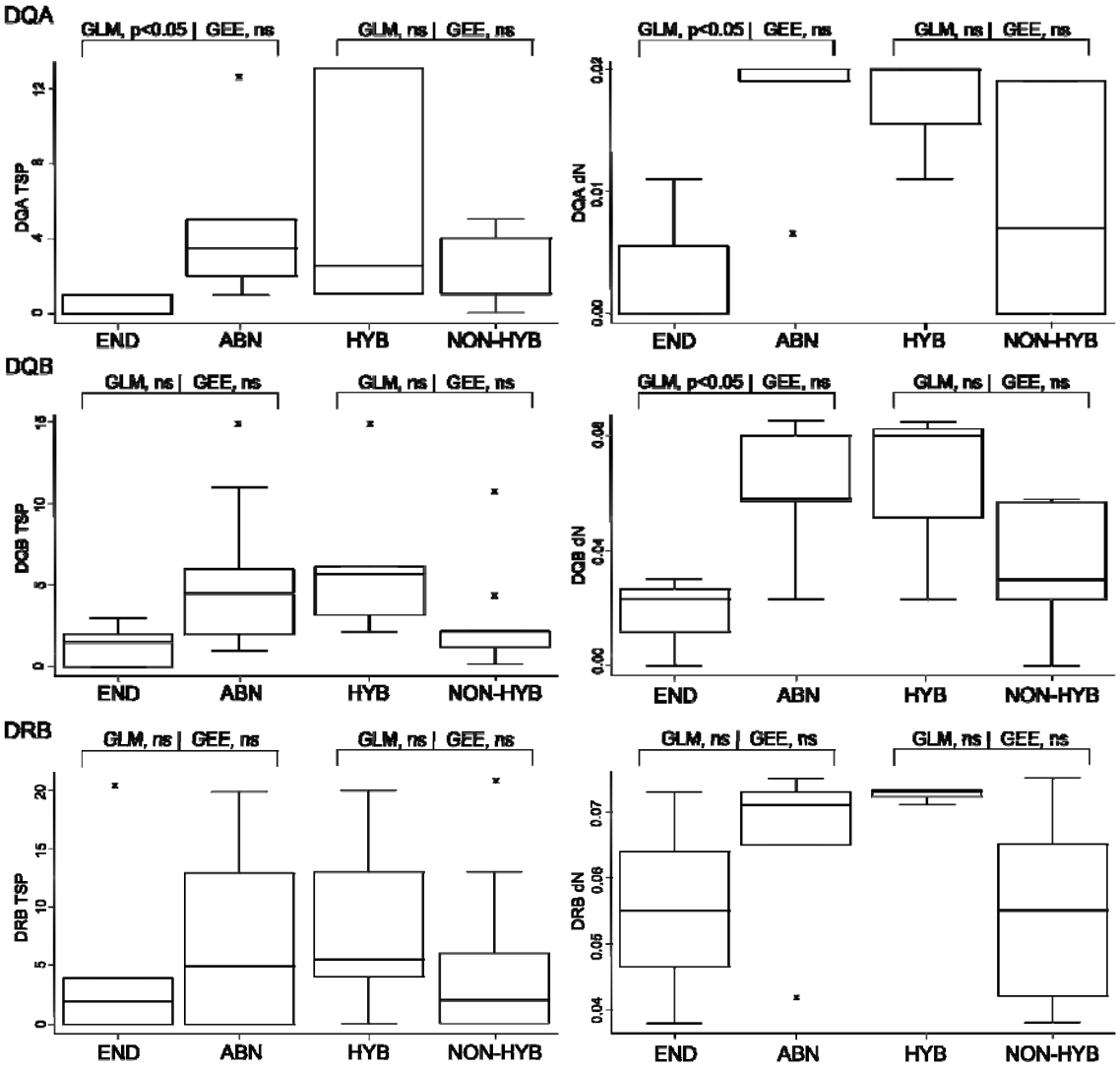


Figure 5.3: Levels of MHC trans-specific polymorphism (TSP), and MHC diversity (dN) for endangered (END, n=3) and abundant (ABN, n=5) wolf-like canids and hybridising (HYB) and non-hybridising (NON-HYB, n=) wolf-like canids. Box plots show 25%, 50% and 75% quartiles. Significant difference in levels of trans-specific polymorphism (or MHC diversity) between END and ABN and HYB and NON-HYB are indicated above the box plots. Results are based on a standard GLM, as well as using GEE, where phylogeny was corrected for.

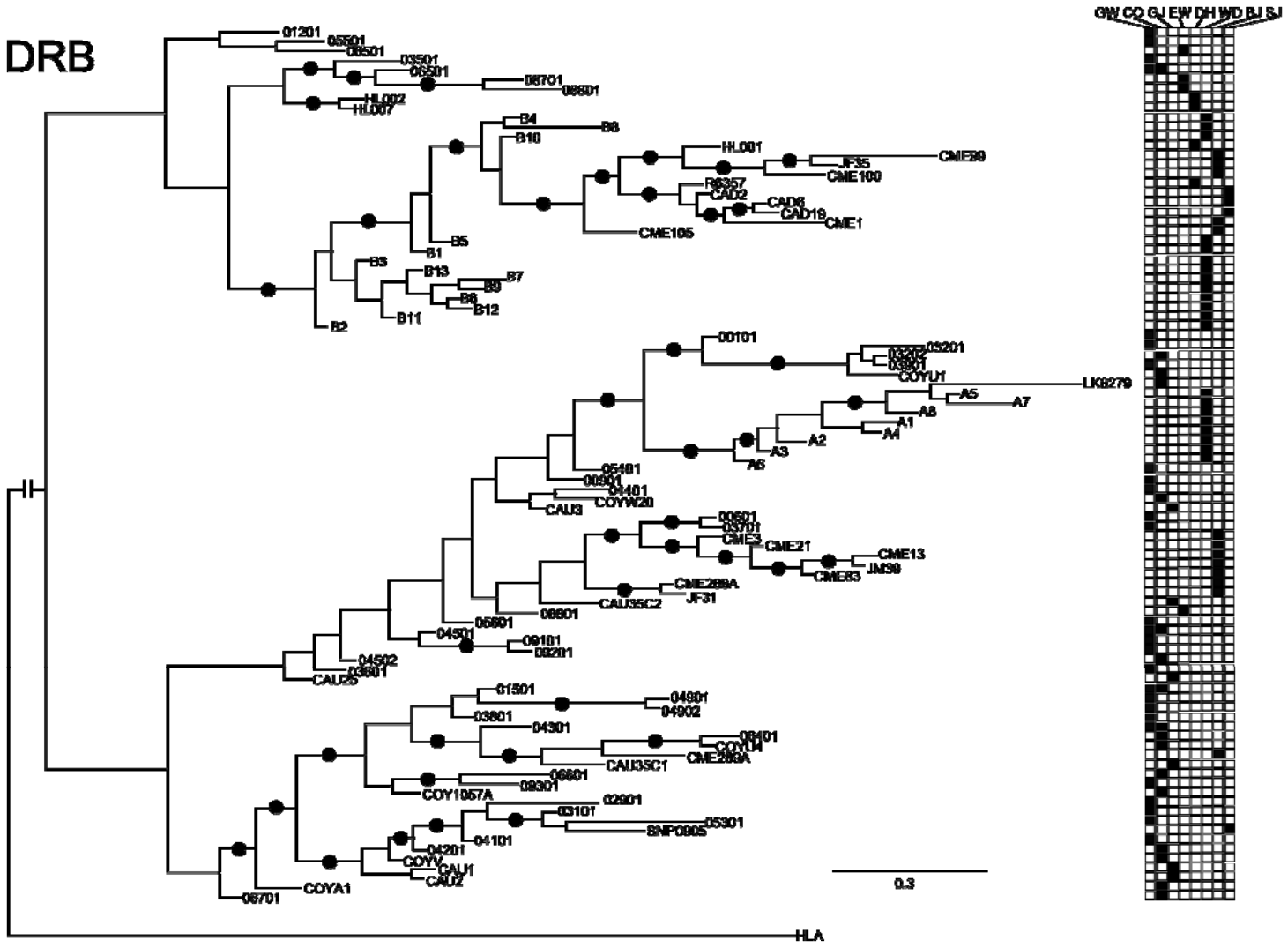

DRB
01201
05501
06501
03501
06501
06701
08801
HL002
HL007
B4
B8
B10
HL001
JF35
CME100
CME89
R6357
CAD2
CAD6
CAD19
CME1
CME105
B5
B1
B3
B13
B7
B9
B6
B12
B11
B2
00101
05201
COYU1
LK8279
A5
A7
A8
A1
A4
A2
A3
A6
05401
00901
04401
COYW20
CAU3
00501
05701
CME3
CME21
CME13
CME83
JM98
CME288A
JF31
CAU35C2
08801
06601
04501
09101
05201
04502
03601
CAU25
01501
04901
04902
03601
04301
06101
COYU4
CME288A
CAU35C1
06601
09301
COY1057A
02901
05107
05301
SNP0905
04101
04201
COYV
CAU1
CAU2
COYA1
06701
0.3
HLA
GN CO GJ EW DH WD BJ SJ

Figure 5.4: Mr Bayes tree of wolf-like canid DRB alleles, rooted with an HLA sequence. Presence of alleles in different species (codes from Table 5.1) are shown by filled boxes, which shows allele sharing between different species and the clustering or scattered distribution of alleles in different species. Solid circles depict branches with posterior probabilities ≥70%.

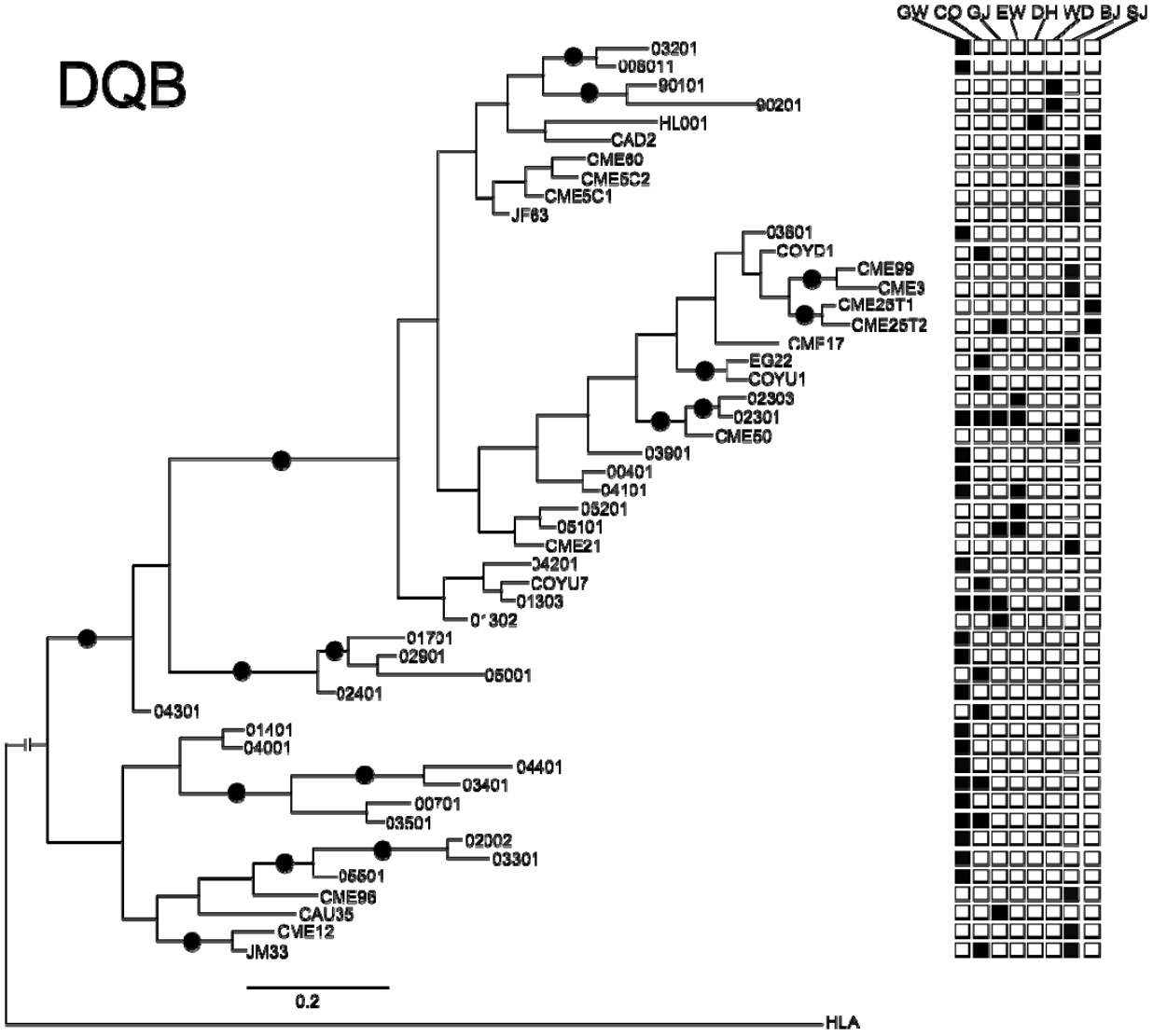

DQB
GW CO GJ EW DH WD BJ SJ
03201
008011
90101
90201
HL001
CAD2
CME60
CME5C2
CME5C1
JF63
03801
COYD1
CME99
CME3
CME25T1
CME25T2
CMF17
EG22
COYU1
02303
02301
CME50
03901
00401
04101
05201
05101
CME21
04201
COYU7
01303
01302
01701
02901
05001
02401
04301
01401
04001
04401
03401
00701
03501
02002
03301
05501
CME96
CAU35
CME12
JM33
0.2
HLA

Figure 5.5: Mr Bayes tree of wolf-like canid DQB alleles, rooted with an HLA sequence. Presence of alleles in different species (codes from Table 5.1) are shown by filled boxes, which show allele sharing between different species and the clustering or scattered distribution of alleles in different species. Solid circles depict branches with posterior probabilities ≥70%.

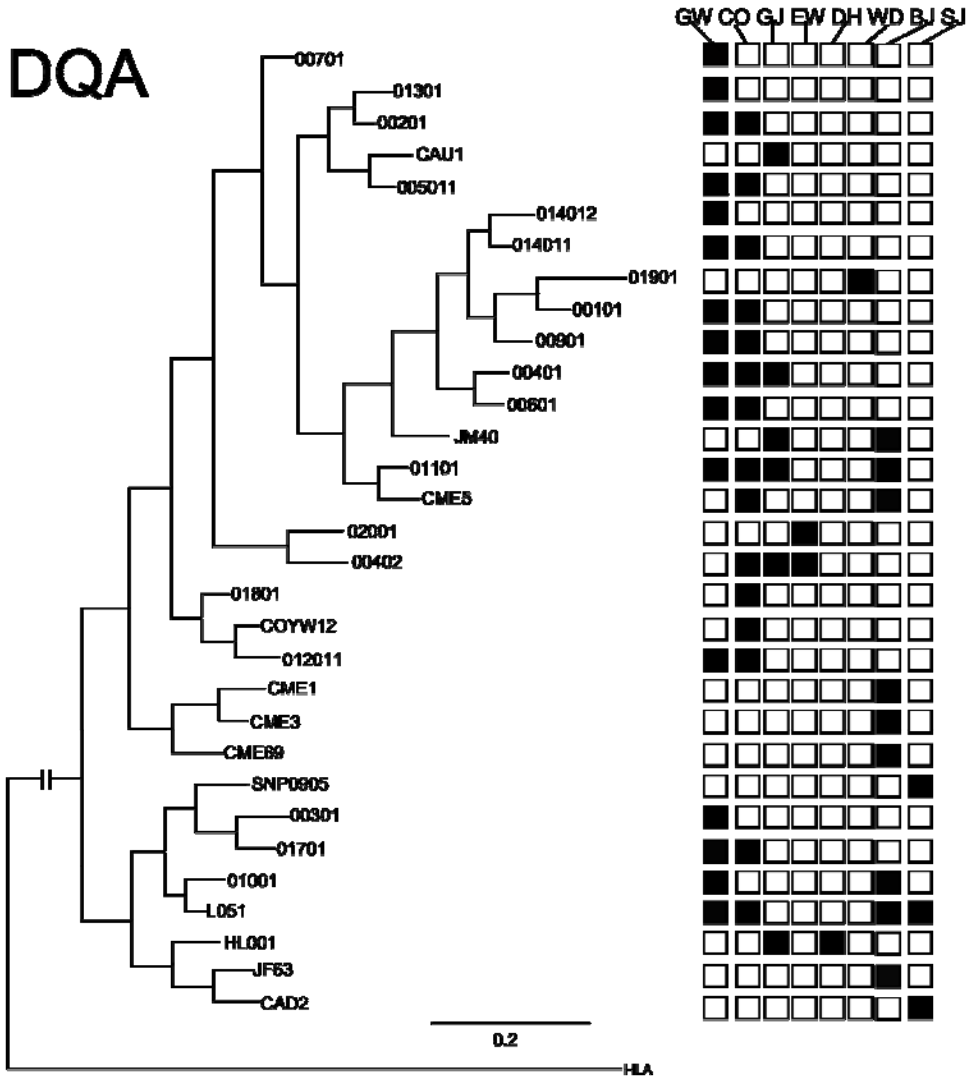

DQA
GW CO GJ EW DH WD BJ SJ
00701
01301
00201
CAU1
005011
014012
014011
01901
00101
00901
00401
00601
JM40
01101
CME5
02001
00402
01801
COYW12
012011
CME1
CME3
CME89
SNP0905
00301
01701
01001
L051
HL001
JF53
CAD2
0.2
HLA

Figure 5.6: Mr Bayes tree of wolf-like canid DQA alleles, rooted with an HLA sequence. No branches were resolved with posterior probabilities ≥70%. Presence of alleles in different species (codes from Table 5.1) are shown by filled boxes, which show allele sharing between different species and the clustering or scattered distribution of alleles in different species.

Table 5.1: Number of DRB, DQB, DQA alleles found in the wolf-like canids. The number of samples used at each locus is shown in brackets. Provenance of the samples for each species is given. Two-letter abbreviations are shown for each species.

Species	Abb	Provenance	DRB	DQB	DQA
Grey wolf	GW	Europe and North America	27 (~514)	21 (~504)	16 (~521)
Coyote	CO	California, Washington, Utah[1]	17 (19)	11 (15)	15 (29)
Golden jackal	GJ	E&S Africa, Israel	6 (9)	6 (13)	5 (11)
Ethiopian wolf	EW	Bale Mountains, Ethiopia[2]	4 (99)	5 (99)	2 (99)
Dhole	DH	Captive and unknown	4 (11)	1 (11)	1 (11)
African wild dog	WD	Eastern & Southern Africa	21 (541)	2 (398)	1 (271)
Black-backed Jackal	BJ	Eastern & Southern Africa	13 (77)	13 (81)	9 (83)
Side-striped Jackal	SJ	Eastern Africa	4 (6)	3 (8)	2 (6)

[1] These coyote populations were selected as they outside the current range of Grey wolves, a species which Coyotes can hybridise with. [2] Sampled from a single population.

Table 5.2: Allele sharing amongst the wolf-like canids at the DRB (a), DQB (b) and DQA (c) loci. Shown is the number of alleles that the species at the top of the column shares with the corresponding species on each row, out of the total number of alleles found in that species. Species abbreviations are from Table 5.1. Zero values have been coloured grey to improve clarity.

a) DRB	GW↓	CO↓	GJ↓	EW↓	DH↓	WD↓	BJ↓	SJ↓
GW	-	**4/17**	0/6	0/4	0/4	0/21	0/13	0/4
CO	**4/27**	-	0/6	0/4	0/4	0/21	0/13	0/4
GJ	0/27	0/17	-	0/4	0/4	0/21	0/13	0/4
EW	0/27	0/17	0/6	-	0/4	0/21	0/13	0/4
DH	0/27	0/17	0/6	0/4	-	0/21	0/13	0/4
WD	0/27	0/17	0/6	0/4	0/4	-	0/13	0/4
BJ	0/27	0/17	0/6	0/4	0/4	0/21	-	0/4
SJ	0/27	0/17	0/6	0/4	0/4	0/21	0/13	-

b) DQB	GW↓	COY↓	GJ↓	EW↓	DH↓	WD↓	BJ↓	SJ↓
GW	-	**4/11**	**2/6**	**2/5**	0/1	0/1	**1/13**	0/3
CO	**4/21**	-	**2/6**	**1/5**	0/1	0/1	**2/13**	0/3
GJ	**2/21**	**2/11**	-	**2/5**	0/1	0/1	**1/13**	**1/3**
EW	**2/21**	**1/11**	**2/6**	-	0/1	0/1	0/13	0/3
DH	0/21	0/11	0/6	0/5	-	0/1	0/13	0/3
WD	0/21	0/11	0/6	0/5	0/1	-	0/13	0/3
BJ	**1/21**	**2/11**	**1/6**	0/5	0/1	0/1	-	0/3
SJ	0/21	0/11	**1/6**	0/5	0/1	0/1	0/13	-

c) DQA	GW↓	CO↓	GJ↓	EW↓	DH↓	WD↓	BJ↓	SJ↓
GW	-	**11/15**	**2/5**	0/2	0/1	0/1	**3/9**	**1/2**
CO	**11/16**	-	**3/5**	**1/2**	0/1	0/1	**2/9**	**1/2**
GJ	**2/16**	**3/15**	-	**1/2**	0/1	0/1	**2/9**	0/2
EW	0/16	**1/15**	**1/5**	-	0/1	0/1	0/9	0/2
DH	0/16	0/15	0/5	0/2	-	0/1	0/9	0/2
WD	0/16	0/15	0/5	0/2	0/1	-	0/9	0/2
BJ	**3/16**	**3/15**	**2/5**	0/2	0/1	0/1	-	**1/2**
SJ	0/16	**1/15**	0/5	0/2	0/1	0/1	**1/9**	-

Table 5.3: Results from tests for selection at the DRB locus: PARRIS was used to test for selection across the whole sequence; MEGA was used to test for selection in the PBR and non-PBR separately; FEL was used to test for evidence of selection at individual codons. Significance is indicated by asterisks: for PARRIS and FEL * <0.1; **<0.5; for MEGA *<0.5, **<0.01. Species with just one allele were not tested; species codes are from Table 5.1.

DRB		PARRIS	MEGA						FEL
			PBR			non-PBR			
	# alleles	ω (SD)	dS	dN	ω	dS	dN	ω	Codon
GW	27	3.401±2.288 **	0.177±0.102	0.383±0.075	2.164 **	0.013±0.010	0.028±0.009	2.154	11**, 13** (PBR), 58* (non-PBR)[1]
CO	17	2.719±2.045 **	0.189±0.111	0.405±0.088	2.143 *	0.017±0.012	0.023±0.008	1.353	11* (PBR)
GJ	6	1.795±0.962	0.171±0.123	0.320±0.083	1.871	0.017±0.013	0.023±0.009	1.353	11* (PBR)
EW	4		0.188±0.108	0.425±0.092	2.261 *	0.000±0.000	0.035±0.013	nc **	
DH	4		0.079±0.055	0.181±0.085	2.291	0.013±0.013	0.012±0.009	0.923	
WD	21	1.542±1.128	0.183±0.098	0.303±0.089	1.656	0.011±0.010	0.013±0.008	1.182	
BJ	13	1.622±1.225 *	0.217±0.128	0.343±0.086	1.581	0.031±0.180	0.032±0.012	1.032	13** (PBR)
SJ	4		0.139±0.077	0.244±0.096	1.755	0.010±0.010	0.006±0.006	0.600	

[1]Codon 58 is adjacent to PBR codon #57

Table 5.4: Results from tests for selection at the DQB locus in different species, as for the DRB (Table 5.3).

DQB		***PARRIS***	***MEGA***						***FEL***
			PBR			non-PBR			
	# alleles	ω (SD)	dS	dN	ω	dS	dN	ω	Codon
GW	21	3.374 ± 2.521 **	0.114±0.07	0.367±0.097	3.219 **	0.012±0.007	0.024±0.009	2	13**, 85** (PBR)
CO	11	2.682 ± 2.024 **	0.113±0.068	0.341±0.094	3.018 **	0.007±0.005	0.024±0.01	3.42857	not sig
GJ	6	2.668 ± 2.461	0.084±0.053	0.241±0.079	2.869 *	0.006±0.007	0.017±0.009	2.83333	not sig
EW	5	20 ± 0	0.026±0.028	0.118±0.06	1.573 *	0±0	0±0	Nc	not sig
WD	2		0.195±0.186	0.094±0.059	0.482	0±0	0.013±0.009	Nc	-
BJ	13	3.994 ± 3.483 **	0.087±0.055	0.243±0.072	2.793 *	0.003±0.003	0.015±0.007	5	63* (non PBR)[1]
SJ	3		0.012±0.013	0.023±0.012	1.917	0±0	0.008±0.006	Nc	-

[1]Codon #63 is adjacent to PBR codon #62.

Table 5.5: Results from tests for selection at the DQA locus in different species, as for the DRB (Table 5.3).

DQA		***PARRIS***	***MEGA***						***FEL***
			PBR			non-PBR			
	# alleles	ω (SD)	dS	dN	ω	dS	dN	ω	Codon
GW	16	2.630±6.527 **	0.000±0.000	0.083±0.041	nc *	0.005±0.003	0.005±0.003	1.000	52** (PBR)
CO	15	20.00±0.000 **	0.000±0.000	0.081±0.038	nc *	0.000±0.000	0.006±0.004	0.000	52* (PBR)
GJ	5	1.590±3.894	0.000±0.000	0.062±0.032	nc *	0.000±0.000	0.009±0.005	nc *	
EW	2	-	0.000±0.000	0.057±0.061	nc	0.000±0.000	0.000±0.000	0.000	
BJ	9	2.123±3.395	0.000±0.000	0.074±0.039	nc *	0.009±0.007	0.007±0.004	0.000	
SJ	3	-	0.000±0.000	0.019±0.019	nc *	0.028±0.02	0.004±0.004	0.000	

Table 5.6: Diversity at three MHC class II loci across the wolf-like canids. Results are split into three tables: a) DRB; b) DQB; c) DQA. Diversity based on numbers of alleles, sequence diversity (π, θ), amino acid diversity (a.a.), variability in PBR (var PBR) and Observed heterozygosity (Ho).

a) DRB	n	# Alleles	π (SD)	θ (SD)	a.a. (SE)	# var PBR	H_o% (n)
GW	514	27	0.073±0.003	0.049±0.017	13.18±2.24	16	72.4[1]
CO	19	17	0.071±0.006	0.050±0.020	12.26±2.21	16	59.1[2]
GJ	9	6	0.065±0.010	0.055±0.027	11.6±2.34	13	53.8
EW	99	4	0.073±0.019	0.069±0.038	13.83±2.6	14	Na
DH	11	4	0.038±0.011	0.034±0.020	5.33±1.7	7	27.3
WD	541	21	0.055±0.004	0.032±0.012	6.92±1.65	13	75.0
BJ	77	13	0.075±0.007	0.051±0.021	12.56±2.29	15	65.5
SJ	6	4	0.042±0.019	0.044±0.025	5.83±1.62	10	37.5

b) DQB	n	# Alleles	π (SD)	θ (SD)	a.a. (SE)	# var PBR	H_o% (n)
GW	504	21	0.067±0.003	0.046±0.015	11.36±2.19	15	70.4[1]
CO	15	11	0.063±0.007	0.049±0.020	10.71±2.06	15	78.3[2]
GJ	13	6	0.047±0.014	0.046±0.023	7.33±1.70	11	76.9
EW	99	5	0.019±0.003	0.014±0.008	3.20±1.38	6	na
DH	11	1	0	0	0	0	0
WD	398	2	0.030±0.015	0.030±0.022	5.00±2.15	3	14.0
BJ	81	13	0.045±0.008	0.039±0.016	7.67±1.62	14	65.5
SJ	8	3	0.017±0.007	0.017±0.012	2.67±1.29	2	37.5

c) DQA	n	# Alleles	π (SD)	θ (SD)	a.a. (SE)	# var PBR	H_o% (n)
GW	521	16	0.016±0.002	0.015±0.007	3.23±1.20	5	69.5 [1]
CO	29	15	0.015±0.002	0.015±0.007	3.16±1.12	6	76.9 [2]
GJ	11	5	0.015±0.003	0.016±0.009	3.20±1.17	4	41.7
EW	99	2	0.008±0.004	0.008±0.007	1.00±0.99	3	na [3]
DH	11	1	0	0	0	0	0
WD	271	1	0	0	0	0	0
BJ	83	9	0.016±0.002	0.015±0.008	3.06±1.12	4	62.7
SJ	6	3	0.011±0.004	0.011±0.008	1.33±0.09	4	0

[1] Based on 141 European wolves from (Seddon & Ellegren 2004). [2] It is possible that there are null alleles at Coyote DRB and DQB loci, which would reduce Ho. [3] This information was not available.

Chapter 6: Combining studbook information, neutral markers and MHC data to assess the genetic status of the European African wild dog captive breeding programme.

6.1 Abstract

Captive breeding programs potentially play an important role in species conservation by serving as both an insurance against extinction in the wild as well as a source of individuals for re-introductions. To achieve these roles captive programmes must establish a healthy and self-sustaining population that represents the genetic diversity from wild populations. However, in reality most programmes are initiated from a small number of initial founders and thus may become inbred and suffer losses in genetic diversity. In this study I combined studbook information with genetic data from neutral markers and the major histocompatibility loci (MHC) to assess the genetic status of the European captive population of African wild dogs. These analyses showed that the captive population is derived from Southern African countries, and that a large proportion of the genetic diversity from the wild populations is represented in captivity. This high diversity appears to be the result of both the diverse origin of the wild founders as well as recent imports of 'new blood' from South African captive facilities. However, it was also shown that many founder lineages are currently over or under-represented and that ~10% of the population was produced by recent first order inbreeding events. Genetic management suggestions are proposed to prevent further losses of diversity and to reduce rates of inbreeding. Overall, these results highlight the value of combining studbook, neutral and MHC data, as each data type provided different information. Incongruence of diversity metrics based on MHC and neutral markers advocates the importance of combining neutral and adaptive markers in assessments of diversity.

6.2 Introduction

Conservation programmes are designed to increase the long-term persistence of species (Ramirez et al. 2006). Most commonly this is achieved through *in situ* conservation, which aims to reduce threats to extant populations and reintroduce a species where populations have been extirpated. However, *ex situ* captive breeding programmes also play a crucial role in endangered species conservation. Indeed, 38 species exist only in captivity, having gone extinct in the wild (IUCN 2010), and many other species have only survived extinction due to successful reintroduction of captive animals into the wild (e.g. Przewalski's horse (*Equus ferus przewalskii;* (Hedrick et al. 1999), black footed ferret (*Mustela nigripes* (Wisely et al. 2003). The two primary roles of captive breeding programmes are to act as an insurance against extinction in the wild and as a source of individuals for reintroductions (Bauman et al. 2004; Glatson 2001). To successfully achieve these roles, they must not only establish a viable, healthy and self-sustaining *ex situ* population but must also conserve the adaptive genetic diversity present in wild populations, as this retains the evolutionary and adaptive potential of the species (Bradshaw 1991; Ramirez et al. 2006). In reality though, breeding programmes are derived from a small number of initial founders and subsequently maintained as a relatively small population because of space restrictions (Hedrick et al. 1999). These characteristics result in a small effective population size (N_e), which increases risk of inbreeding and loss of genetic diversity; threats which pose a considerable risk to both the viability and success of *ex situ* conservation (Breen et al. 1995). Genetic management is thus fundamental to the maintenance of healthy and genetically diverse captive populations.

Traditionally, genetic management of captive breeding programmes has relied solely on studbooks, but these are often incomplete and they do not provide any information on genetic diversity. The increasing availability of molecular markers has led to a rise in genetic assessments of captive breeding programmes, which are used to both fill in gaps in the studbook and to address questions relating to genetic diversity (e.g. Gautschi et al. 2003; Nsubuga et al. 2010; Ramirez et al. 2006). However, most commonly, genetic assessments utilise only neutral markers (e.g. microsatellites and mtDNA). Whilst these can address many useful questions (e.g. source of founders, inbreeding, relatedness), the ability of neutral markers to reflect adaptive genetic diversity has been heavily criticised (Lynch 1996; Reed & Frankham 2003). The Major Histocompatibility Complex (MHC) is a group of genes known to be important for immune responses and is one of the best understood set of adaptive genes (Bernatchez & Landry 2003). As a result, some captive genetic studies have incorporated the MHC as a proxy for adaptive genetic diversity (Hedrick et al. 2000; Zhang et al. 2006; Zhu et al. 2007).

However, there is a paucity of studies comparing MHC and neutral markers within captive breeding programmes. As such, the value of adding MHC markers to genetic studies of captive breeding programmes is not known.

Assessments of MHC variation are also pertinent due to their role in immunity and growing concerns about the threat that infectious diseases pose to endangered species (Daszak et al. 2000). The risk of diseases to wildlife has increased as human populations and their accompanying domesticated animals have spread, resulting in spill-over infections between domesticated and wild animals (Daszak et al. 2000; Randall et al. 2004). For example, rabies virus transmission between domestic dogs and Ethiopian wolves (Randall et al. 2006), and feline leukaemia virus between domestic cats and the Iberian lynx (López et al. 2009).

African wild dogs, hereafter wild dogs, are a red list endangered species which has suffered extensive declines in the wild (IUCN 2010). They now persist over just 7% of their former range in a few small and fragmented populations which together total less than 8,000 individuals (IUCN/SSC 2008, 2009; McNutt et al. 2008; Woodroffe et al. 1997). In addition to wild populations, there is also a growing captive population of wild dogs. Wild dogs were first brought into captivity in 1901, and in 2008 there were 636 captive wild dogs (364♂; 261♀; 11 unknown) in 108 institutions that contribute to wild dog captive breeding programmes (ISIS 2010; Verberkmoes & Verberkmoes 2008). These are managed as four separate regional populations: the European Endangered Species Programme (EEP) in Europe, as well as by comparable programmes in Northern America, Australia and South Africa. Zoos contributing to the European captive breeding programme have held almost half of world's captive population (n~270 in 2008 (Verberkmoes & Verberkmoes 2008) and therefore could form a critical component of *ex situ* conservation efforts. However, informed genetic management of the wild dog captive population has been difficult because it relies solely on an incomplete studbook.

The aim of this study was to assess the genetic status of the European captive wild dog population, using a combination of studbook data and genetic information based on neutral and MHC markers for >80% of the extant population. I specifically addressed four questions: 1) How much of the diversity from wild populations is represented in the European captive population?; 2) How have recent imports from South African captive facilities affected the genetic status of the European captive population?; 3) How can genetic management of the European captive population be improved?; 4) Are genetic diversity metrics from neutral and MHC markers correlated?

6.3 Methods

6.3.1 Studbook and pedigree analyses

Studbook analyses were based on the 2008 European regional African wild dog studbook provided by the studbook coordinator (H. Verberkmoes). There was too much missing data for analysis in the standard studbook analysis programme, PM2000 (Pollack et al. 2001). Therefore, the studbook was converted into spreadsheet format and analysis was conducted by hand. Death dates were missing for many individuals in the studbook, but this information was needed to calculate the number of individuals which died without leaving descendents. As such, any individuals in the studbook that were more than 20 years old were considered dead, as no wild dog has lived > 15 years in European zoos (H.Verberkmoes, *pers comm*.). The number of wild founders brought into the European zoos was estimated based on individuals in the studbook whose parents had been listed as 'wild'. However, this should be considered as a conservative estimate of the number of wild founders as the origin of many individuals in the studbook was listed as unknown. A pedigree of the European captive population was built manually, using the drawing programme SMARTDRAW (www.smartdraw.com), as the studbook data were not appropriate for standard pedigree programmes which typically are designed for humans and thus have difficulty accommodating multiple partners and inbreeding. To assist navigation of the pedigree, I refer to the centre, left or right side of the pedigree. This literally refers to the centre, right or left side of the printed out pedigree. To improve clarity of the pedigree, all individuals that died without leaving descendents were removed, unless they had been sampled for this study. For demonstration, I also retained a group representing two founder lines that recently died out (boxed by dashed lines at the top centre of the pedigree). The pedigree was used to calculate the individuals from which the extant population is descendent. These are referred to as 'putative founders', as the relatedness between many of these individuals was unknown. Following standard convention, captive wild dogs are referred to here by their studbook ID, which is indicated by the prefix #. Numbers beginning with 'T' are temporary studbook ID numbers. For individuals without a studbook ID, a local ID without a # prefix was used instead.

6.3.2 Samples

With the assistance of the studbook keeper (H. Verberkmoes), I compiled and sent genetic sampling kits to all 41 zoos participating in the European breeding programme in 2007. Sampling kits consisted of: 1) recommended procedures for anesthetisation and detailed instructions on how to collect, store and send samples; 2) blood, tissue and serum collection tubes; 3) pre-printed adhesive tube labels specific for the wild dogs of the

zoo receiving each kit; 4) UK import licence; 5) pre-labelled shipping envelope with specimen bags. In total, 223 samples were sent from 34 zoos (Table 6.1) over three years, which represents approximately 83% of the population (n~270) and >75% of zoos (34/44). Blood and tissue samples were preserved in TES buffer (50ul - 500mM Tris Base, 50ul - 500mM EDTA, 50ul-10% SDS, 350ul dH_20) and serum and hair samples were frozen. DNA was extracted from blood, tissue and serum samples using DNeasy extraction kits according to the manufacturer's instructions (Qiagen, Crawley, UK). Hair was extracted according to a user-developed protocol available from Qiagen (Qiagen 2006).

Fifteen of the 223 wild dogs sampled in European zoos were recently imported from South African captive facilities: nine from the Hoedspruit captive breeding facility in South Africa (also known as Kapama, but hereafter, Hoedspruit); six from the De Wildt Cheetah and Wildlife Centre in South Africa (also known as Ann van Dyk Cheetah Centre, but hereafter, DeWildt). I acquired DNA from a further 60 South African captive wild dogs as follows: 45 samples from DeWildt collected in 1993 (n=15; R.K.Wayne), and 2002 (n=30; A.Bastos, T.Matjila, H.Strydom). Fifteen samples were available from Hoedspruit collected in 1993-4 (R.K.Wayne).

Finally, to enable comparisons of genetic diversity between free ranging and captive populations and zoos, I included samples from six study populations in Southern Africa (n=257), which is the region of Africa from which the European zoo captive wild dog population is descendent (Woodroffe et al. 1997). However, the number of samples that were successfully typed in each population varied according to the genetic marker: Kruger, South Africa (n=90-94); NE Namibia (n=4-10); Ghanzi, Botswana (n=1); Okavango, Botswana (n=56-90); Lowveld, Zimbabwe (n=15); Hwange, Zimbabwe (n=14-47) (Chapter 4). Whilst the sample sizes from Ghanzi and NE Namibia were small, they were useful for the distribution of mtDNA haplotypes and MHC haplotypes, where they improved geographic coverage. However, these samples were excluded from genetic structure and population diversity analyses.

For the purposes of genetic analyses the samples above were subdivided into five categories:

- "EUZ": All samples from European zoo wild dogs.
- "EUZ_pure": European zoo wild dogs excluding recent (1990+) South African imports and any of their descendents.
- "SAZ_imp": Wild dogs recently imported from South Africa to EU zoos.

- "SAZ": Wild dogs from South African captive facilities (Hoedspruit and Dewildt), including the 15 SAZ_imp.
- "WILD": Free ranging wild dogs from populations in Southern Africa.

6.3.3 Genetic typing

mtDNA

To assess the provenance of the founders of the EUZ I amplified a 327 bp segment of mtDNA D-Loop control region 1 by PCR using a modified version of the canid-specific primers, Thr-L and DLH, which were redesigned specifically for African wild dogs (Leigh 2005): forward 5' ACT ATT CCC TGA TCT CCC CC 3'; reverse CAG GAA ACA GCT ATG ACC CCT GAA GTA AGA ACC AGA TGC C. The underlined section of the reverse primer marks an M13 tag, which was used to permit sequencing in a single direction. PCRs were performed in a 20-µl reaction volume containing: 1.25 x Q solution (Qiagen); 1.25 x PCR buffer (containing 15mM $MgCl_2$); 3.1 mM MgCl; 0.2 mM of each dNTP (Invitrogen); 0.19 µM of each primer; 1 unit of Hot Star *taq* (Qiagen); and approximately 10ng of template DNA (except for negative controls). PCR was conducted according to the following protocol: 5 min at 95°C, 30 cycles of 95°C for 30s, 55°C for 30s, and 72°C for 30s. The protocol ended with a final extension of 72°C for 10 minutes. The number of amplification cycles was increased from 30 to 37 cycles for weak DNA samples derived from hair and serum. PCR products were cleaned with ExoSAP-IT (USB) according to the manufacturer's instructions and then sequenced using an ABI 3730 automated sequencer (by The Sequencing Service, University of Dundee). Sequences were aligned and analysed using Geneious Pro v 4.5.5 (Biomatters Ltd). In total, I typed 33 samples that were representative of 5/5 SAZ_imp founder lines and 9/12 EUZ_pure mtDNA founder lines. It was not possible to sample 3/12 EUZ_pure founders because there were no extant female lineages of these founders. The data from captive animals was compared with mtDNA data from WILD samples (n=266; Chapter 4; (Girman et al. 2001).

Microsatellite genotyping

I assessed diversity at 10 microsatellite loci shown to be polymorphic in free-ranging populations of wild dogs (Chapter 4): PEZ08, PEZ12, PEZ15 (J. Halverson in Neff *et al*. 1999); FHC2010, FHC2054, FHC2611, FHC2658, FHC2785, FHC3399, FHC3965 (Guyon et al. 2003; Neff et al. 1999); Breen *et al*. 2001). All selected loci were located on different chromosomes (Neff et al. 1999). The forward primer of each pair was dye-labelled with ABI fluorescent dyes: NED (yellow), 6-FAM (blue) or HEX (green). Samples

were amplified alongside negative controls by multiplex PCR using Qiagen Multiplex PCR mix. I followed default reagent concentrations recommended by the manual except in cases of DNA derived from serum, hair and blood spots, where 0.4ul of 10mM Bovine Serum Albumin (Promega) was added per PCR reaction. PCR was performed on PTC-200 (MJ Research) theromcyclers with the following touchdown protocol: 15 min at 95°C, 12 touchdown cycles of 94°C for 30 s, followed by 1 min 30 s annealing, starting at 60°C and reducing at 0.5°C per cycle, and 72°C for 1 min. This was followed by 33 cycles of 89°C for 30 s, 55°C for 1 min, and 72°C for 1 min. The protocol ended with a final extension of 60°C for 30 minutes. Samples were run with a ROX 500 size standard on an ABI 3730 (by The Sequencing Service, University of Dundee). Microsatellites were analysed using GENEMAPPER 4.7 (Applied Biosystems). Weak DNA samples have a higher probability of allelic drop out; therefore, I amplified and genotyped DNA samples derived from hair and serum up to seven times. I also re-amplified and genotyped a further 50% of blood and tissue samples to verify results. Any samples that had data missing for three or more loci were excluded. I used pedigree analyses to check for the expected segregation of alleles within families, and visually assessed the data for evidence of lack of amplification at loci which could indicate individuals homozygous for null alleles. The data from captive animals were compared with microsatellite data for WILD samples (n=174; Chapter 4).

MHC typing

Sequence-based typing was conducted on exon 2 of the DLA-DRB1 locus (hereafter referred to as DRB), which was previously shown to be highly variable in African wild dogs (Marsden et al. 2009). The DRB locus was typed according to Marsden et al. (2009), which is included as Chapter 2. In brief, DRB sequence data were analysed using Match Tools and Match Tools Navigator (Applied Biosystems), as described in Kennedy et al (2002b). Alleles were named as in Chapter 4. The data from captive animals were compared with MHC data from WILD samples (n=185; Chapter 4). The studbook-based pedigree was used to confirm that a single locus was amplified at DRB (see Marsden et al. 2009).

6.3.4 Genetic Analyses

Effective population size and genetic diversity

Effective population size (N_e) estimates for the EUZ and SAZ were conducted using NeEstimator v1.3 (Peel et al. 2004). Specifically, I used the linkage disequilibrium method (Hill 1981), which tests for evidence of linkage disequilibrium between alleles at

different loci arising as a result of increased genetic drift at smaller effective population sizes (Hill 1981). N_e estimates for WILD populations were taken from Chapter 4.

Microsatellite and DRB diversity were measured as allelic richness (A_R), fixation index (F_{IS}), observed (H_o) and expected heterozygosity (H_e), as calculated in GENALEX6 (Peakall & Smouse 2006). Allelic richness standardised for sample size (R_S) was calculated using rarefaction in HP-RARE (Kalinowski 2005). I tested for significant differences in microsatellite H_o, H_e, F_{IS} between sample categories using two sample t-tests in Minitab 15 (Minitab 2007). Genetic diversity estimates for WILD populations were taken from Chapter 4.

Genetic structure

I assessed genetic structure of the EUZ and SAZ samples using Bayesian clustering analysis conducted in STRUCTURE v 2.2 (Pritchard et al. 2000). STRUCTURE uses a Bayesian clustering model-based algorithm to elucidate the number of genetic clusters (K) within a sample set. For each K value, the model generates an estimated log probability of the data, which is used to determine the most likely value among the range of K values tested (Pritchard et al. 2000). A likelihood of assignment value is also generated for each individual to each cluster (Pritchard et al. 2000). Runs were conducted assuming no prior population information, with correlated allele frequencies and admixture, 200,000 burn-in cycles, 1,000,000 Markov chain Monte Carlo runs (MCMC) for K = 1-10, with ten replicates per K. I plotted likelihood values and variance amongst the 10 replicates at the different values of K, as well as the ΔK statistic, which assesses the rate of change in the log probability of the data between successive K values (Evanno et al. 2005) using STRUCTURE HARVESTER v.05. (Earl 2009). The value of K that best fit the data was selected based on the ΔK statistic and consistency amongst replicates (Pritchard et al. 2000). I also conducted assignment tests on the 15 recent South African Imports, by implementing the USEPOPINFO option in STRUCTURE. In this mode STRUCTURE calculated the posterior probabilities of membership of the 15 South African Imports to four Southern African populations: Hwange; Okavango; Lowveld; and Kruger. Pairwise genetic differentiation between EUZ_pure, SAZ and WILD populations was calculated using the estimator D_{est} (Jost 2008) computed in the programme SMOGD (Crawford 2009), using 1000 bootstraps.

Parentage

I used microsatellite data to verify parentage across seven litters (2-8 pups/litter) from four EUZ (Aalborg, Boras, Kolmarden and Pontscorff). These litters were specifically selected because samples were available for the studbook dam and sire, as well as all

other potential parents (i.e. other adults in the same enclosure) with the exception of one unsampled male for the 2008 Ponscorff litter and one unsampled female for the 2007 Kolmarden litter. For each offspring, putative parents were excluded when mismatches were found at two or more loci (Jones & Ardren 2003).

6.4 Results

6.4.1 Studbook and pedigree analyses

Between 1901 and 2008, 2175 wild dogs have been held in EUZ, including 80 wild caught founders (42 males; 36 females) (Verberkmoes & Verberkmoes 2008) and 41 wild dogs imported from captive South African facilities since 1990. Of these 2175 individuals, only 415 have produced offspring (19%) and only 271 of 415 breeders, and just 18/80 founders, have descendents alive today .

The EUZ currently (2008) have 44 zoo spaces for wild dogs (including 26 spaces for breeding groups), which collectively hold ~270 wild dogs (155 males; 114 females; 1 unknown sex (Verberkmoes & Verberkmoes 2008). The reconstructed pedigree (Figure 6.1) shows that the extant EUZ population is descendent from 38 putative founders (17 females; 21 males): 18 wild born founders; 10 recent (1990+) imports from South African captive facilities (Hoedspruit-4, DeWildt-4, Johannesburg zoo-1) and 11 individuals of unknown parentage and origin. Given that some of these 38 individuals may be related, especially those imported from South African captive facilities, this may be an over-estimate of the true number of founders of EUZ.

The pedigree (Figure 6.1) demonstrates four important points about the genetic status of the EUZ: 1) There are very unequal founder contributions, with a heavy bias to founders from one side of the pedigree (the right side). For example, founders #896 and #897 (boxed by dashed lines at the centre and top of the pedigree) produced few descendents, the last of which died in 2007. 2) There is a large skew in reproductive contributions, with a few individuals producing a very large number of offspring but most individuals producing none. 3) There are a number of close inbreeding events (sibling matings/parent-offspring, indicated by red lines). For example, four litters totalling 29 offspring were produced by full sibling matings at a single zoo between 2004 and 2005, 21 of which survived (Figure 6.1; centre of pedigree, #5705-5709; #T5819-T5829; #5719-5724). In total, 10% (27/270) of the extant EUZ population are the offspring from first order relative (sibling-sibling/parent-offspring) inbreeding events. 4) Most breeding groups are formed within the left side of the pedigree, or within the right side of the pedigree, which further increases inbreeding. Overall, the bias in founder representation, large reproductive skew and close inbreeding events will have reduced N_e in the EUZ, thus accelerating the loss of genetic diversity as well as increasing the likelihood of inbreeding depression.

6.4.2 Effective population size and genetic diversity

mtDNA

I used mtDNA sequencing to trace the origin of 5 SAZ_imp and 12 wild founders of EUZ_pure. Both the 5 SAZ_imp and 12 EUZ_pure wild founders were found to share three mtDNA haplotypes (E2, S2 and Z1; Table 6.2;Table 6.3). The studbook lists the origins of the EUZ_pure wild founders as South Africa, Namibia and Unknown (no information was available for SAZ_imp). Comparison of EUZ mtDNA haplotypes with wild population data (Table 6.3) shows that the presence of the E2, S2 and Z1 haplotypes are inconsistent with the origin of wild founders being just Namibia and South Africa. Rather, the founder base of the EUZ most likely includes individuals from Namibia, South Africa, Zimbabwe and/or Botswana. Two Namibian founders (#895 caught ~1982; #1986 caught ~1993) were found to have mtDNA haplotypes that are not known from wild populations in Namibia (Table 6.3). This may be the result of either errors in the studbook or incomplete sampling of Namibian populations.

Microsatellites

In total, I amplified 244 EUZ and SAZ wild dogs at 10 microsatellite loci and combined these data, with data from 174 wild dogs derived from wild Southern African populations (Chapter 4). Amongst the EUZ samples, segregation of alleles in pedigreed zoo families detected allelic drop out at only a single locus (Pez08). Whilst the number of individuals affected by drop out at the Pez08 locus was small, the locus was excluded as it may have significantly affected parentage assessments.

Based on the linkage disequilibrium method, N_e for EUZ and SAZ was estimated to be 26.6 and 15.6 respectively, which are similar to estimates from WILD populations (7.8-19.6; Table 6.5). Furthermore, N:N_e was similar between EUZ (0.10) and WILD populations (0.05-0.1; Table 6.5).

Microsatellite diversity in EUZ was comparable to that in WILD populations: A_R (6.4), R_S (5.2), H_o (0.76) and F_{IS} (-0.05) in EUZ compared to A_R (4.6-7.4), R_S (4.1-5.7), H_o (0.62-0.80), F_{IS} (-0.11- -0.02) in WILD populations (Table 6.5). Furthermore, 72.3% of microsatellite alleles from WILD populations in Southern Africa were represented in EUZ (Table 6.5), including two WILD private alleles; one from the Lowveld and one from Okavango.

Comparison of diversity before (EUZ_pure) and after (EUZ) recent South African Imports, indicates that these imports have increased the proportion of wild microsatellite

alleles represented in EUZ from 57.4 to 72.3%, but resulted in little change in microsatellite diversity: R_S (4.2 vs. 4.8); H_o (0.77 vs. 0.76); and F_{IS} (-0.11 vs -0.05). Microsatellite diversity (R_S) was found to be higher in SAZ than EUZ_pure (R_S: 5.2, 4.2), although both were within the range found in WILD populations (R_S: 4.1-5.7; Table 6.5). However, microsatellite H_o was significantly higher for the EUZ_pure (H_o = 0.77) than SAZ individuals (H_o 0.63; t test, p = 0.042). Furthermore, F_{IS} in SAZ (F_{IS} +0.04) was outwith the range of wild populations (-0.11, -0.02), and significantly higher than EUZ_pure (EUZ_pure, F_{IS} -0.11; t-test, p=0.027).

MHC

Amongst 211 EUZ wild dogs I detected 10 of the 13 DRB alleles known in free-ranging Southern African populations (Table 6.4), including three private MHC alleles; one from Okavango and two from Kruger. The EUZ had two further DRB alleles; one (B1) described only in free-ranging Eastern African populations (Selous, Masai Steppe, Serengeti-Mara-refer to table Chapter 4) and one allele (B13) that has only previously been detected in a single individual of unknown origin, which was translocated into Hwange. The frequency of DRB alleles in EUZ were highly variable; allele A1 was very common (32.2%; detected in 115/211 individuals; Figure 6.1), but there were also four DRB alleles (A7, B2, B7, B11) that were very rare (<5%; Table 6.4). Overall, DRB H_o (0.83) was high and comparable to that in wild populations (H_o 0.67-0.84; Table 6.5). F_{IS} (+0.01) in EUZ was also within the range of wild populations F_{IS} -0.08 - +0.05 (Table 6.5). Overall, there were just 36 homozygotes amongst 211 EUZ individuals (Table 6.5). It is noteworthy that 20/36 homozygotes were homozygous for allele A1 and 31/36 homozygotes were the result of close inbreeding events or inbreeding within the right side of the pedigree (Figure 6.1).

Comparison of DRB diversity between EUZ_pure and EUZ shows that recent South African imports have also increased DRB diversity: A_R increased from 7 to 12; % wild diversity represented increased from 38.5% to 76.9%; and H_o rose from 0.76 to 0.83 (Table 6.5). This diversity increase can be attributed to the very high DRB allelic richness in SAZ (A_R, 11; R_S 9.2 Table 6.5); 77% of Southern African wild DRB diversity was represented in just 66 SAZ wild dogs. However, whilst SAZ allelic diversity was high, DRB H_o (0.68) was lower than EUZ_pure (0.78), and F_{IS} (+0.20) was both much higher than EUZ_pure (-0.12) and outside the range of wild populations (-0.08 – +0.05; Table 6.5), similar to the pattern found at microsatellite loci. Specifically, the low H_o is largely attributable to the early 1993/4 SAZ individuals, 22/30 of which were homozygous at DRB compared with just 1/36 for the recent SAZ individuals.

Genetic structure

Bayesian clustering analysis of EUZ and SAZ wild dogs in STRUCTURE found that these two captive populations represented different genetic populations (Figure 6.2). The *ΔK* statistic found a single prominent peak at K=2 (data not shown), which correlated with separate clusters for EUZ_pure and SAZ. However, the recent SAZ_imp to EUZ mean that the EUZ now includes wild dogs from both genetic clusters, as well as admixed individuals (Figure 6.2). It is noteworthy that I ran STRUCTURE with the excluded locus (Pez08) incorporated back into the data set, and the same result was obtained (data not shown).

I used assignment tests in STRUCTURE to indicate the origin of the fifteen wild dogs recently imported from South Africa. The results found all 15 individuals to be of admixed ancestry; probability of assignment ≤55% for each population (data not shown).

Estimates of genetic differentiation based on D_{est} showed that the EUZ_pure and SAZ were differentiated (D_{est} 0.320) to the same degree as isolated free ranging wild dog populations (D_{est} 0.277-0.483;Table 6.6). The SAZ were quite genetically similar to Kruger (D_{est} 0.149), whereas EUZ_pure did not seem similar to any wild populations sampled (D_{est}>0.2;Table 6.6).

Parentage

I genetically assessed parentage in eight EUZ litters. In 5/8 litters I could not exclude the putative studbook dam and sire, although in 4/5 of these litters there was more than one equally well matched sire for one or more pups which means a multiple paternity litter is a possibility. Given that alternative sires were generally first order relatives (brothers, father-son), many more microsatellite loci would be required to identify a single sire/pup. For the 6th litter (Boras 2006), both the studbook dam and sire appeared to be incorrect. Here, #5517 (not #5513) was the only genetically possible dam of all of the eight pups. There also appeared to be at least two sires (#T6048, #T6049) for the litter as no single male matched all pups, although neither of the two genetically possible sires was the studbook sire (#T6047). For the seventh litter (Kolmarden 2005), genetic data confirmed behavioural data that two females (#5533, #5534) had whelped pups, despite all being raised by the alpha female. In this case, the studbook sire (#5605) appeared to be incorrect, as he was excluded as the sire of 4/7 pups, and was not the closest match of any pup. Rather, #5606 was the closest matched sire to all pups (except one pup where #5606 and another male, #5607, matched equally). For the 8th litter (Kolmarden 2007), the studbook sire (#5605) again appeared to be incorrect, as he was excluded as the sire of 4/8 pups. In this case, a multiple paternity litter is possible, as

there were multiple equally well matched sires for 6/8 pups, although #5606 was the only sire that was a match to all pups. The studbook dam was the only mature female present at Kolmarden in 2007.

6.5 Discussion

6.5.1 Genetic diversity in wild vs captive populations

One of the primary aims of captive breeding programmes is to provide an insurance against extinction in the wild. As such, it is important that a large proportion of genetic diversity from the wild is conserved in the captive population, in particular, adaptive genetic diversity which is important for evolutionary potential. Analysis of the EUZ studbook highlights that unequal founder representation and a large reproductive skew has resulted in the loss of a substantial portion of the wild founder lineages initially brought into the EUZ (62/80, 77.5%) i.e. they have no descendents alive today. Whilst this is regrettable, it is unfortunately not atypical for captive breeding programmes; 194/242 (80.2%) of the original wild Golden lion tamarin (*Leontopithecus rosalia*) founders lineages are not represented in the extant population (Frankham et al. 2002). Contributions from 18 wild founders, 11 individuals of unknown ancestry and 10 recent imports from SAZ, mean that in total the current EUZ is descendent from 38 putative founders. Mitochondrial haplotypes and the distribution of both neutral microsatellite and potentially adaptive MHC private alleles indicate that these putative founders were sourced from a number of countries in Southern African. This wide founder base appears to have resulted in a large proportion of microsatellite (72.3%) and MHC diversity (~77%) from wild Southern African populations being maintained in EUZ (Table 6.5). Furthermore, microsatellite and MHC diversity metrics (R_S, H_o, H_e and F_{IS}) and N_e/N estimates of the EUZ were within the range reported for wild populations (Table 6.5). These results are highly encouraging, as they suggest that despite the loss of so many founder lineages, the extant EUZ population has a genetic status comparable to wild populations, and adequately represents the diversity of wild populations. Such findings contrast with many other captive breeding programmes, which have been found to suffer from low genetic diversity compared to wild populations (e.g. Giant panda, *Ailuropoda melanoleuca*, (Shen et al. 2009), Asiatic lions, *Panthera leo persica* (Sachdev et al. 2005), high levels of inbreeding (e.g. Siam eld's deer *Cervus eldi siamensis* (Thevenon & Couvet 2002) and Puerto Rican crested toad *Peltophryne lemur* (Beauclerc et al. 2001), and severe founder effects (e.g. high frequencies of chondrodystrophy in Californian condors, *Gymnogyps californianus*, (Ralls et al. 2000); poor sampling of *Hymenoxys acaulis* var *glabra* resulted in an ex situ population that was unable to set seed because all founders were monomorphic for the same allele at self-incompatibility loci (Barrett & Kohn 1991). Nonetheless, similar positive results are known from other captive breeding programmes: neutral genetic diversity was comparable between wild and captive populations of Gorillas (*Gorilla gorilla* (Nsubuga et al. 2010), Iberian wolves (Canis lupus, (Ramirez et al. 2006), and Amur tigers (*Panthera tigris altaica* (Henry et al. 2009), and wild MHC diversity was

well represented in a population of Mauritian Cynomolgus macaques *Macaca fascicularis* used for medical research (Mee et al. 2009).

Despite the positive indicators from the EUZ, the pedigree highlighted unequal founder representation and close inbreeding events in the extant population, which may reduce both genetic health and genetic diversity in the future. For example, a single MHC DRB allele (A1) dominates the allelic repertoire of the EUZ (32%), despite being relatively rare in wild populations (5.8%; Table 6.4). Examination of MHC variation mapped on to the pedigree (Figure 6.1) demonstrates that the high frequency of allele A1 is likely a consequence of unequal bias in founder contributions to, and inbreeding within, the right hand side of the pedigree, where A1 is common. It is also possible that the high frequency of this allele may be a consequence of adaptation to captivity (Frankham *Pers comm.*), whereby the captive environment selects for different components of variation compared with the wild (Frankham 2008). Careful genetic management is required to prevent further losses of diversity.

6.5.2 Have recent imports of wild dogs from South African improved the genetic status of EU zoos?

Captive breeding programmes essentially represent a closed system, and therefore, inbreeding is eventually inevitable. To counter this, zoos sometimes import 'new blood', that is, individuals from outside of their specific breeding programme. Since 1990, 41 wild dogs have been imported from SAZ (Johannesburg zoo – 6; DeWildt – 23; Hoedspruit – 12), 17 of which have bred and produced viable offspring (depicted by purple highlighting in the pedigree; Figure 6.1).

Genetic structure analyses showed that EUZ_pure and SAZ represent different genetic populations that are quite differentiated from each other (Figure 6.2; D_{est} 0.320). As such, the imports would be expected to increase genetic diversity, which was found to be the case: the imports introduced five new MHC alleles, and increased the proportion of wild diversity represented at MHC and microsatellite loci by 38% and 15% respectively (Table 6.5). Nonetheless, estimates of H_o and F_{IS} based on SAZ as a whole (not just the imports) suggested that inbreeding was higher than in EUZ (Table 6.5), which is concordant with assessments of SAZ studbook data (~38% of breeding pairs were close relatives (Frantzen et al. 2001). This is a particular concern for the EUZ, because in general the ancestry of imports is not known, which makes genetic management of potentially inbred individuals difficult. As such, it raises the question of whether the advantage of new diversity outweighs the potential genetic health problems associated with inbred individuals?

Another concern pertaining to the import of wild dogs from South Africa relates to trafficking of wild born animals into captivity, which is a problem for many endangered species (e.g. Chimpanzees, *Pan troglodytes*; (Goodall 2003). Reports state that litters of wild born wild dog pups are being dug out of dens in Southern Africa countries and subsequently transported to dealers and onto captive facilities to supply demand for new blood lines (Rasmussen 2010; Scott 1991). Indeed, I conducted genetic testing to verify that a litter of dug up pups seized in Zimbabwe had come from the Lowveld region (Hwange-Unknown samples; Chapter 4). The removal of litters presents a significant threat to the persistence of wild populations, which are already very small. Unfortunately, though, there is also a huge financial incentive to remove litters from the wild (a Chinese zoo paid $80,000 for a group of 20 wild dogs, ~$4,000/pup), and as yet no CITES listing to protect them (G.Rasmussen *Pers comm*). Clearly, the EUZ would not endorse the purchase of wild dogs that had been removed from the wild, but there is no mechanism by which they can check either. Genetic tracing could be particularly useful in this respect. For example, 35 wild dogs from Hoedspruit and DeWildt sampled in 1993/4 had just one mtDNA haplotype, (S2; (Girman et al. 2001), whereas three mtDNA haplotypes (S2, Z1 and E2) were found amongst just 9 wild dogs imported from Hoedspruit and DeWildt after 2002. The presence of two new and non-South African mtDNA haplotypes amongst just 9 samples suggests that the products of this trade or their descendents may have entered Hoedspruit and DeWildt, and thus the EUZ, although the admixed origin of the imports (assignment tests) indicates that the imports themselves were not wild caught. Given that this unethical trade presents a significant risk to wild populations and demand for wild dogs from zoos will only perpetuate it, I propose a complete moratorium be placed on the import of wild dogs from South Africa to EUZ. Since all 41 recent imports were conducted without the approval of the studbook keeper, the onus and responsibility of this issue clearly lies with individual zoos. Currently, the genetic status of the EUZ wild dog population can be defined as healthy, and thus there is not the need (nor the space) for further imports. It is important to note that this is not the situation for many captive species where more wild founders (Giant panda (Shen et al. 2009) and mixing of currently isolated captive lineages (Puerto Rican crested toad (Beauclerc et al. 2001); Mexican wolf (Hedrick et al. 1997) has been recommended.

6.5.3 How can genetic management of zoo population be improved?

Studbook analyses show that most wild founder lineages were lost, and many first order relative matings occurred, early on in the breeding programme. However, some founder lineages have been lost recently (e.g. #896, #897 in 2007) and others, particularly those on the left side of the pedigree, are under-represented and thus at risk (Figure 6.1). Furthermore, the offspring of recent first order relative inbreeding events account for ~10% of individuals in the EUZ. Clearly, the current priority for the EUZ must be to

manage the population to maximise the retention of genetic diversity, whilst minimising inbreeding.

The combination of studbook, neutral (microsatellite) and adaptive (MHC) genetic diversity data described here, provides a strong basis for informed genetic management. However, the behavioural ecology of wild dogs makes genetic management challenging. They live in packs typically formed by the fusion of a group of closely related females, and a group of closely related males, which are unrelated to each other (Creel & Creel 2002). Breeding in the pack is typically monopolised by an alpha pair, which produce large litters of 10-15 pups every year, and commonly remain dominant for several years (Creel & Creel 2002; Girman et al. 1997). As such, it is not possible to specifically designate which individuals breed, and a large reproductive skew is inevitable. Unfortunately, genetic management of group-living species is one of the least-understood areas of captive breeding management (Frankham et al. 2002), and as such, there is not a generalised optimal management strategy for these types of species akin to minimising kinship for solitary species (Ballou et al. 1995; Frankham 2007; Frankham et al. 2002). The most logical solution for wild dogs might be to hold them as pairs. However, this is not viable because it would infinitely increase space requirements, conflict with the educational role of zoos, dramatically reduce breeding success (80% of EUZ wild dogs held as pairs did not breed or failed to rear pups) and potentially create welfare or ethical issues because social species can become stressed when held in small groups or isolation (H. Verbokemoes *Pers comm*). The formation and separation of breeding packs is also restricted due to high mortality and failure rates associated with formation of packs of wild dogs in zoos; ~ 50% of attempts to integrate male and female groups to form breeding packs fail due to fighting, which is often fatal (Verbokemoes *Pers comm.*). Lastly, breeding opportunities within the pack are unlikely to be equalised by contraception or sterilisation of the alpha pair because of the strict dominance hierarchy; subordinate males are physically prevented from mating by the alpha male, and many subordinate females appear to either not to ovulate (Creel et al. 1997), or have their litters killed (Malcolm 1979). As such, genetic management is limited to control over which sibling groups are selected to breed. However, even here, one is restricted due to the necessity to select: 1) groups that are of the appropriate age and character (e.g. not hand reared) for breeding; 2) opposite sex groups of similar ages and numbers of animals, to reduce the likelihood of aggression and death when forming breeding packs problems; 3) from countries where movement of animals is logistically possible.

Taking into consideration the management restrictions exerted by the social behaviour of this species, a specific detailed breeding programme is not realistic. In light of this, the following generalised guidelines were devised with the EUZ EEP management committee in May 2008 and subsequently have been implemented:

- Breeding groups should be selected with the aim to maximise mixing different founder lines, e.g. by breeding sibling groups from opposite sides of the pedigree, whilst at the same time equalising representation of founders.
- To prevent over-representation of some families, no more than two siblings/sibling groups from the same parents should be given breeding opportunities and a breeding pair should be limited to producing no more than two successful litters, unless the zoo is willing to cull the surplus litters.
- To maximise genetic diversity, breeding groups should not be formed by reciprocal swaps, whereby a zoo exchanges half its males with another zoo, in exchange for half of their females.
- To increase space availability, groups of same sex relatives should not be split unless necessary and offspring should be kept with parents to assist with rearing of at least one, and preferably two, subsequent litters. This will also ensure that parenting skills are passed on. Sterilised animals and the offspring from first-order matings, are of little genetic value, and could be culled to increase space availability.
- Fragmentation of captive programmes into sub-populations that are managed separately, with occasional migration, has been shown to reduce loss of genetic diversity and adaptation to captivity (Frankham 2008; Margan et al. 1998). Therefore, it is not recommended that North American, European, Australian and South African captive breeding programmes are merged with an international studbook. However, occasional exchanges with North American and/or Australian zoos could be considered as a management option to reduce inbreeding, as an alternative to further imports from South Africa.

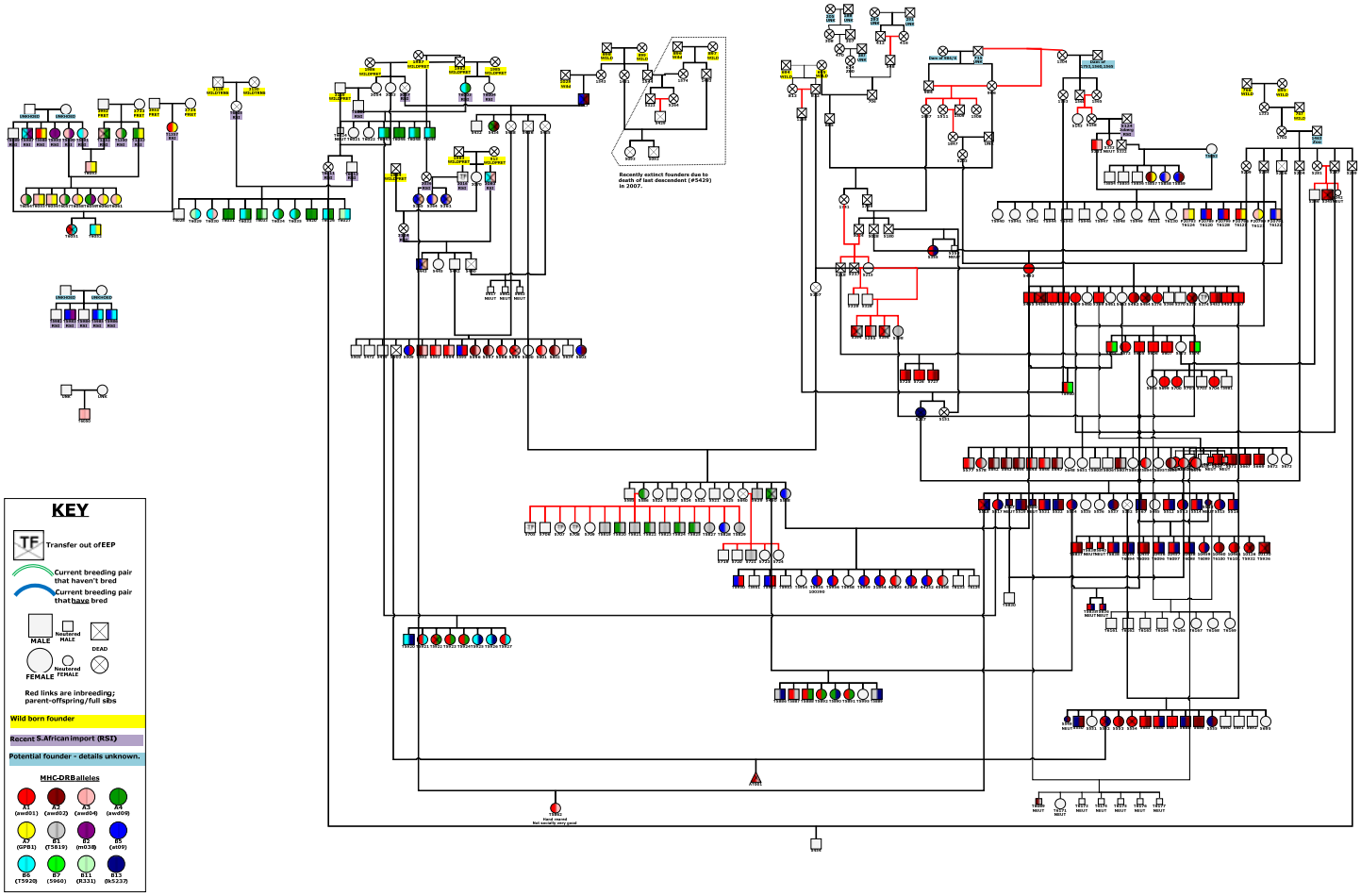
These recommendations should increase effective population size, reduce inbreeding and loss of genetic diversity. However, space restrictions, social breeding structure, and large litter sizes produce a combination of challenges that make loss of genetic diversity inevitable. Artificial reproductive technologies (e.g. artificial insemination) are being developed for wild dogs (Thomassen & Farstad 2009). Such methods may enable managers to exert more control over breeding in the future (Howard et al. 2003), as well as offering the opportunity to conserve reproductive material from genetically valuable individuals that could be killed during pack formation or be prevented from breeding due to space restrictions.

6.5.4 MHC data for captive management

Conservation of adaptive genetic diversity is critical to prevent the loss of the evolutionary and adaptive potential of species (Lacy 1997). However, the use of neutral markers as a proxy for adaptive diversity has been heavily criticized because neutral loci

are not under selection and therefore are unlikely to be well correlated with adaptive traits as found by (Bekessy et al. 2003; Ennos et al. 1997). For this reason, assessments of adaptive loci, such as the MHC, have been advocated as an alternative proxy for adaptive diversity, which was pioneered in endangered species by Hedrick (1998). However, as far as I am aware, a comparative analysis based on MHC and microsatellite markers has not been conducted for a mammalian captive breeding programme (except the Giant panda where MHC and microsatellite analyses were conducted on different sample sets (Zhu et al. 2007). In this study, genetic diversity metrics for captive populations were calculated based on both MHC and microsatellite markers for the same individuals. Whilst, there were some similar results between the two marker sets (e.g. H_o and F_{IS} for EUZ_pure, Table 6.5), there were a number of important differences (Table 6.5): 1) A_R and R_S were much higher for MHC than microsatellites; 2) microsatellites suggested higher diversity for EUZ than EUZ_pure, whereas the opposite pattern was found for the MHC; 3) the proportion of wild diversity represented in EUZ_pure was much higher for microsatellites (57.4%) than at the MHC (38.5%); 4) EUZ H_o, and SAZ F_{IS} were much higher for the MHC than microsatellites. Overall, these results suggest that MHC and neutral diversity metrics were not completely correlated in captive wild dogs. As such, these findings advocate the value of incorporating MHC markers into genetic assessments of captive breeding programmes. Nonetheless, the specific role of MHC data in genetic management has been the subject of much debate. Hughes (1994) proposed that captive breeding programmes should be managed to maintain MHC in preference to management of loci that are selectively neutral (e.g. microsatellites, mtDNA). However, this has been strongly opposed primarily due to concerns that conservation of a small number of loci may accelerate loss of diversity at other loci, as well as a lack of knowledge over the relative advantage of different alleles. In this study, mapping of MHC data onto the pedigree provided a very visually intuitive representation of patterns of diversity. Specifically, the MHC data was found to be an accurate indicator of both inbreeding and founder representation in the EUZ wild dog population. For example, the high frequency of DRB allele A1 reflected the over-representation of, and inbreeding amongst, founders from the left side of the pedigree, whereas 3/4 low frequency DRB alleles were found in individuals from under-represented founder lines. As such, these results advocate the potential for variable MHC loci as indicators of genetic status of captive breeding programmes. Overall, however, genetic management of captive breeding programmes will likely be strongest when combining multiple data types, that is, studbook information and data from neutral and adaptive loci, because each provides different types of information.

Figure 6.1: Current African EUZ wild dog pedigree; a full sized version of this figure has been attached as an appendix to the thesis. For clarity most individuals that died leaving no descendents have been removed. Individuals typed at the MHC have their alleles depicted by colours. Full details are in the key.



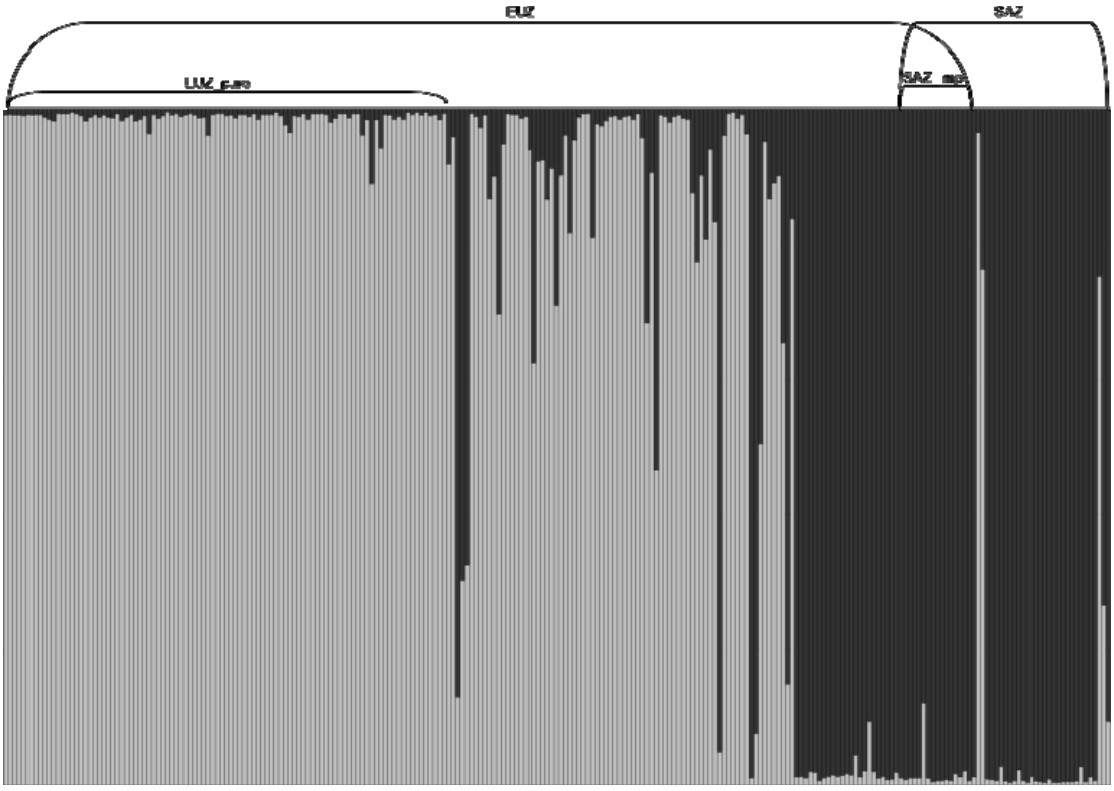


Figure 6.2: Genetic Structure of European (EUZ) and South African (SAZ) captive wild dogs using 9 microsatellite loci. Shown is the most likely level of population clustering (K=2) as indicated by the ΔK statistic. Columns are individuals, with the proportion of an individual's genotype assigned to each cluster (K) denoted by different colours. Admixture is indicated where both colours are well represented in an individual. Groupings are depicted as follows: EUZ_ pure = EUZ individuals with no ancestry to recent SAZ imports; EUZ = all individuals in EUZ including SAZ_imp and their descendents; SAZ_imp = animals imported from South African zoos; SAZ – individuals from South African zoos.

Table 6.1: List of samples contributed by different EUZ.

Contributing EUZ	Zoo code	No. samples
Aalborg zoo, Aalborg, Denmark	Aarlborg	9
Artis zoo, Amsterdam, Netherlands	Artis	4
Attica Zoological Park, Spata, Greece	Attica	2
Beekse Bergen Safarai Park, Hilvarenbeek, Netherlands	Hilvaren	11
Borås Djurpark Zoo, Alvsborg, Sweden	Boras	12
Bretagne Zoo Sarl, Pont-Scorff, France	PontScorff	13
Centre d'Etude Rech Zool Augeron, Lisieux, France	Lisieux	4
City of Belfast Zoo, Belfast, UK	Belfast	3
Colchester zoo, Essex, UK	Colchester	4
Ebeltoft zoo, Ebeltoft, Denmark	Ebletoft	5
Edinburgh Zoo, Edinburgh, UK	Edinburgh	6
Friguia Zoo, Hammam, Tunisia	Friguia	1
Fondazione Bioparco di Roma, Rome, Italy	Roma	3
GaiaPark Kerkrade Zoo, Kerkrade, Netherlands	Kerkrade	7
Kolmården Djurpark AB, Kolmården, Sweden	Kolmarden	20
La Palmyre Zoo, Royan, France	La Palmyre	3
Le Pal Parc Animalier, Dompierre-sur-Besbre, France	Le Pal	3
London zoo, London, UK	London	4
Munchener Tierpark Hellabrunn, Muenchen, Germany	Munich	7
Parken Zoo Eskilstuna AB, Sodermanland, Sweden	Eskilstun	2
Port Lymne Wild Animal Park, Hythe, UK	Lympne	28
Quinta De Santo Inacio, Avintes, Portugal	Quintas	1
Reserve Africaine De Sigean, Sigean, France	Sigean	10
Rostock Zoologischer Garten, Rostock, Germany	Rostock	5
Safari De Peaugres, Peaugres, France	Peaugres	7
Warsaw Zoological Garden, Warszawa, Poland	Warsaw	4
West Midland Safari & Leisure Park, Worcester, UK	Bewdley	20
Zoo Basel, Basel, Switzerland	Basel	5
Zoo Dortmund, Dormund, Germany	Dortmund	4
Zoo D'Amneville, Amneville, France	Amneville	6
Zoo Dvůr Králové, Dvůr Králové nad Labem, Czech Republic	Dvur Kralove	4
Zoo Duisburgh Ag, Duisburg, Germany	Duisburg	3
Zoological Center Tel Aviv, Ramat Gan, Israel	Ramat Gan	1
Zoological Society of Ireland-Dublin, Dublin, Ireland	Dublin	2
Total	34	223

Table 6.2: mtDNA haplotypes of the 17 females that have contributed to the extant EUZ population. Where the female did not have an ID #, they were referred to by their offspring ID#, specifically Dam of offspring ID#.

Studbook ID	Origin	mtDNA
#205	Unknown	Only male descendents
#293	Unknown	Only male descendents
#695	Wild born, Unknown	Only male descendents
#895	Wild born, Namibia	E2
#899	Wild born, Unknown	S2
#913	Wild born, Unknown	S2
#1985	Wild born, Etosha, Namibia	S2
#1986	Wild born, Etosha, Namibia	Z1
#1987	Wild born, Unknown	S2
#3130	Wild born, Transvaal, South Africa	S2
Dam of #984/6	Unknown	Z1
Dam of #T5853	Unknown	S2
#3733	SAZ, De Wildt	S2
#3734	SAZ, De Wildt	E2
Dam of #T5983	SAZ, Hoedspruit	S2
Dam of #T5987	SAZ, Hoedspruit	S2
Dam of #T5993	SAZ, Hoedspruit	Z1

Table 6.3: Sample sizes (n) and distribution of mtDNA haplotypes across wild African wild dog populations and in EUZ and SAZ.

		mtDNA haplotype						
	n	**S1**	**S2**	**S5**	**Z1**	**Z2**	**E1**	**E2**
Wild								
Kruger, South Africa	94	35	59					
NE Namibia	10		10					
Okavango, Botswana	90		13	5	4	6	59	3
Ghanzi, Botswana	1		1					
Lowveld, Zimbabwe	15		15					
Hwange, Zimbabwe	47		16		11	17	1	2
Captive								
EUZ	9		6		1			1
SAZ	5		3		1			1

Table 6.4: DRB1 allele frequencies for all EUZ wild dogs (EUZ); wild dogs of EUZ excluding recent S. African imports and their descendents (EUzoos_pure); wild dogs from South African captive facilities (SAZ) and wild populations in Southern Africa (WILD).

DRB allele	EUZ (n=211)	EUZ_pure (n=104)	SAZ (n=66)	WILD (n=185)
A1	32.2	50.0	8.3	5.8
A2*	12.1	19.7		7.1
A3	8.8	1.0	13.6	11.4
A4	8.3		17.4	13.8
A6			1.5	
A7*	3.3		5.3	0.5
B1	11.6	11.1		
B2	1.4		20.5	9.5
B3				0.3
B4			3.0	5.8
B5	6.2	2.4	4.5	6.1
B6	5.2		22.7	22.5
B7	0.7	1.4	2.3	5.6
B9				1.9
B11*	0.9		0.8	9.8
B13	9.2	14.4		

* Private MHC alleles in WILD populations: A2 = Okavango; A7, B11 = Kruger.

Table 6.5: Effective populations size (Ne) compared to census size (Npop), and microsatellite and MHC diversity metrics for EUZ, SAZ and WILD populations, where n = number of samples; AR = allelic richness; Rs = allelic richness standardised for sample size; % wild = % of wild diversity from Southern African populations represented; observed (Ho) and expected (He) heterozygosity; fixation index, FIS =1-Ho/He).

	Effective pop size		**Microsatellite diversity**						***MHC diversity***					
	N_e	N_{pop}/N_e	*n*	A_R	R_S	*% wild*	H_o:H_e	F_{IS}	*n*	A_R	R_S	*% wild*	H_o:H_e	F_{IS}
EUZ_pure	20.2 (17.4-23.5)	NA	98	5.0 (0.58)	4.2	57.4	0.77:0.70	-0.11 (0.03)	104	7	8.1	38.5	0.76:0.68	-0.12
EUZ	26.6 (24.5-28.5)	0.10	212	6.4 (0.8)	4.8	72.3	0.76:0.73	-0.05 (0.03)	211	12	5.1	76.9	0.83:0.84	+0.01
SAZ	15.6 (13.7-17-9)	NA	46	6.5 (0.67)	5.2	69.9	0.63:0.69	+0.04 (0.04)	66	11	9.2	76.9	0.68:0.85	+0.20
WILD[1]	7.8-19.6 (6.1-22.2)	0.05-0.10	169	4.6-7.4	4.1-5.7	NA	0.62-0.80: 0.59-0.76	-0.11 – +0.02	186	3-9	3.0-8.3	NA	0.67-0.84: 0.60-0.86	-0.08 - +0.05

[1] Wild data are based on data from Kruger, Okavango, Hwange and Lowveld populations (Table 4.3 and Table 4.5), NE Namibia (n=4) and Ghanzi (n=1) were excluded for these population level diversity metrics.

Table 6.6: Pairwise bootstrapped Dest estimates of microsatellite genetic differentiation between captive (EUZ_pure; SAZ) and free-ranging wild dog (Kruger, Okavango, Hwange, Lowveld) populations from Southern Africa.

--	EUZ_pure	SAZ	Kruger	Okavango	Hwange	Lowveld
EUZ_pure		0.320	0.284	0.259	0.218	0.483
SAZ			0.149	0.251	0.247	0.376
Kruger				0.319	0.277	0.338
Okavango					0.074	0.450
Hwange						0.383
Lowveld						

Chapter 7: General Discussion

7.1 Genetics in conservation

Genetic information and molecular tools have many potential applications in conservation. However, conservation genetics is an emerging and expensive field and many conservationists question how much genetics adds to conservation. One particular criticism has been that conservationists often fail to translate their data into 'practical conservation solutions' (Vernesi et al. 2008). In this respect, criticism is probably warranted. However, there are also conservation genetics success stories and it would be useful to identify common factors of these. In the conservation of canids, for example, molecular data has been used to show that genetic exchange has not occurred amongst subpopulations of the reintroduced Rocky Mountain Grey wolf, and as such the population did not meet the requirements for delisting and so were re-listed (Bergstrom et al. 2009). Another high profile conservation genetics success story was presented in Chapter 3. Here, genetic data provided new information on one of the most contentious population extinctions in conservation history; wild dogs in the Serengeti in 1991. The suggestion that animal handling during vaccination against a rabies outbreak caused the extinction resulted in restrictions in wildlife handling and the use of vaccination in wildlife and domestic species in a number of countries, despite the fact that the cause of the extinction could not be shown due to insufficient data (Woodroffe 2001). In 2001, wild dogs naturally re-colonised the Serengeti and I used molecular data to assess where the colonists came from. Contrary to expectations and current opinion, these results suggest that wild dogs did not go extinct in the Serengeti. Rather, they continued to cryptically reside in the monitoring area or persisted elsewhere in the Serengeti region. Essentially, it was the absence of knowledge rather than human intervention that resulted in the apparent extinction of wild dogs in the Serengeti. High profile case studies such as the Serengeti wild dogs and Rocky Mountain Grey wolves, are incredibly important in advocating the role and value of conservation genetics. In both of these case studies, very specific hypotheses were tested, and conservation genetics was able to provide novel data that could not have been elucidated with any other approach. These characteristics appear to have been central to their success.

Elucidating population structure is important for designating management units for translocation strategies and for understanding connectivity between populations and here also, genetics has been important for changing perspectives. To some extent, connectivity between populations can be elucidated by non-genetic methods, such as radio-collaring and assessments of habitat distributions. However, non-genetic methods typically assess migration, which is not the same as gene flow. For example, a study of bobcats and coyotes in Southern California showed that populations either side of a major freeway were highly differentiated, despite migration being observed across the freeway

(Riley et al. 2006). The reason was that the freeway caused a phenomenon called 'territory pile-up' which meant that although animals dispersed across the freeway, few manage to establish a territory and reproduce (Riley et al. 2006). In my study, I showed that Kruger in North Eastern South Africa was genetically isolated from South Eastern Zimbabwe, despite continuous habitat between these two areas, which had led to an assumption of connectivity by previous researchers. The reason for the genetic discontinuity remains to be shown, as no natural or human barriers to wild dog movement have been identified. However, elucidating the cause of the separation is important because genetic isolation puts Kruger at higher risk of genetic threats. Sampling of wild dogs in Northern Kruger, and Southern Zimbabwe is planned to elucidate the causal factors behind this genetic discontinuity (R. Groom). Overall, this example reiterates the importance of genetic studies in conservation, and more specifically it highlights the importance of spatial sampling; if samples had not be available from South Eastern Zimbabwe (i.e. Lowveld), the genetic isolation of Kruger would not have been revealed.

It is not only spatial sampling that is important, temporal sampling is also vital to track changes in status. Wild dogs have been described as a boom bust species (Woodroffe *Pers* comm) that are heralded for their ability to rapidly recover from demographic declines (Pole 2000; Woodroffe 2010). This ability appears to derive from their high fecundity and dispersal abilities; for example, the Laikipia population in Southern Kenya recovered from local extinction to >300 animals in less than 10 years (Woodroffe 2010). However, genetic analysis of populations that have experienced declines or extinction and recolonisation show that these demographic fluctuations are associated with reduced genetic diversity and increased inbreeding (Chapter 4). These data highlight that although wild dogs show demographic resilience to extinction, this is not matched by genetic resilience. As such, demographic fluctuations and population extinction and recolonisation events should be perceived as significant threats to wild dogs.

Overall, it is clear that, in general, conservation genetics must assign more priority to providing practical management suggestions. However, this must be reciprocated by willingness from policy makers to heed advice and implement solutions (Vernesi et al. 2008). My PhD was conducted in partnership with the Royal Zoological Society of Scotland, who provided funding to assess the genetic status of the EU zoo wild dog population, and subsequently to use these data to assist with the genetic management of this species. As a geneticist, the approach was simple; collecte the appropriate genetic data, and devise a genetic management plan for the zoos based on standard genetic principles. However, by working closely with the studbook manager, I got a valuable insight into the practical difficulties of implementing genetic management strategies. One of the biggest challenges for studbook managers is their lack of power. Essentially, if a zoo refuses to conform to requests there is not much the studbook manager can do.

Furthermore, wild dogs exist in some captive facilities in Europe that have not signed up to the European breeding programme (EEP), and are not part of the European Association of Zoos and Aquaria (EAZA) and thus they fall out of any jurisdiction. This would not present a problem if those wild dogs were kept separate. However, in reality, some of these animals are transferred into the breeding programme. It is not clear to me how the studbook manager can manage a population if they do not have jurisdiction over the population. Another challenge pertains to the fact that decisions by individual zoos are often taken without prior consultation with the studbook keeper. In recent years, wild dogs have been castrated, imported from South Africa, and sold to dealers, all without permission from the studbook manager. It is not generally helpful to point fingers of blame. However, I think it is important to raise these examples because the wild dogs that were sold or castrated were some of the most genetically valuable in the breeding programme, and therefore these actions have resulted in a large loss of potential genetic diversity. Furthermore, as discussed in Chapter 6, it is likely that importation of wild dogs from South Africa is fuelling an illegal trade of wild dog pup trafficking that is putting wild populations at risk.

A more general obstacle to genetic management comes from the fact zoos not only participate in captive breeding programmes but also have educational and research roles, and they must also generate sufficient revenue to sustain themselves. In any captive breeding programme, space restrictions limit the number of individuals that can breed, and amongst individuals, some will have a higher breeding priority than others. However, pups and cubs and chicks attract customers, so problems can arise when zoos with low breeding priority animals request/demand breeding groups. The studbook manager can control transfers of breeding priority animals between zoos; therefore, it might be possible to at least try to share breeding opportunities. But if species are problematic to breed, do you keep things ‘fair’? Or, do you send valuable animals to zoos with high success at breeding and in doing so risk that the zoos that are not given breeding opportunities refuse to cooperate further with the breeding programme? Whilst I understand why the zoos have multiple aims, education, research, captive breeding and generating revenue will not always be parsimonious aims. As such I think it is exceptionally challenging for zoos to achieve all of these aims well, and perhaps a consensus is needed across the zoo community on which of these has priority.

The last challenge to captive breeding relates to problems specific to the species. As discussed in Chapter 6, the social behaviour of wild dogs means that they must be held as single sex groups or as a breeding pack formed by the merging of two unrelated single sex groups. Formation of breeding packs of wild dogs in zoos is also associated with high aggression; approximately half of attempts to integrate male and female groups to form breeding packs fail due to aggression, which is often fatal (Verbokemoes *Pers*

comm.). However, aggression is also common in single sex groups and sometimes fatal amongst individuals particularly during the breeding season (despite the absence of the opposite sex) (Boutelle & Bertschinger 2010; van Heerden et al. 1996), which results in pressure on the studbook keeper to provide breeding recommendations for such groups.

Overall, it can be see that genetic management of captive breeding programmes is very challenging. Geneticists will always propose the ideal genetic management plan, and it is important to strive to achieve this. However, failing to consider factors such as logistics, financial costs, and behavioural factors, renders geneticists as guilty as studbook managers that ignore genetics. My initial genetic management plan for the European captive wild dog population was sound from a genetic perspective but unrealistic from a management perspective because I had failed to appreciate: 1) the severity of the aggression problem during pack formation; 2) the importance of keeping offspring with parents to learn pup rearing skills; 3) the degree to which breeding success differs between zoos; 4) the variance between zoos in their willingness to cooperate (most do, a few don't); 5) the value placed on the educational role of wild dogs in zoos. It would be naïve to expect the management committee to appreciate the importance of genetic threats, if I were to ignore the issues they raised. As such, together we devised a new genetic management plan that was feasible from a management perspective and acceptable from a genetics perspective. This did require some genetic compromises. However, the ideal genetic management plan would never have been implemented, and whilst not perfect, the new management plan will go a long way in improving the genetic status of the breeding programme. Overall, this example indicates that closer relationships between the conservation geneticists and managers would likely facilitate better integration of genetic data in conservation policy.

In terms of application of genetic data, there is one last issue that needs to be raised. There appears to be a big elephant in the room of conservation genetics that goes by the name of politics. Whilst this is not unique to the field of conservation genetics, indeed politics seems to thwart many charismatic species conservation programmes, it is a concern. Alvarez (1993) provided a damning review of the management of the Florida panther, California condor, black footed ferret and dusky seaside sparrow endangered species programmes which is summed up in this quote:

> *"And so it goes...the actors come and go; decisions are reversed, often without explanation; no one is in charge; the different agencies and factions pursue their separate objectives; motives are sometimes discernible and sometimes not; the recovery program is a case of strategic aversion and operational chaos, organized only to the extent that it can avoid any action deemed undesirable by its component factions, as they project an image of industry and purpose while consuming a perennial flow of revenue." (Alvarez 1993)*

Clearly, there are many obstacles to implementing genetic research into endangered species conservation, but the value of well-sampled and well-analysed genetic data collected in close collaboration with conservationists on the ground should go a long way towards ensuring that knowledge of genetics has the potential to inform management decisions.

7.2 Broader implications of this work

Loss of adaptive genetic variation is thought to be one of the greatest threats to the persistence of species because it may impair their ability to respond to changes in the abiotic and biotic environments (Lacy 1997). As one of the most well understood adaptive loci (Bernatchez & Landry 2003; Miller et al. 2001), assessments of the MHC are increasingly being conducted on endangered species to assess whether species have lost adaptive diversity. However, the results presented in Chapter 5 suggest that the relationship between MHC diversity and endangered status is complex. Overall, it was shown that patterns of MHC diversity in endangered species were not only affected by their demographic history, but also by selection and evolutionary history. The importance of evolutionary history as a factor influencing MHC diversity is intuitive but has rarely been considered when interpreting differences in MHC diversity between species, even if they are distantly related. I expect that evolutionary history (along with differences in selection pressures) may help to explain why some endangered species appear depauperate at the MHC, whereas others have considerably more variation than expected given their demographic history. The results from Chapter 5 have important implications for assessments of MHC diversity in endangered species where determining whether an endangered species has lost MHC diversity is based on comparisons of diversity with a reference taxon. Clearly, unless very closely related species are compared, it will be difficult to disentangle whether any differences in diversity between species are the result of population declines rather than evolutionary history.

Use of the MHC as a proxy for adaptive genetic variation has been advocated because neutral markers are predicted to be uncorrelated to adaptive genetic diversity. However, like many other studies of the MHC in endangered species (e.g. Campos et al. 2006), I found that MHC and neutral markers were often correlated, indicating that contemporary patterns of MHC variation are predominantly affected by neutral processes. Clearly this raises the question of whether assessments of MHC variation are worthwhile in endangered species? I personally think the answer to this question is yes. Firstly, because although there was a general correlation between MHC and neutral markers, this was not always the case, which suggests that selection is likely weak or variable, rather than absent. For example, patterns of genetic differentiation across wild populations indicated by MHC and microsatellites were significantly different. Secondly, in the case of the zoos, the MHC data was invaluable tool for presenting genetic data to a non-genetics audience. In particularly, mapping MHC alleles onto the pedigree enabled me to demonstrate a range of genetic concepts in a visually intuitive format, which would not have been possible with microsatellite data. Overall, it is likely that the strongest

approach would be to combine multiple marker types, because each is associated with different strengths and weakness and each provides different types of information.

7.3 Future work

Whilst this thesis addressed a number of questions with the MHC, these were approached from a genetic perspective. Given the importance of the MHC in immunity, the next logical step seems to be to use the MHC data towards applied disease questions. Infectious diseases have the potential to cause catastrophic population declines and are increasingly recognised as posing a significant threat to endangered species (Acevedo-Whitehouse 2009; Smith 2009). Disease epidemics have resulted in both population and extinctions of endangered species (reviewed in Daszak et al. 2000; Smith 2009), and newly emerged diseases such as the chytrid fungus pandemic in amphibians (Berger et al. 1998) and infectious facial tumour disease in Tasmanian devils (Hawkins et al. 2006) have resulted in species of little conservation concern transitioning to endangered status. Overall, the threat of diseases to wildlife species has increased as human populations and their accompanying domesticated animals have spread, resulting in spill-over infections between domesticated and wild animals (Randall et al. 2006), such as rabies virus between domestic dogs and Ethiopian wolves, and feline leukaemia virus between domestic cats and the Iberian lynx (López et al. 2009), which makes it a key conservation concern.

As for most canids, African wild dogs are highly susceptible to a number of diseases (Woodroffe et al. 2004a; Woodroffe & Ginsberg 1997) and epidemics have resulted in dramatic demographic impacts (Woodroffe & Ginsberg 1997). Rabies caused the death of 21/23 AWD pack members in the Masai Mara in 1989 (Kat et al. 1995), and was implicated in the loss of the pre-extinction Serengeti-Mara population in 1991 (Burrows et al. 1995; Ginsberg et al. 1995). Canine distemper virus (CDV) resulted in the death of 10/12 pack members in Chobe National Park, Botswana in the early 1990's (Creel & Creel 2002). Interestingly, though, wild dogs populations appear to show substantial variation in the extent to which they are influenced by disease. Research is currently being conducted to investigate potential ecological differences (e.g. proximity to domestic dogs), for these differences. However, genetic variation, such as differences at the MHC, could also be a causal factor underlying these patterns. It would be interesting to investigate this further. Given that many of the most severe disease threats to wild dogs are viral, it would be very worthwhile to investigate MHC class I variation, as well as genetic variation in disease receptors such as the SLAM gene, which is important for responses to CDV (McCarthy et al. 2007).

Vaccination strategies have been instigated successfully into conservation programmes for some endangered species. For example, in the Ethiopian wolf, *Canis simensis,* a low coverage targeted vaccination strategy successfully prevented the spread

of rabies between subpopulations (Haydon et al. 2006). However, the protection afforded by vaccination is dependent on an efficacious response to the vaccination and non-responders present a problem to many vaccination programmes (for example, measles in humans (Hayney et al. 1998). Vaccination programmes have been implemented on a number of occasions in wild dogs. Unfortunately, wild dogs appear to exhibit extreme variation in vaccination response, with a large proportion failing to seroconvert. For example, a CDV outbreak in a captive breeding facility in Tanzania resulted in the death of 49/52 of the dogs, all of which had been vaccinated (van de Bildt et al. 2002). Similarly, wild dogs have died despite rabies vaccination on a number of other occasions (reviewed in Creel & Creel 2002). The cause for this variation in vaccination response is currently unknown, but must be ascertained if vaccinations programmes are to be implemented effectively. A number of factors are known to influence post-vaccination antibody levels. For example, a recent study of antibody responses to rabies in >10,000 domestic dogs showed that animal size, breed, sampling time and vaccine brand significantly influenced whether antibody responses were efficacious or not (Kennedy et al. 2007b). Less understood is the contribution of genetics to variation in response to vaccination. In light of MHC disease associations, and the similarity in the immunological response to vaccination and that of normal infection, it is possible that variation in vaccination response may, in part, be explained by the MHC. That is, like disease resistance, efficacious vaccination response may require specific MHC alleles or heterozygosity (Hayney et al. 1998). The influence of genetics on vaccination response has been best studied in humans. Tan (2001) conducted a comparative study on monozygotic and dizygotic twins and found that the response to measles vaccination had a large genetic component (heritability - 88.5%). Furthermore, non response to single-dose measles vaccination has been associated with both levels of homozygosity and specific alleles at the MHC (Hayney et al. 1998). More recent research on MHC measles vaccinations have shown that, although non-response to measles vaccination is associated with the MHC, an efficacious response can normally be achieved with a second dose of vaccine dose (St. Sauver et al. 2005). Consequently, two doses of vaccine are recommended to afford protection, regardless of MHC type (St. Sauver et al. 2005). This research highlights the potential contribution of MHC studies in increasing vaccination efficacy but that careful research would be required. The use of vaccination in free-ranging wild dogs remains contentious. However, assessments of antibody responses to vaccination in captive animals could be an alternative. In this situation, it would be more feasible to assess additional questions relating to heritability of vaccine response.

Another area requiring further research concerns understanding the causes of aggression between wild dogs in captivity, as this greatly restricts management in the zoos. There has been interest as to whether this may in part reflect ‘MHC-related mate incompatibility’. Unfortunately, this would be difficult to test because each zoo

approaches introductions in different ways, and in most cases the details and outcomes of the introductions are not documented. Whilst it is possible that the MHC may be involved, a thorough behavioural study investigating husbandry approaches would be the best starting point to approaching this problem.

Appendices

Appendix 1: Official DLA-DRB1 and DLA-DQB1 names for wild dog MHC alleles. An abbreviated name is provided for the DRB alleles; the letter refers to whether the allele is derived from either lineage A or B. Also listed are local names which were used in original files.

Official name	Abbreviated name	Local name
DRB1*90101	A1	awd01
DRB1*90102	A2	awd02
DRB1*90201	A3	awd04
DRB1*90202	A4	awd09
DRB1*90203	A5	at04
DRB1*90204	A6	awd04v
DRB1*90301	A7	GPB1
DRB1*90205	A8	STAR
DRB1*90401	B1	T5819/at09v
DRB1*90402	B2	m038
DRB1*90501	B3	awd34
DRB1*90601	B4	awd08
DRB1*90602	B5	at09
DRB1*907011	B6	T5920/bru40
DRB1*90801	B7	5960
DRB1*91101	B8	K2MozNa
DRB1*91001	B9	R504
DRB1*90403	B10	DISH
DRB1*90901	B11	R331
DRB1*907012	B12	lk5237v
DRB1*90702	B13	lk5237
DQB1*90102		awd14
DQB1*90201		awd16

Appendix 2: Details of monitoring populations in Chapter 4

The Kruger samples were derived from a monitoring area in the Southern section of Kruger National Park in North Eastern South Africa [31º6'E, -25º1'N]. Samples were collected at two different time periods, 1990-1995 (Gus Mills) and 2007 (Janet Edwards, Paulette Bloomer, Harriet Mostert-Davies). Population size in the monitoring area was ~400 (1990-1995) and ~100 (2007).

The Lowveld is located in South Eastern Zimbabwe [32º1'E, -20º8'N], and the monitoring area encompassed both the Savé Valley Conservancy monitoring population (3487km^2) and Malilangwe Wildlife Reserve (384 km^2), which is an adjacent, virtually contiguous protected area. The Lowveld samples were collected in 2008-2009 (Rosemary Groom).

The Hwange monitoring area includes Hwange National Park and Nyamandlovu (an area of land that is contiguous with Hwange national Park and part of the Hwange ecosystem) in North Western Zimbabwe [26º9'E, -18º8'N]. Hwange has been influenced by artifical translocations of wild dogs from South Africa and other areas of Zimbabwe. Therefore, I only classified animals as Hwange, where I were confident they were of pure Hwange origin. I collected 19 samples from Hwange; 10 samples from Hwange National Park (n=5, unknown packs, sampled prior to 1993 and n=5 from 3 packs, sampled 2006-7), and 9 samples from Nyamandlovu (n=9, unknown packs, 1995-96) (Joshua Ginsberg, Greg Rasmussen). Sample sizes were too small for a temporal comparison, therefore these samples were pooled into a single data set.

The Okavango encompasses the Okavango delta area including the Moremi Game Reserve in Northern Botswana [23º4'E, -19º3'N]. Samples were collected at two different time periods, 1991-1993 and 2000-2007 (Tico McNutt). The five migrant animals with the new S5 (T1) haplotype had the following local ID's; Mank [(T)06-197], Warne [(T)04-188], Nino [(T)02-169; [(T)03-176] Lyra [(T)02-152], Cygnus [(T)01-149].

The Selous sampling area was the Northern parts of the Selous game reserve (43,600 km^2) [38º1'E, -7º6'N]. Samples were collected between 1991-1995 (Scott Creel). The three migrant animals with the new S4 (C3) haplotype had the following local ID's; Pluto, Thor and Neptune.

The Masai Steppe samples were derived from a group of three litters of wild born pups that were taken from three dens in the Masai Steppe area in the 1980's [37º0'E, -4º5'N] and subsequently moved to a captive facility at the Mkomazai Game Reserve.

Specifically, I were provided with DNA from 18 wild born pups, and 15 samples from the descendents of a further 5 wild born pups for which DNA samples were not available (Aart Visee).

The Serengeti-Mara is located on the border of Southern Kenya and Northern Tanzania. Samples were collected from wild dog packs residing in or near the Serengeti National Park (SNP), Masai Mara Nature Reserve (MMNR), Ngorongoro Conservation Area and Loliondo Game Controlled Area, [35º1'E, -2º3'N]. Samples were collected at two time periods; Old, n = 20, 1981-1990 (Sarah Cleaveland, Pieter Kat); Recent, 2005 and 2007 (Emmanuel Masenga). For more details of Serengeti-Mara see Chapter 3.

The Laikipia samples were derived from a monitoring area in the Laikipia district and parts of the Samburu, Isiolo and Baringo Districts in Kenya [37º2'E, 0º6'N]. Samples were collected 2003-2008 (Rosie Woodroffe).

Monitoring areas with ≤5 samples were collected in the following years; Kajiado district 2001 (Rosie Woodroffe), NE-Namibia pre-1990 (Flip Stander) and 2006-2007 (Robin Lines), Niassa National Reserve 2008 (Colleen Begg; 35º3'E, -18º2'N), Northern Sofala Province 2004 (35º4'E, -18º1'N; Jean-Marc André), Ghanzi district Western Botswana 2006 (Mark Swarner).

Appendix 3: Information on microsatellite primers used in this study.

Primer Name	Sequence (5'---3')	Chromosome	Approx # alleles	Label and primer mix
FH2611F	GAAGCCTATGAGCCAGATCA	36	20	Ned Multiplex 1
FH2611R	TGTTAGATGATGCCTTCCTTCT	36		
PEZ12F	GTAGATTAGATCTCAGGCAG	3	7	Ned Multiplex 1
PEZ12R	TAGGTCCTGGTAGGGTGTGG	3		
PEZ08F	TATCGACTTTATCACTGTGG	17	10	6-Fam Multiplex 2
PEZ08R	ATGGAGCCTCATGTCTCATC	17		
FH2785F	ATGGCAGGTCAAGAGTATGG	28	12	6-Fam Multiplex 2
FH2785R	GATAGATCCAAGCCAACACC	28		
FH3965F	GTCGCTCAGCAGTTAAGCTC	02	20	6-FAM Multiplex 1
FH3965R	GAATCCTGGCTCTGCTACTTAC	02		
FH2054F	GCCTTATTCATTGCAGTTAGGG	12	7	Hex Multiplex 1
FH2054R	ATGCTGAGTTTTGAACTTTCC	12		
FH2658F	TCTTAGAAATTGCTGGTGGG	14	11	Hex Multiplex 1
FH2658R	TAAGAAACTGCCAGTCTGTGG	14		
FH3399F	TCTCTATGCCTGCAGTTTCC	38	32	Hex Multiplex 1
FH3399R	TTCTGATGCCCTCATAAAGC	38		
FH2010F	AAATGGAACAGTTGAGCATGC	24	5	Ned Multiplex 2
FH2010R	CCCCTTACAGCTTCATTTTCC	24		
PEZ15F	CTGGGGCTTAACTCCAAGTTC	05	9	Hex Multiplex 2
PEZ15R	CAGTACAGAGTCTGCTTATC	05		

Appendix 4: Pairwise bootstrapped Dest (Jost 2008) estimates of genetic differentiation between African wild dog populaitons. DRB data above the diagonal, msat below the diagonal. Values <0.2 are underlined. O and R depict Old and Recent samples for those populations.

	KruO	***KruR***	***OkaO***	***OkaR***	***Hwa***	***Low***	***Sel***	***MaSt***	***SerO***	***SerR***	***Lai***
KrugerO		0.06	0.88	0.82	0.85	0.46	0.48	0.83	0.99	0.85	0.59
KrugerR	0.03		0.93	0.89	0.91	0.26	0.47	0.78	0.99	0.91	0.74
OkavangoO	0.36	0.27		0.07	0.42	0.63	0.83	0.75	0.78	0.95	0.85
OkavangoR	0.32	0.26	0.02		0.54	0.90	0.78	0.54	0.61	0.88	0.72
Hwange	0.27	0.26	0.09	0.06		0.60	0.82	0.44	0.64	0.65	0.69
LoweId	0.37	0.28	0.46	0.39	0.35		0.61	0.82	1.00	1.00	1.00
Selous	0.37	0.33	0.51	0.45	0.36	0.45		0.51	0.77	0.44	0.67
Masai	0.32	0.36	0.32	0.30	0.36	0.55	0.43		0.28	0.53	0.72
SerengetiO	0.30	0.36	0.37	0.28	0.27	0.35	0.23	0.28		0.63	0.76
SerengetiR	0.24	0.23	0.21	0.25	0.20	0.46	0.47	0.27	0.14		0.75
Laikipia	0.39	0.46	0.37	0.30	0.48	0.50	0.33	0.36	0.25	0.28	

Appendix 5: Pairwise bootstrapped F_{st} estimates of genetic differentiation between African wild dog populaitons. DRB data above the diagonal, msat below the diagonal. Values <0.2 are underlined. O and R after three letter codes, depict Old and Recent samples for those populations.

	KruO	***KruR***	***OkaO***	***OkaR***	***Hwa***	***Low***	***Sel***	***MaSt***	***SerO***	***SerR***	***Lai***
KrugerO		0.02	0.18	0.17	0.20	0.17	0.10	0.21	0.25	0.21	0.23
KrugerR	0.03		0.23	0.21	0.26	0.13	0.12	0.25	0.30	0.28	0.32
OkavangoO	0.17	0.15		0.01	0.08	0.19	0.11	0.15	0.16	0.18	0.28
OkavangoR	0.14	0.12	0.03		0.10	0.24	0.11	0.11	0.13	0.17	0.24
Hwange	0.12	0.11	0.06	0.04		0.21	0.14	0.12	0.17	0.16	0.26
LoweId	0.17	0.19	0.23	0.17	0.17		0.17	0.28	0.33	0.33	0.42
Selous	0.16	0.16	0.19	0.17	0.15	0.23		0.10	0.15	0.09	0.22
Masai	0.18	0.19	0.19	0.16	0.16	0.29	0.20		0.09	0.15	0.29
SerengetiO	0.13	0.16	0.14	0.11	0.09	0.18	0.13	0.16		0.18	0.30
SerengetiR	0.14	0.14	0.11	0.10	0.09	0.23	0.19	0.16	0.09		0.30
Laikipia	0.18	0.21	0.16	0.14	0.17	0.23	0.15	0.18	0.13	0.16	

Appendix 6: mtDNA haplotypes. Listed first are the original sequences from Girman et al. 2001, followed by the corrected/longer sequences identified in Chapter 4.

Original Girman et al. 2001 mtDNA haplotypes (AF335724-AF335731).

>E1-Girman
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCCACCCTCTTTTCTCCCCT
ATGTACGTCGTGCATTAATGGCTTGCCCCATGCATATAAGCATGTACATGATATTATATTCTTACATAGGACATATCTACT
TAATCTCACAATCTCATTGACCTACAGCAGCAATGAAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCCTC
GAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATACCAATGTGGGGG

>E3-Girman
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCCACCCTCTTTTCTCCCCT
ATGTACGTCGTGCATTAATGGCTTGCCCCATGCATATAAGCATGTACATGATATTATATTCTTACATAGGACATACCTGC
TTAACCTCACAATCTCATTGACCTACAACAGCAATGAAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCCT
CGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATACCAATGTGGGGG

>E2-Girman
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCCACCCTCTTTTCTCCCCT
ATGTACGTCGTGCATTAATGGCTTGCCCCATGCATATAAGCATGTACATGATATTATATTCTTACATAGGACATACCTACT
TAACCTCACAATCTCATTGACCTACAACAGCAATGAAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCCTC
GAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATACCAATGTGGGGG

>Z2-Girman
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCCACCCTCTTTTCTCCCCT
ATGTACGTCGTGCATTAGTGACTTGCCCCATGCATATAAGCATGTACATAGTATTATACTCTTACATAGGACATACCTAC
TTAGTCTCACAATCTCATTAACCTACAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCCT
CGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGGGGG

>Z1-Girman
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCCACCCTCTTTTCTCCCCT
ATGTACGTCGTGCATTAATGACTTGCCCCATGCATATAAGCATGTACATAGTATTATACTCTTACATAGGACATACCTACT
TAGTCTCACAATCTCATTAACCTACAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCCTC
GAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGGGGG

>S2-Girman
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCCACCCTCTTTTCTCCCCT
ATGTACGTCGTGCATTAATGACTTGCCCCATGCATATAAGCATGTACATAGTATTATACTCTTACATAGGACATATCTACT
TAGTCTCACAATCTCATTAACCTATAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCCTC
GAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGGGGG

>S1-Girman
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCCACCCTCTTTTCTCCCCT
ATGTACGTCGTGCATTAATGACTTGCCCCATGCATATAAGCATGTACATAATATTATACTCTTACATAGGACATATCTACT
TAGTCTCACAATCTCATTAACCTATAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCCTC
GAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGGGGG

>S3-Girman
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCCACCCTCTTTTCTCCCCT
ATGTACGTCGTGCATTAATGACTTACCCCATGCATATAAGCATGTACATAATATTATACTCTTACATAGGACATATCTACT
TAGTCTCACAATCTCATTAACCTATAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCCTC
GAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGGGGG

Corrected mtDNA haplotype sequences and new sequences identified in Chapter 4.. Sequencing errors were detected for E1, E2, E3. The two new haplotypes identified were S4 and S5. All of the sequences have been extended by a further 21 bp.

>E1-Marsden (aka K1)
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCCCATCCTCTTTTCTCCCC
TATGTACGTCGTGCATTAATGGCTTGCCCCATGCATATAAGCATGTACATGATATTATATTCTTACATAGGACATATCTAC
TTAATCTCACAATCTCATTGACCTACAGCAGCAATGAAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCCT
CGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATACCAATGTGGGGGTCGCTATAACGG
AACTATAC
>E2-Marsden (aka M1)
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTCACTATGCCACGCCGGCACCCCATCCTCTTTTCTCCCC
TATGTACGTCGTGCATTAATGGCTTGCCCCATGCATATAAGCATGTACATGATATTATATTCTTACATAGGACATACCTAC
TTAACCTCACAATCTCATTGACCTACAACAGCAATGAAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCCT
CGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATACCAATGTGGGGGTCGCTATAATGGA
ACTATAC

>E3-Marsden (aka K2)
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTCACTATGCCACGTCGGCACCCCATCCTCTTTTCTCCCC
TATGTACGTCGTGCATTAATGGCTTGCCCCATGCATATAAGCATGTACATGATATTATATTCTTACATAGGACATACCTG
CTTAACCTCACAATCTCATTGACCTACAACAGCAATGAAATGCATATCACCTAGTCCAATAAGGGATTAATCACCATGCC
TCGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATACCAATGTGGGGGTCGCTATAATGG
AACTATAC

>Z2-Marsden
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCC-
ACCCTCTTTTCTCCCCTATGTACGTCGTGCATTAGTGACTTGCCCCATGCATATAAGCATGTACATAGTATTATACTCTTA
CATAGGACATACCTACTTAGTCTCACAATCTCATTAACCTACAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGG
ATTAATCACCATGCCTCGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGG
GGGTCGCTACAATGGAACTATAC

>Z1-Marsden
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCC-
ACCCTCTTTTCTCCCCTATGTACGTCGTGCATTAATGACTTGCCCCATGCATATAAGCATGTACATAGTATTATACTCTTA
CATAGGACATACCTACTTAGTCTCACAATCTCATTAACCTACAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGG
ATTAATCACCATGCCTCGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGG
GGGTCGCTACAATGGAACTATAC

>S2-Marsden
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCC-
ACCCTCTTTTCTCCCCTATGTACGTCGTGCATTAATGACTTGCCCCATGCATATAAGCATGTACATAGTATTATACTCTTA
CATAGGACATATCTACTTAGTCTCACAATCTCATTAACCTATAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGG
ATTAATCACCATGCCTCGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGG
GGGTCGCTACAATGAAACTATAC

>S1-Marsden
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCC-
ACCCTCTTTTCTCCCCTATGTACGTCGTGCATTAATGACTTGCCCCATGCATATAAGCATGTACATAATATTATACTCTTA
CATAGGACATATCTACTTAGTCTCACAATCTCATTAACCTATAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGG
ATTAATCACCATGCCTCGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGG
GGGTCGCTACAATGAAACTATAC

>S3-Marsden
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCC-
ACCCTCTTTTCTCCCCTATGTACGTCGTGCATTAATGACTTACCCCATGCATATAAGCATGTACATAATATTATACTCTTA
CATAGGACATATCTACTTAGTCTCACAATCTCATTAACCTATAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGG
ATTAATCACCATGCCTCGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGG
GGGTCGCTACAATGAAACTATAC

>S5-Marsden (aka T1) NEW
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCC-
ACCCTCTTTTCTCCCCTATGTACGTCGTGCATTAATGACTTGCCCCATGCATATAAGCATGTACATAGTATTATACTCTTA
CATAGGACATATCTACTTAGTCTCACAATCTCATTAACCTACAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGG
ATTAATCACCATGCCTCGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGG
GGGTCGCTACAATGAAACTATAC

>S4-Marsden (aka C1) NEW
CTATTCCCTGATCTCCCCCATATTCACATATTGAGTCAACCTTACTATGCCACGTCGGCACCC-
ACCCTCTTTTCTCCCCTATGTACGTCGTGCATTAATGACTTACCCCATGCATATAAGCATGTACATAATATTATACTCTTA
CATAGGACATATCTACTTAGTCTCACAATCTCATTAACCTACAACAGCAATGGAATGCATATCACCTAGTCCAATAAGGG
ATTAATCACCATGCCTCGAGAAACCATCAATCCTTGCTCGTAATGTCCCTCTTCTCGCTCCGGGCCCATATTAATGTGG
GGGTCGCTACAATGAAACTATAC

Appendix 7 African wild dog MHC class II sequences identified in this PhD. Shown are the official nomenclature names, with local names in brackets where applicable.

DLA-DQA1:
>DQA1-01901
GACCATGTTGCCAACTACGGCATAAATGTCTACCAGTCTTACGGTCCCTCTGGCCAGTTCACCCATGAATTTGATGGCG
ATGAGGAGTTCTATGTGGACCTGGAGAAGAAGGAAACTGTCTGGCGGCTGCCTGTGTTTAGCACATTTAGAAGTTTTGA
CCCACAGGGTGCACTGAGAAACTTGGCTATAATAAAACAAAACTTGAACATCCTGACTAAAAGGTCCAaCCAAAcTGCtG
CTaCCAaT

DLA-DQB1:
>dqb-90101(awd14)
gATTtCGTgTaCcAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGCTTCTGACTAAACACATCTA
TAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCGAC
GCTGAGTACTGGAACCGGCAGAAGGACGAGGTGGACCGGGTACGGGCCGAGGTGGACACGGTGTGCAGACACAACT
ACGGGATGGAGGAGCTCACCACGTTGCAGCGGCGA
>dqb-90201 (awd16)
gATTtCgTGTaCcAGTTTAaGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTCGTGGACAGATACATCTA
TAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCGAC
GCTGAGTACTGGAACCGGCAGAAGGACGAGGTGGACCGGGTACGGGCCGAGGTGGACACGGTGTGCAGACACAACT
ACGGGATGGAGGAGCTCACCACGTTGCAgCGGCGA

DLA-DRB1:
>DRB-90101 (AWD01)
CATTTCTTGAACGTGGCAAAGTCCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTCGTGGACAGATACATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACCTGAACCGGCAGAAGGAGATCTTGGAGCAGGAGCGGGCCGCGGTGGACACCTACTGCAGACACAAC
TACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90102 (AWD02)
CATTTCTTGAACGTGGCAAAGTCCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTCGTGGACAGATACATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACCTGAACCGGCAGAAGGAGATCTTGGAGCAGGAGCGGGCCGCGGTGGACACCTACTGCAGACACAAC
TACGGGGTGATTGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90201 (AWD04)
CATTTCTTGAACGTGGCAAAGTCCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTCGTGGACAGATACATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACTGGAACCGGCAGAAGGAGATCTTGGAGCAGGAGCGGGCCGCGGTGGACACCTACTGCAGACACAA
CTACGGGGTGATTGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90202 (AWD09)
CATTTCTTGAACGTGGCAAAGTCCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTCGTGGACAGATACATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACTGGAACCGGCAGAAGGAGATCTTGGAGCAGGAGCGGGCCGCGGTGGACACCTACTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90203 (AT04)
CATTTCTTGAACGTGGCAAAGTCCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTCGTGGACAGATACATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACTGGAACCGGCAGAAGGAGATCTTGGAGCAGGAGCGGGCCGCGGTGGACACGGTGTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90204 (AWD04V)
CATTTCTTGAACGTGGCAAAGTCCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTCGTGGACAGATACATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACTGGAACCGGCAGAAGGAGATCTTGGAGCAGGAGCGGGCCGCGGTGGACACCTACTGCAGACACAA
CTACGGGGTGATTGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90205(STAR)
CACATTTCTTGAACGTGGCAAAGTCCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTCGTGGACAGATACA
TCTATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCC
CGACGCTGAGTACTGGAACCGGCAGAAGGAGATCTTGGAGCAGGAGCGGGCCGCGGTGGACACGGTGTGCAGACAC
AACTACGGGGTGATTGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90301 (GPB1)
CATTTCTTGAACGTGGCAAAGTCCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTCGTGGACAGATACATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACTGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGAGGTGGACACGGTGTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90401 (T5819/at09v)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACTGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGAGGTGGACACGGTGTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90402 (M038)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACTGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGAGGTGGACACCTACTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90403 (DISH)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACTGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGAGGTGGACACGGTGTGCAGACACAA
CTACGGGGTGATTGAGAGCTTCACGGTGCAGCGGCGA

>DRB-90501 (AWD34)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGTTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACCGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGAGGTGGACACCTACTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90601(AWD08)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGCTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACTGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGAGGTGGACACGGTGTGCAGACACAA
CTACGGGGTGATTGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90602 (AT09)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGCTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACTGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGAGGTGGACACGGTGTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-907011 (T5920/bru40)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGCTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACCGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGCGGTGGACACCTACTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-907012 (lk5237V2)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACAGAGCGGGTGCGGCTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACCGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGCGGTGGACACCTACTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90702 (LK5237)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGCTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACCGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGCGGTGGACACCTACTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90801(5960)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGCTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACCTGAACCGGCAGAAGGAGATCTTGGAGCAGGAGCGGGCCGCGGTGGACACCTACTGCAGACACAAC
TACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-90901 (R331)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGCTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACCGGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGAGGTGGACACCTACTGCAGACACAA
CTACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-91001 (R504)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGCTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACCTGAACCGGCAGAAGGAGCTCTTGGAGCAGAGGCGGGCCGAGGTGGACACCTACTGCAGACACAAC
TACGGGGTGGGCGAGAGCTTCACGGTGCAGCGGCGA
>DRB-91101 (K2MOZNA)
CATTTCGTGTACCAGTTTAAGGGCGAGTGCTATTTCACCAACGGGACGGAGCGGGTGCGGCTTCTGGCGAGAAGCATC
TATAACCGGGAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTCACGGAGCTCGGGCGGCCCG
ACGCTGAGTACCTGAACCGGCAGAAGGAGATCTTGGAGCAGGAGCGGGCCGCGGTGGACACGGTGTGCAGACACAA
CTACGGGGTGATTGAGAGCTTCACGGTGCAGCGGCGA

Appendix 8: Donator ID and my corresponding lab ID for Lycaon pictus samples collated for this PhD (N.B. not all samples were used/useable), plus full details of EU zoo samples.

South Africa - Kruger (Old) National Park Lycaon pictus samples from Southern section: Donated by Gus Mills via RK Wayne (Girman study)

Donator sample ID	Clare's lab ID	Donator sample ID	Clare's lab ID
PF7	RKW 274	DM22	RKW 470
AF2	RKW 283	DM25	RKW 471
AF6	RKW 284	DM24	RKW 472
SKAF	RKW 287	DM23	RKW 473
SKAF	RKW 288	DF27	RKW 474
SRM5	RKW 289	DF31	RKW 475
AF5	RKW 290	GF1	RKW 476
SKAF	RKW 295	DF1	RKW 487
DPNK9	RKW 297	DN-M1,	RKW 5301
NM1	RKW 306	SF-7,	RKW 5302
TM2,	RKW 310	SM-10	RKW 5303
SRF2	RKW 314	SM-3	RKW 5304
Lpi	RKW 321	WD7	RKW 5305
FM5	RKW 323	DF18	RKW 550
DPNK2	RKW 330	MF23	RKW 564
AM12	RKW 331	MF22	RKW 565
SRM3	RKW 337	DF9	RKW 566
MF7	RKW 338	DM15	RKW 567
MM21	RKW 342	DM19	RKW 568
SKAF	RKW 344	DF16	RKW 569
DPNK8	RKW 345	DM16	RKW 570
SM12	RKW 346	DF22	RKW 571
SRM4	RKW 347		
MF20	RKW 421		
AF20	RKW 422		
AM10	RKW 423		
PM12	RKW 425		
TF1	RKW 427		
TM3	RKW 428		
DM21	RKW 429		
SF13	RKW 431		
PM11	RKW 432		
DM20	RKW 433		
ZM3	RKW 442		
PM10	RKW 452		
DF10	RKW 453		
DF20	RKW 454		
ZM2	RKW 455		
DF13	RKW 456		
MM10	RKW 458		
AF8	RKW 459		
ZM1	RKW 460		
AM8	RKW 461		
DF11	RKW 462		
NF2	RKW 463		
SM13	RKW 464		
RM1	RKW 466		
EM1	RKW 467		
MF10	RKW 468		
PF12,	RKW 469		

South Africa - Kruger (Recent) National Park Lycaon pictus samples, from Southern section
Donated by P.Bloomer, H.Mostert-Davies, J.Edwards.

Donator sample ID	Clare's lab ID
VL02074	K2074
VL02075	K2075
VL02076	K2076
VL02077	K2077
VLO2078	K2078
VL02083	K2083
VL02087	K2087
VL02088	K2088
VL02089	K2089
VL02090	K2090
VL02093	K2093
VL02094	K2094
VL02095	K2095
VL02096	K2096
VL02138	K2138
VL02139	K2139
VL02140	K2140
VL02141	K2141
VL02142	K2142
VL02143	K2143
VL02144	K2144
VL02145	K2145
VL02146	K2146
VL02147	K2147

Botswana – Okavango (Old) Lycaon pictus samples
Donated by Tico McNutt via RK Wayne (Girman study)

Donator sample ID	Clare's lab ID
Bot41 Negro	RKW 496
Bot42 Leah	RKW 497
Bot44 PACT	RKW 498
Bot45 FETCH	RKW 499
Bot46 Cider	RKW 500
Bot47 Schooner	RKW 501
Bot48 Tag	RKW 502
Bot49	RKW 503
Bot50 Twilight	RKW 504
Bot51 Isaac	RKW 505
Bot53 LUNA	RKW 506
Bot54 Bullseye	RKW 507
Bot55 Carley (Carliog??)	RKW 508
Bot57 Polyspock	RKW 509
Bot59 Rugby	RKW 510
Bot60	RKW 511
Bot61	RKW 512
Bot62	RKW 513
Bot63	RKW 514
Tinman	RKW 6209

Botswana – Okavango (Recentycaon pictus samples
Donated by Tico McNutt

Donator sample ID	Other sample ID/tube label	Clare's lab ID
Ander	Lpi 144	T00-144
Moonstone	LPI 172	T00-172
Scarp	LPI 173	T00-173
Ts	LPI 01-143	T01-143
Dundee	LPI 01-144	T01-144
Meg	LPI 01-145	T01-145
	LPI 01-146	T01-146
Rap*	LPI 01-148	T01-148
Cygnus	LPI 01-149	T01-149
Hewitt	LPI 01-150	T01-150
	LPI 02-151	T02-151
Lyra	LPI 02-152	T02-152
Soya	LPI 02-153	T02-153
Neve	LPI 02-154	T02-154
Safer	LPI 02-155	T02-155
Pooni*	LPI 02-156	T02-156
Hinich	LPI 02-157	T02-157
Chacha	LPI 02-158	T02-158
Dombek	LPI 02-159	T02-159
Lady	LPI 02-160	T02-160
	LPI 02-161	T02-161
Ponni*	LPI 02-162	T02-162
Njooki	LPI 02-164	T02-164
Minnie*	LPI 02-165	T02-165
Holmes	LPI 02-166	T02-166
Agate*	LPI 02-167	T02-167
Nino	LPI 02-169	T02-169
Manka	LPI 02-170	T02-170
Todaz	LPI 03-171	T03-171
	LPI 03-174	T03-174
Sharpa	LPI 03-175	T03-175
Nino	LPI 03-176	T03-176
Rakuka	LPI 03-177	T03-177
Rap*	LPI 03-178	T03-178
Costello*	LPI 03-179	T03-179
Butterfly	LPI 04-182	T04-182
Minnie*	LPI 04-183	T04-183
Euph	LPI 04-184	T04-184
Gangs	LPI 04-185	T04-185
Agate*	LPI 04-186	T04-186
Matopi	LPI 04-187	T04-187
Warne	LPI 04-188	T04-188
Hinkley	LPI 05-189	T05-189
Mathews	LPI 05-190	T05-190
Punter	LPI 191-05	T05-191
Costello*	LPI 05-192	T05-192
?	LPI 05-193	T05-193
?	LPI 05-194	T05-194
Turkana	LPI 06-195	T05-195
?	LPI 06-196	T06-196
?	LPI 06-197	T06-197
Yocco	LPI 07-198	T07-198
Galileo	Lpi 07-199	T07-199
Pavav	Lpi 07-200	T07-200
?	Lpi 07-201	T07-201
Verreaux	Lpi 07-203	T07-203
?	Female pup ioivias	T-IPF
?	Male pup ioivia's	T-IPM

Zimbabwe - Hwange Lycaon pictus samples
Donated by Greg Rasmussen

Donator ID	Clare's Lab ID
A137	G-A137
ROMA 33	G-ROM33
M038 08-08-07	G-M038
Marble Hwange	G-MARBLE
PITA, Alpha Male	G-PITA
UM39	G-UM39

Zimbabwe - Hwange Lycaon pictus samples
Donated by Joshua Ginsberg via RK Wayne (Girman study)

Donator ID	Clare's Lab ID
ZIM CB Lpi DNA	RKW 518
Zim CR Lpi DNA	RKW 516
ZIM NO Lpi DNA	RKW 517
ZIM XX Lpi DNA	RKW 519
ZIM ZZ Lpi DNA	RKW 520

Zimbabwe - Hwange Nyamandlovu Lycaon pictus samples
Donated by Greg Rasmussen via Linda Munson

Donator ID	Lab ID
ZWD10	RKW 6211
ZWD11	RKW 6212
ZWD12	RKW 6213
ZWD13	RKW 6214
ZWD14	RKW 6215
ZWD15	RKW 6216
ZWD16	RKW 6217
ZWD17	RKW 6218
ZWD18	RKW 6219
ZWD19	RKW 6220
ZWD20	RKW 6221
ZWD21	RKW 6222
ZWD22	RKW 6223
ZWD23	RKW 6224
ZWD24	RKW 6225
ZWD25	RKW 6226
ZWD26	RKW 6227
ZWD27	RKW 6228
ZWD9	RKW 6210

Zimbabwe –African wild dogs translocated into Hwange
Donated by Greg Rasmussen

Donator ID	Clare's Lab ID
ANGELA	G-ANGEL
#40 Brutus	G-BRU40
JOHN	G-JOHN
LONDZI / LOND	G-LONDZI
#30 Minto	G-MIN30
PB1	G-PB1
PB22	G-PB22
PB29	G-PB29
PB3	G-PB3
PBS 08-05-07	G-PBS1
PBS 05-08-07	G-PBS2
37 Male Rambo	G-RAM30
SIPHO	G-SIPHO
ULAKA	G-ULAK
B 13-8-07	G-B138

Zimbabwe – Lowveld (Savé & Malilangwe) Lycaon pictus samples
Donated by R.Groom

Donator sample ID	Other ID	Clare's Lab ID
LWDP F1	Spanners	ZLF1
LWDP F2	Racoon	ZLF2
LWDP F3	Sandy	ZLF3
LWDP M1	Blackie	ZLM1
LWDP M2	Survivor	ZLM2
LWDP M3	Darkie	ZLM3
DWD 001		ZL01
DWD 003		ZL03
DWD 004		ZL04
DWD 005	Sandy	ZL05
DWD 007		ZL07
DWD 008		ZL08

Samples from Malilangwe

MAL Pups 1-4		Mal-1
MAL Pups 1-4		Mal-2
MAL Pups 1-4		Mal-3
MAL Pups 1-4		Mal-4

Non-Lycaon pictus samples

DWD 002		ZL02

Namibia –Lycaon pictus samples Donated by Robin Lines

Donator sample ID	Clare's Lab ID
AWD#92	RKW 6203
DJOXY, AWD#89	RKW 6200
TJEKA M1, AWD#91	RKW 6202
TJEKA M2, AWD#90	RKW 6201

Namibia –Lycaon pictus samples Donated by F.Stander

Donator sample ID	Clare's Lab ID
Mangetti	ZG-Mang
Canal	ZG-Canal

Bostwana – Ghanzi Lycaon pictus samples Donated by M.J. Swarner

Donator sample ID	Clare's Lab ID
MJS-F2	BK-MJSF2

Mozambique –Ciné/Sofala Lycaon pictus samples Donated by J-M André

Donator sample ID	Clare's Lab ID
M0Z SOF1	AT003, CMAWMZ001T
M0Z SOF2	AT004, CMAWMZ002T
M0Z SOF3	AT005, CMAWMZ003T

Mozambique – Niassa Reserve Lycaon pictus samples Donated by C.Begg

Donator sample ID	Clare's Lab ID
	MozNa/MozNb

Tanzania – Captive Lycaon pictus samples (Mkomzai) Donated by R.Woodroffe

Donator sample ID	Clare's Lab ID
	WDP14a

Kenya – Magadi Road Lycaon pictus samples Donated by R.Woodroffe

Donator sample ID	Clare's Lab ID
LPI020	S-MagRd

Tanzania – Selous Lycaon pictus samples
Donated by Scott Creel via R.K. Wayne (Girman study)

Donator sample ID	Clare's Lab ID
ARCHER	RKW 12157
BLONDY	RKW 12158
BRACKY	RKW 6204
CALL	RKW 12159
Dish	RKW 5306
EF30	RKW 6208
HAT	RKW 6205
K2	RKW 12164
KING	RKW 12156
MAGNUM	RKW 12150
MBILI	RKW 12165
MM1	MM1
MPIRA	RKW 12151
Neptune	Neptune
PILGRIM	RKW 12152
PLUTO	RKW 12149
SALLY	RKW 6207
Shiva	RKW 12595
SOCKO	RKW 6206
Star	RKW 12564
Thor	RKW 12162
Wengi Pup	RKW 5307
Wishbone	RKW 12163

Tanzania – Masai Steppe Lycaon pictus samples
Donated by Aart Visée

Donator ID	Clare's Lab ID	Donator ID	Clare's Lab ID
104	MK104	298	MK298
202	MK202	300	MK300
251	MK251	305	MK305
261	MK261	308	MK308
262	MK262	310	MK310
263	MK263	333	MK333
264	MK264	335	MK335
265	MK265	336	MK336
269	MK269	337	MK337
274	MK274	348 Kisi	MK348 Kisi
276	MK276	348 Kimo	MK348 Kimo
284	MK284	368	MK368
285	MK285	389	MK389
289	MK289	448	MK448
291	MK291	651	MK651
294	MK294	730	MK730
296	MK296		

Tanzania –Old Serengeti Lycaon pictus samples
Donated by Sarah Cleaveland via R.K. Wayne (Girman study)

Donator ID	Clare's Lab ID
SRI 15, Collar	RKW 154
SRI 14, Mwenzi	RKW 165
SRI 3	RKW 166
D.HGm	RKW 5309
D583	RKW 5310
Fleur	RKW 5311
Legs	RKW 5312
Limp	RKW 5314
MDM	RKW 5315
J1 SRI ♀	RKW 328

Kenya –Old Masai Mara Lycaon pictus samples
Donated by Pieter Kat via R.K. Wayne (Girman study)

Donator ID	Clare's Lab ID
Umlaut	RKW 171
Saddle	RKW 176
Ring mara	RKW 177
SCPHP	RKW 179
Bat- 11	RKW 236
Arror	RKW 285
Saddle	RKW 176 & 294
Bald ♂	RKW 444
Scorpian Hp	RKW 446
H Lpi ♂	RKW 447
B.Guards	RKW 5308

Tanzania – Re-established Serengeti - Mara Lycaon pictus samples
Donated by Sarah Cleaveland & E.Masenga

Donator ID	Clare's Lab ID
WD 12.	S-WD1205
WD 13.05	S-WD1305
Lycaon Dog 1	S-D1LOS
Lycaon Dog 2	S-D2LOS
Lycaon	S-FMUT
LPI-1537	S-1537
LPI-1539	S-1539
LPI-1541	S-1541
LPI-1543	S-1543
LPI-1545	S-1545
LPI-1547	S-1547
OL0707	S-OL0707
WD 09.05	S-WD0905
WD 10.05	S-WD1005
WD 11.05	S-WD1105
WD 14.05	S-WD1405

Kenya – Laikipia Lycaon pictus samples
Donated by R.Woodroffe

Donator ID	Clare's Lab ID	Other ID's
WDF 12, LPI018	AB001	CMAWLA001B
WDF 28, LPI035	AB002	CMAWLA002B
WDM 30, LPI037	AB003	CMAWLA003B
WDM 32, LPI043	AB004	CMAWLA004B
WDF 33, LPI044	AB005	CMAWLA005B
WDM 34, LPI045	AB006	CMAWLA006B
WDF 35, LPI046	AB007, WDF 35	CMAWLA007B
WDM 36, LPI047	AB008	CMAWLA008B
WDM 37, LPI048	AB009	CMAWLA009B
WDF 38, LPI049	AB010	CMAWLA010B
WDM 39, LPI050	AB011	CMAWLA011B
WDM 40, LPI051	AB012	CMAWLA012B
WDF 41, LPI052	AB013	CMAWLA013B
WDF 42, LPI053	AB014	CMAWLA014B
WDF 43, LPI054	AB015	CMAWLA015B
WDM 44, LPI055	AB016	CMAWLA016B
WDM 45, LPI056	AB017	CMAWLA017B
WDM 46, LPI057	AB018	CMAWLA018B
WDM 47, LPI058	AB019	CMAWLA019B
WDM 48, LPI059	AB020	CMAWLA020B
WDM 49, LPI060	AB021	CMAWLA021B
WDM 50, LPI061	AB022	CMAWLA022B
WDM 51, LPI062	AB023	CMAWLA023B
WDF 52, LPI063	AB024	CMAWLA024B
WDF 2, LPI002	AB025	CMAWLA025B
WDF 2, LPI002	AB026	CMAWLA026B
WDM 5, LPI005	AB027	CMAWLA027B
WDF 6, LPI006	AB028	CMAWLA028B
WDF 7, LPI011	AB029	CMAWLA029B
WDF 8, LPI012	AB030	CMAWLA030B
WDF 8, LPI012	AB031	CMAWLA031B
WDF 9, LPI013	AB032	CMAWLA032B
WDM 10, LPI014	AB033	CMAWLA033B
WDM 11, LPI015	AB034	CMAWLA034B
WDF 12, LPI016	AB035	CMAWLA035B
WDM 13, LPI017	AB036	CMAWLA036B
WDF 14, LPI020	AB038	CMAWLA037B
WDM 15, LPI021	AB039	CMAWLA038B
WDM 16, LPI022	AB041	CMAWLA039B
WDF 17, LPI023	AB042	CMAWLA040B
WDM 18, LPI024	AB043	CMAWLA041B
WDM 19, LPI025	AB044	CMAWLA042B
WDF 20, LPI026	AB045	CMAWLA043B
WDF 21, LPI027	AB046	CMAWLA044B
WDF 21, LPI027	AB047	CMAWLA045B
WDM 22, LPI028	AB048	CMAWLA046B
WDM 23, LPI030	AB049	CMAWLA047B
WDF 24, LPI031	AB050	CMAWLA048B
WDF 25, LPI032	AB051	CMAWLA049B
WDF 28, LPI035	AB052	CMAWLA050B
WDF 25, LPI032	AB053	CMAWLA051B
WDF 26, LPI033	AB054	CMAWLA052B
WDF 27, LPI034	AB055	CMAWLA053B
WDM 29, LPI036	AB056	CMAWLA054B
WDM 30, LPI037	AB057	CMAWLA055B
WDM 31, LPI038	AB058	CMAWLA056B
WDM 53, LPI064	AB059	CMAWLA057B
WDM 54, LPI065	AB060	CMAWLA058B
WDM 55, LPI066	AB061	CMAWLA059B
WDM 56, ????	AB062	CMAWLA060B
WDM 57, ????	AB063	CMAWLA061B
WDF 58, LPI069	AB064	CMAWLA062B
LPP 025 LPI009, LPI009	AB065	CMAWLA063B
LPP 019 LPI002/008, LPI008	AB066	CMAWLA064B
LPP 026 LPI 010, LPI010	AB067	CMAWLA065B
LPI 032	AB071	
LPI 015, LPP 093	AB072	
LPI029, LPI029	AT002	CMAWLA066T
WDF 59, LPI 070	WDF 59	
WDF 62, LPI 073	WDF 62	
WDF 66, LPI 077	WDF 66	
WDF 67, LPI 078	WDF 67	
WDF 69, LPI 080	WDF 69	
WDF 70, LPI 081	WDF 70	
WDF 72, LPI 083	WDF 72	
WDF 72, LPI 083	WDF 72	
WDF 73, LPI 084	WDF 73	
WDF 74, LPI 085	WDF 74	
WDF 77, LPI 088	WDF 77	
WDF 78, LPI 089	WDF 78	
WDF 79, LPI 090	WDF 79	
WDM 49, LPI 060	WDM 49	
WDM 54, LPI 065	WDM 54	
WDM 60, LPI 071	WDM 60	
WDM 61, LPI 072	WDM 61	
WDM 63, LPI 074	WDM 63	
WDM 64, LPI 075	WDM 64	
WDM 65, LPI 076	WDM 65	
WDM 68, LPI 079	WDM 68	
WDM 71, LPI 082	WDM 71	
WDM 75, LPI 086	WDM 75	
WDM 76, LPI 087	WDM 76	

EUzoos samples:

B=blood, S=Serum, T=tissue. SAI = South African Imports; SAH = Descendent from SAI. #T numbers are temporary ID. Most cases studbook ID = DNA ID.

Studbook #	Zoo	Local/other ID	Sex	Sampling date	SAI	SAH	DOB	Sample type	DNA Extracted
#3083	LYM	P97062 - Benguela	M	26/02/2002	Y		28/09/1995	B, S	05/03/2008
#5117	LYM	P93069 - Rafiki	M	03/09/2005			22/11/1993	S	18/02/2008
#5237	AMS	M95168	F	01/10/2007			29/10/1995	B, S, T	17/10/2007
#5331	LYM	P20048	M	10/01/2008		Y	04/11/2005	B, S	21/01/2008
#5333	ATT	MLYP05	F	16/01/2008		Y	08/01/1998	B, S	01/02/2008
#5345	ROM	3207	M	19/02/2008			16/11/1998	B, S, T	05/03/2008
#5353	AMN	LP18	M	25/06/2008				B, S	12/08/2008
#5355	AMN	LP20	F	25/06/2008			13/10/1999	B, S	12/08/2008
#5361	EBE	LYC003	F	01/11/2007		Y	17/10/1999	B, S, T	22/11/2007
#5364	MUN	044015-Pointl	F	16/10/2007		Y	17/10/1999	B, S	17/10/2007
#5365	MUN	044016-Salome	F	16/10/2007		Y	17/10/1999	B, S	17/10/2007
#5369	HIL	M99332	M	Jul-08				B, S	12/08/2008
#5373	BAS	393365	M					B, S	12/08/2008
#5376	LIS	M00113	F	14/01/2008			26/10/1999	B, S	28/01/2008
#5394	LIS	M01112	M	14/01/2008			30/10/1999	B, S	01/02/2008
#5395	LIS	M01113	M	14/01/2008			30/10/1999	B, S	01/02/2008
#5396	LIS	M01114	M	14/01/2008			30/10/1999	B, S	01/02/2008
#5398	BAS	010333	F					B, S	12/08/2008
#5432	FRI	P97005 - Kassanga	M	15/10/2001		Y	01/03/1997	S	18/02/2008
#5434	Lym	H20237 - Kassama	F	23/12/2002		Y	01/03/1997	S	18/02/2008
#5438	KNO	P97011 - Kassala	F	26/02/2002		Y	01/03/1997	S	18/02/2008
#5442	KNO	P98044 - Kang	M	26/02/2002		Y	15/06/1998	S	18/02/2008
#5443	ATT	MLYP02	M	16/01/2008		Y	08/01/1998	B, S	01/02/2008
#5450	AAL	LYC11	M	05/02/2008		Y	22/11/2000	T	13/03/2008
#5452	BEL	5072	M	16/01/2008			29/11/2000	B, S	01/02/2008
#5453	PON	100304	M	17/03/2008				B, S	12/08/2008
#5455	PON	100303	M	17/03/2008				B, S	13/08/2008
#5456	BEL	5075	M	16/01/2008			29/11/2000	B, S	01/02/2008
#5457	PON	100302	M	17/03/2008				B, S	13/08/2008
#5458	BEL	5077	M	16/01/2008			29/11/2000	B, S	01/02/2008
#5459	War		F				29/11/2000		
#5462	War		F				29/11/2000		
#5464	LE	262	F	?			29/11/2000	B, S	21/01/2008
#5467	COL	BLL895, BRAM	M	27/02/2008			11/04/2000	B, S	05/03/2008
#5472	CAR	P20081 - Kippa	M	25/02/2002		Y	13/11/2000	S	18/02/2008
#5474	LYM	P20077 - Shue	F	26/02/2002		Y	13/11/2000	S	18/02/2008
#5476	COL	P20079 - Krane	F	26/02/2002		Y	13/11/2000	S	18/02/2008
#5477	COL	P20080 - Depti	F	26/02/2002		Y	13/11/2000	S	18/02/2008
#5493	LAP	3193	F	15/10/2007			29/10/2001	B, S	17/10/2007
#5503	EDI	P21050 - Snake	M	28/08/2002		Y	24/11/2001	S	18/02/2008
#5509	AMN	LP24	F	25/06/2008		Y	24/11/2001	B, S	12/08/2008
#5512	MUN	044017	M	15/10/2007			02/11/2001	B, S	17/10/2007
#5513	ESK	1714	F	18/10/2007			02/11/2001	B, S	12/11/2007
#5514	MUN	044018	M	16/10/2007			02/11/2001	B, S	17/10/2007
#5515	ESK	1715	F	18/10/2007			02/11/2001	B, S	12/11/2007
#5516	MUN	044019	M	15/10/2007			02/11/2001	B, S	17/10/2007
#5517	BOR	RA0011	F	05/11/2007			02/11/2001	B, S	13/11/2007
#5518	MUN	044020	M	16/10/2007			02/11/2001	B, S	17/10/2007
#5527	HIL	M02341	M	16/06/2008			24/11/2002	B, S	11/08/2008
#5529	COL	BLL896, BEM	M	27/02/2008			24/11/2002	B, S	05/03/2008
#5530	HIL	M02344	M	16/06/2008			24/11/2002	B, S	11/08/2008
#5531	COL	BLL897, MARANI	M	27/02/2008		Y	24/11/2002	B, S	05/03/2008
#5532	COL	BLL898, CHICANI	M	27/02/2008		Y	24/11/2002	B, S	05/03/2008
#5533	KOL	8904	F	31/03/2008		Y	24/11/2002	B, S	07/04/2008
#5534	AAL	LYC25	F	05/02/2008			24/11/2002	B, S	19/02/2008
#5537	AMS	M02168	F	06/02/2008			24/11/2002	B, S	05/03/2008

#5549	AMS	M06010	M	06/02/2008			04/11/2002	B, S	05/03/2008
#5550	AMS	M06009	M	06/02/2008			04/11/2002	B, S	05/03/2008
#5552	EDI	Pip	F	04/12/2007			04/11/2002	B, S	07/01/2008
#5553	EDI	Spot	F	04/11/2002				B	
#5554	EDI	Spice	F	04/11/2002				B	
#5555	EDI	Ginger	F	04/11/2002				B	
#5572	RAM	200982	F	22/10/2007			27/01/2003	B, S	12/11/2007
#5574	Sig	Ring	M	10/03/2010			27/01/2003		
#5575	LAP	S5575	M	15/10/2007			27/01/2003	B, S	17/10/2007
#5577	KER	M02140	M	18/12/2007			09/11/2002	B, S	07/01/2008
#5578	HIL	M02328	F	Jul-08				B, S	12/08/2008
#5579	HIL	M02329	F	Jul-08				B, S	12/08/2008
#5580	HIL	M02330	F	16/06/2008			09/11/2002	B, S	11/08/2008
#5586	DVU	033288 Bianca	F	15/01/2008		Y	09/11/2002	B, S	28/01/2008
#5588	PON	100306	F	16/06/2008		Y		B, S	13/08/2008
#5592	DUB	A4M037 ASSEGAI	M	17/01/2008		Y	02/11/2002	B, S	01/02/2008
#5593	DUB	P22049	M	27/01/2003		Y	02/11/2002	S	18/02/2008
#5594	DUB	A4M039 YELLA	M	17/01/2008		Y	02/11/2002	B, S	01/02/2008
#5595	EDI	M07B03	M	20/06/2008		Y		B, S	12/08/2008
#5596	KNO	P22052	F	27/01/2003		Y	02/11/2002	S	18/02/2008
#5597	KNO	P22053	F	27/01/2003		Y	02/11/2002	S	18/02/2008
#5598	KNO	P22054	F	27/01/2003		Y	02/11/2002	S	18/02/2008
#5599	KNO	P22055	F	27/01/2003		Y	02/11/2002	S	18/02/2008
#5601	KNO	P22057	F	27/01/2003		Y	02/11/2002	S	18/02/2008
#5602	KNO	P22058	F	27/01/2003		Y	02/11/2002	S	18/02/2008
#5603	KNO	P22059	F	27/01/2003		Y	02/11/2002	S	18/02/2008
#5605	KOL	8900	M	31/03/2008			27/01/2003	B, S	07/04/2008
#5606	KOL	8901	M	31/03/2008			27/01/2003	B, S	07/04/2008
#5607	KOL	8902		31/03/2008			27/01/2003	B, S	07/04/2008
#5620	Sig	Charly	M	10/03/2010			21/10/2003		
#5639	LE	M06033	M	?		Y	30/11/2003	B, S	21/01/2008
#5642	KER	M03141	M	18/12/2007			07/12/2003	B, S	07/01/2008
#5643	KER	M03142	M	18/12/2007			07/12/2003	B, S	07/01/2008
#5644	KER	M03143	M	18/12/2007			07/12/2003	B, S	07/01/2008
#5645	KER	M03144	M	18/12/2007			07/12/2003	B, S	07/01/2008
#5646	KER	M03145	M	18/12/2007			07/12/2003	B, S	07/01/2008
#5647	KER	M013146	M	18/12/2007			07/12/2003	B, S	07/01/2008
#5666	LON	4537	M	20/11/2007			17/11/2004	B	22/11/2007
#5667	LON	4538	M	20/11/2007			17/11/2004	B	22/11/2007
#5669	LON	4539	M	20/11/2007			17/11/2004	B	22/11/2007
#5671	LON	4540	M	20/11/2007			17/11/2004	B	22/11/2007
#5685	ROS	4471	M	14/01/2008			15/11/2004	B, S	28/01/2008
#5686	ROS	4472	M	14/01/2008			15/11/2004	B, S	28/01/2008
#5687	ROS	4473	M	14/01/2008			15/11/2004	B, S	28/01/2008
#5688	ROS	4474	M	14/01/2008			15/11/2004	B, S	28/01/2008
#5689	ROS	4475	M	14/01/2008			15/11/2004	B, S	28/01/2008
#5696	DOR	6CC138	F	06/03/2008			05/11/2004	B, S	03/04/2008
#5699	DOR	6C14EB	F	06/03/2008			05/11/2004	B, S	03/04/2008
#5700	DOR	710SF6	F	06/03/2008			05/11/2004	B, S	03/04/2008
#5704	DOR	6C2BF8	F	06/03/2008			05/11/2004	B, S	03/04/2008
#5721	LE	M06034-Max	M	?		Y	23/02/2005	B, S	21/01/2008
#5725	AMN	LP25	M	25/06/2008				B, S	12/08/2008
#5726	AMN	LP26	M	25/06/2008				B, S	12/08/2008
#5727	AMN	LP27	M	25/06/2008				B, S	12/08/2008
#T5932	KOL	10126	F	31/03/2008			29/12/2006	B, S	07/04/2008
#T5936	KOL	10130	M	31/03/2008			29/12/2006	B, S	07/04/2008
#T6094	KOL	10454	M	14/01/2008			30/10/2007	B, S	21/01/2008
#T6095	KOL	10455	M	14/01/2008			30/10/2007	B, S	21/01/2008
#T6096	KOL	10456	M	14/01/2008			30/10/2007	B, S	21/01/2008
#T6097	KOL	10457	M	14/01/2008			30/10/2007	B, S	21/01/2008

#T6098	KOL	10458	M	14/01/2008			30/10/2007	B, S	21/01/2008
#T6099	KOL	10459	F	14/01/2008			30/10/2007	B, S	21/01/2008
#T6100	KOL	10460	F	14/01/2008			30/10/2007	B, S	21/01/2008
#T6101	KOL	10461	F	14/01/2008			30/10/2007	B, S	28/01/2008
31844	PON	250229600031844	F	06/08/2008			05/06/2006	B, S	13/08/2008
40405	PON	250229600040405	F	06/08/2008				B, S	13/08/2008
42698	PON	250229600042698	F	06/08/2008			05/06/2006	B, S	13/08/2008
44252	PON	250229600044252	F	06/08/2008				B, S	13/08/2008
45658	PON	250229600045658	F	06/08/2008				B, S	13/08/2008
100390	PON	(T5955) 100390	F	21/04/2008		Y	05/06/2006	B, S	13/08/2008
AT001	Edi	DEAD PUP	?	28/08/2007		Y	28/08/2007	T	
#768		AT006	F	NA			c.1980	T	12/09/2007
AT007			M					T	12/09/2007
#892		AT008	F	NA			c.1982	T	12/09/2007
AT009			M	NA				T	12/09/2007
DUIS	DUI		F	15/07/2008				B, S	
M07114	HIL	T6089	m	16/06/2008				B, S	12/08/2008
M07118	HIL	(T6090) M07118	F	Jul-08			31/10/2007	B, S	12/08/2008
M07119	HIL	(T6091) M07119	F	Jul-08			31/10/2007	B, S	12/08/2008
M07120	HIL	(T6092) M07120	m	Jul-08			31/10/2007	B, S	12/08/2008
M07121	HIL	(T6093) M07121	m	Jul-08			31/10/2007	B, S	12/08/2008
#T6120	LYM	P20789	M			Y	04/11/2007	H	10/04/2008
#T6121	LYM	P20790	F			Y	04/11/2007	H	10/04/2008
#T6122	LYM	P20791	M			Y	04/11/2007	H	10/04/2008
#T6124	LYM	P20793	M			Y	04/11/2007	H	10/04/2008
#T6127	LYM	P20796	M			Y	04/11/2007	H	10/04/2008
#T6128	LYM	P20799	F			Y	04/11/2007	H	10/04/2008
#T1350	BEW	DOG1	F	18/02/2008		Y	01/03/2002	B, S	19/02/2008
#T1352	BEW	DOG3	M	18/02/2008		Y	01/03/2002	B, S	19/02/2008
#T1353	BEW	DOG4	M	18/02/2008		Y	01/03/2002	B, S	19/02/2008
#T1357	BEW	DOG5	F	18/02/2008		Y	01/03/2002	B, S	19/02/2008
#T5807	BAS	050352	M					B, S	12/08/2008
#T5819	PEA	NICK	M	29/10/2007		Y	25/09/2005	B, S	12/11/2007
#T5820	PEA	RON	M	29/10/2007		Y	25/09/2005	B, S	12/11/2007
#T5821	PEA	REX	M	29/10/2007		Y	25/09/2005	B, S	12/11/2007
#T5822	PEA	OBER	M	29/10/2007		Y	25/09/2005	B, S	12/11/2007
#T5823	PEA	KRYSTO	M	29/10/2007		Y	25/09/2005	B, S	12/11/2007
#T5824	PEA	ROSTA	M	29/10/2007		Y	25/09/2005	B, S	12/11/2007
#T5825	PEA	DAN	M	29/10/2007		Y	25/09/2005	B, S	12/11/2007
#T5827	DVU	033320 Andy	F	15/01/2008		Y	25/09/2005	B, S	28/01/2008
#T5828	DVU	033321 Ela	F	15/01/2008		Y	25/09/2005	B, S	28/01/2008
#T5829	DVU	033322 Irma	F	15/01/2008		Y	25/09/2005	B, S	28/01/2008
#T5835	KOL	9832	M	31/03/2008			18/11/2005	B, S	07/04/2008
#T5836	KOL	9833	M	31/03/2008			18/11/2005	B, S	07/04/2008
#T5837	KOL	9834	M	31/03/2008			24/11/2005	B, S	07/04/2008
#T5838	KOL	9835	M	31/03/2008			24/11/2005	B, S	07/04/2008
#T5839	KOL	9836	M	31/03/2008			24/11/2005	B, S	07/04/2008
#T5840	KOL	9837	M	31/03/2008			24/11/2005	B, S	07/04/2008
#T5857	LYM	P20572	F	10/01/2008		Y	04/11/2005	B, S	19/02/2008
#T5858	LYM	P20573	F	10/01/2008		Y	04/11/2005	B, S	19/02/2008
#T5859	LYM	P20574	F	10/01/2008		Y	04/11/2005	B, S	19/02/2008
#T5862	MUN	044022	F	15/10/2007		Y	01/11/2005	B, S	17/10/2007
#T5886	AAL	LYC27	M	05/02/2008		Y	23/10/2006	B, S	19/02/2008
#T5887	AAL	LYC28	M	05/02/2008		Y	23/10/2006	B, S	19/02/2008
#T5888	AAL	LYC29	M	05/02/2008		Y	23/10/2006	B, S	19/02/2008
#T5889	AAL	LYC30	M	05/02/2008		Y	23/10/2006	B, S	19/02/2008
#T5890	AAL	LYC31	F	05/02/2008		Y	23/10/2006	B, S	19/02/2008
#T5891	AAL	LYC32	F	05/02/2008		Y	23/10/2006	B, S	19/02/2008
#T5892	AAL	LYC33	F	05/02/2008		Y	23/10/2006	B, S	19/02/2008
#T5894	BAS	600354	F					B, S	12/08/2008

#T5896	BAS	600356	F					B, S	12/08/2008
#T5920	BOR	RA0027	M	05/11/2007		Y	22/10/2006	B, S	13/11/2007
#T5921	BOR	RA0028	F	05/11/2007		Y	22/10/2006	B, S	13/11/2007
#T5922	BOR	RA0029	F	05/11/2007		Y	22/10/2006	B, S	13/11/2007
#T5923	BOR	RA0030	F	05/11/2007		Y	22/10/2006	B, S	13/11/2007
#T5924	BOR	RA0031	F	05/11/2007		Y	22/10/2006	B, S	13/11/2007
#T5925	BOR	RA0032	F	05/11/2007		Y	22/10/2006	B, S	13/11/2007
#T5926	BOR	RA0033	F	05/11/2007		Y	22/10/2006	B, S	13/11/2007
#T5927	BOR	RA0034	F	05/11/2007		Y	22/10/2006	B, S	13/11/2007
#T5950	PON	100385	M	21/04/2008		Y	05/06/2006	B, S	13/08/2008
#T5952	PON	100387	M	21/04/2008		Y	05/06/2006	B, S	13/08/2008
#T5953	PON	100388	M	21/04/2008		Y	05/06/2006	B, S	13/08/2008
#T5956	DUI	5386	F	15/07/2008		Y	05/06/2006	B, S	12/08/2008
#T5959	DUI	5389	F	15/07/2008		Y	05/06/2006	B, S	12/08/2008
#T5960	LAP	4322	F	15/10/2007			17/12/2006	B, S	17/10/2007
#T5983	EBE	LYC005	M	12/11/2007		Y	~06/2002	B, S	22/11/2007
#T5984	EBE	LYC006	M	12/11/2007		Y	~06/2002	B, S	22/11/2007
#T5985	EBE	LYC007	M	12/11/2007		Y	~06/2002	B, S	22/11/2007
#T5986	EBE	LYC008	M	12/11/2007		Y	~06/2002	B, S	22/11/2007
#T5987	BEW	DOG6	M	18/02/2008		Y	~/04/2002	B, S	19/02/2008
#T5988	BEW	DOG7	M	18/02/2008		Y	~/04/2002	B, S	19/02/2008
#T5990	BEW	DOG9	F	18/02/2008		Y	~/04/2002	B, S	19/02/2008
#T5991	BEW	DOG10	F	18/02/2008		Y	~/04/2002	B, S	19/02/2008
#T5993	BEW	DOG12	F	18/02/2008		Y	~/04/2002	B, S	19/02/2008
#T6003	Sig	Xenie	F	10/03/2010		Y	01/06/1998		
#T6022	Sig	Zeta	F	10/03/2010		Y	28/11/2000		
#T6026	ROM	4476	M	19/02/2008		Y	12/11/2001	B, S	05/03/2008
#T6027	ROM	4475	M	19/02/2008		Y	12/11/2001	B, S	05/03/2008
#T6029	Sig	Amandine	F	10/03/2010		Y	12/11/2001		
#T6030	Sig	Auridie	F	10/03/2010		Y	12/11/2001		
#T6031	War		M			Y	12/11/2000		
#T6032	Sig	Arthur	M	10/03/2010		Y	12/11/2001		
#T6033	War		M			Y	12/11/2000		
#T6034	Sig	Adeline	F	10/03/2010		Y	12/11/2001		
#T6035	Sig	Ariane	F	10/03/2010		Y	12/11/2001		
#T6046	Sig	Annabelle	F	10/03/2010		Y	25/11/2001		
#T6047	BOR	RA0006	M	05/11/2007		Y	25/11/2001	B, S	13/11/2007
#T6048	BOR	RA0007	M	05/11/2007		Y	25/11/2001	B, S	13/11/2007
#T6049	BOR	RA0008	M	05/11/2007		Y	25/11/2001	B, S	12/11/2007
#T6050	Qui		M				UNK		
#T6051	BEW	DOG13	F	18/02/2008		Y	05/12/2004	B, S	19/02/2008
#T6052	BEW	DOG14	M	18/02/2008		Y	05/12/2004	B, S	19/02/2008
#T6053	BEW	DOG17	M	18/02/2008		Y	10/12/2004	B, S	05/03/2008
#T6054	BEW	DOG20	F	18/02/2008		Y	27/10/2005	B, S	19/02/2008
#T6055	BEW	DOG21	M	18/02/2008		Y	27/10/2005	B, S	05/03/2008
#T6056	BEW	DOG22	M	18/02/2008		Y	27/10/2005	B, S	05/03/2008
#T6057	BEW	DOG23	F	18/02/2008		Y	27/10/2005	B, S	05/03/2008
#T6058	BEW	DOG24	M	18/02/2008		Y	27/10/2005	B, S	05/03/2008
#T6059	BEW	DOG25	F	18/02/2008		Y	27/10/2005	B, S	05/03/2008
#T6060	BEW	DOG26	F	18/02/2008		Y	27/10/2005	B, S	05/03/2008
#T6061	BEW	DOG27	F	18/02/2008		Y	27/10/2005	B, S	05/03/2008
#5509	COU	P21056 - Eva	M	15/02/2002			24/11/2001	S	18/02/2008
#5595	DUB	P22051	M	27/01/2003			02/11/2002	S	18/02/2008
AAL pup	AAL	AALBORG PUP		05/02/2008				H	

List of References

Abdo, Z., A. C. Keith, and P. Joyce. 2004. Evaluating the performance of likelihood methods for detecting population structure and migration. Molecular Ecology **13**:837-851.

Acevedo-Whitehouse, K. 2009. The importance of disease management programmes for wildlife conservation. Animal Conservation **12**:185-186.

Aguilar, A., G. Roemer, S. Debenham, M. Binns, D. Garcelon, and R. K. Wayne. 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. Proceedings of the National Academy of Sciences of the United States of America **101**:3490-3494.

Alexander, K. A., J. W. McNutt, M. B. Briggs, P. E. Standers, P. Funston, G. Hemson, D. Keet, and M. van Vuuren. 2010. Multi-Host Pathogens and Carnivore Mangagement in Southern Africa. Comparative Immunology, Microbiology and Infectious Diseases **33**:249-265.

Allendorf, F. W., and R. S. Waples. 1996. Conservation genetics of salmonid fishes. Pages 238-280 in J. C. Avise, and J. L. Hamrick, editors. Conservation genetics: cases histories from nature. Springer Verlag, New York.

Alvarez, K. 1993. Twilight of the panther; Biology, Bureaucracy and Failure in an Endangered Species Program. Myakka River Publishing, Sarasota, Florida.

Amos, W., and A. Balmford. 2001. When does conservation genetics matter? Heredity **87**:257-265.

Anderson, T. M., B. M. vonHoldt, S. I. Candille, M. Musiani, C. Greco, D. R. Stahler, D. W. Smith, B. Padhukasahasram, E. Randi, J. A. Leonard, C. D. Bustamante, E. A. Ostrander, H. Tang, R. K. Wayne, and G. S. Barsh. 2009. Molecular and Evolutionary History of Melanism in North American Gray Wolves. Science **323**:1339-1343.

Appel, M. J., R. A. Yates, G. L. Foley, J. J. Bernstein, S. Santinelli, L. H. Spelman, L. D. Miller, L. H. Arp, M. Anderson, M. Barr, and a. et. 1994. Canine distemper epizootic in lions, tigers, and leopards in North America. Journal of Veterinary Diagnostic Investigation **6**:277-288.

Arcaro, K. F., and A. Eklund. 1999. A review of MHC-based mate preferences and fostering experiments in two congenic strains of mice. Genetica **104**:241-244.

Arctander, P., C. Johansen, and M. A. Coutellec-Vreto. 1999. Phylogeography of three closely related African bovids (tribe Alcelaphini). Molecular Bioliogy and Evolution **16**:1724-1739.

Ardern, S. L., and D. M. Lambert. 1997. Is the black robin in genetic peril? Molecular Ecology **6**:21-28.

Aspi, J., E. Roininen, J. Kiiskilä, M. Ruokonen, I. Kojola, L. Bljudnik, P. Danilov, S. Heikkinen, and E. Pulliainen. 2009. Genetic structure of the northwestern Russian wolf populations and gene flow between Russia and Finland. Conservation Genetics **10**:815-826.

Aspi, J., E. Roininen, M. Ruokonen, I. Kojola, and C. Vila. 2006. Genetic diversity, population structure, effective population size and demographic history of the Finnish wolf population. Molecular Ecology **15**:1561-1576.

Ballou, J. D., M. Gilpin, and T. J. Foose 1995. Population management for survival and recovery: Analytical methods and strategies in small population conservation. Columbia University Press, New York.

Bardeleben, C., R. L. Moore, and R. K. Wayne. 2005. A molecular phylogeny of the Canidae based on six nuclear loci. Molecular Phylogenetics and Evolution **37**:815-831.

Barrett, S. C. H., and J. Kohn. 1991. Genetic and evolutionary consequences of small population size in plants: implications for conservation. Pages 3-30 in D. A. Falk, and K. E. Holsinger, editors. Genetics and conservation of rare plants. Oxford University Press, Oxford.

Bauman, K. L., C. S. Asa, J. Grisham, and W. Verberkmoes. 2004. Captive canid conservation. Pages 280-288 in C. Sillero-Zubiri, M. Hoffmann, and D. W.

Macdonald, editors. Canids: foxes, wolves, jackals and dogs: status survey and conservation action plan. Gland: IUCN.

Beauclerc, K., B. Johnson, and B. White. 2001. Genetic rescue of an inbred captive population of the critically endangered Puerto Rican crested toad (*Peltophryne lemur*) by mixing lineages. Conservation Genetics **11**:21-32.

Beerli, P. 2004. Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. Molecular Ecology **13**:827-836.

Beerli, P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics **22**:341-345.

Beerli, P. 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? Pages 42-74 in G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi, editors. Population genetics for animal conservatin. Cambridge University Press, Cambridge.

Beerli, P., and J. Felsenstein. 1999. Maximum likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. . Genetics **152**:763-773.

Bekessy, S. A., R. A. Ennos, M. A. Burgman, A. C. Newton, and P. K. Ades. 2003. Neutral DNA markers fail to detect genetic divergence in an ecologically important trait. Biological Conservation **110**:267-275.

Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. 2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

Bere, R. M. 1956. The Africa wild dog. Oryx **3**.

Berger, L., R. Speare, P. Daszak, D. E. Green, A. A. Cunningham, C. L. Goggin, R. Slocombe, M. A. Ragan, A. D. Hyatt, K. R. McDonald, H. B. Hines, K. R. Lips, G. Marantelli, and H. Parkes. 1998. Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. Proceedings of the National Academy of Sciences of the United States of America **95**:9031-9036.

Bergstrom, B. J., S. Vignieri, S. R. Sheffield, W. Sechrest, and A. A. Carlson. 2009. The Northern Rocky Mountain Gray Wolf is Not Yet Recovered. BioScience **59**:991-999.

Bernatchez, L., and C. Landry. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? Journal of Evolutionary Biology **16**:363-377.

Bollmer, J., F. Vargas, and P. Parker. 2007. Low MHC variation in the endangered Galápagos penguin (*Spheniscus mendiculus*). Immunogenetics **59**:593-602.

Bondinas, G. P., A. K. Moustakas, and G. K. Papadopoulos. 2007. The spectrum of HLA-DQ and HLA-DR alleles, 2006: a listing correlating sequence and structure with function. Immunogenetics **59**:539-553.

Bonnell, M. T. 1974. Elephant seal: genetic variation and near extinction. Science **184**:908-909.

Bos, D. H., D. Gopurenko, R. N. Williams, A. J. DeWoody, and L. Knowles. 2008. Inferring Population History and Demography Using Microsatellites, Mitochondrial DNA, and Major Histocompatibility Complex (MHC) Genes. Evolution **62**:1458-1468.

Boutelle, S. M., and H. J. Bertschinger. 2010. Reproductive management in captive and wild canids: contraception challenges. International Zoo Yearbook **44**:109-120.

Bouzat, J. 2010. Conservation genetics of population bottlenecks: the role of chance, selection, and history. Conservation Genetics **11**:463-478.

Boyce, W. M., P. W. Hedrick, N. E. Muggli-Cockett, S. Kalinowski, M. C. T. Penedo, and R. R. Ramey-Ii. 1997. Genetic Variation of Major Histocompatibility Complex and Microsatellite Loci: A Comparison in Bighorn Sheep. Genetics **145**:421-433.

Bradshaw, A. D. 1991. Genostasis and the limits to evolution. Philosophical Transactions of the Royal Society B: Biological Sciences **333**:289-305.

Breen, M., Z. Irvin, P. Downs, I. Gill, P. Bircher, N. Hart, and K. Bell. 1995. Parentage verification of a mountain zebra (*Equus zebra hartmannae*) using polymorphic microsatellite markers. Zoo Biology **14**:475-479.

Breen, M., S. Jouquand, C. Renier, C. S. Mellersh, C. Hitte, N. G. Holmes, A. Chéron, N. Suter, F. Vignaux, A. E. Bristow, C. Priat, E. McCann, C. André, S. Boundy, P. Gitsham, R. Thomas, W. L. Bridge, H. F. Spriggs, E. J. Ryder, A. Curson, J. Sampson, E. A. Ostrander, M. M. Binns, and F. Galibert. 2001. Chromosome-Specific Single-Locus FISH Probes Allow Anchorage of an 1800-Marker Integrated Radiation-Hybrid/Linkage Map of the Domestic Dog Genome to All Chromosomes. Genome Research **11**:1784-1795.

Briggs, W. H., and I. L. Goldman. 2006. Genetic Variation and Selection Response in Model Breeding Populations of Brassica rapa Following a Diversity Bottleneck. Genetics **172**:457-465.

Brooks, S. P., Gelman, A. 1998. General Methods for Monitoring Convergence of Iterative Simulrations. Journal of Computational and Graphical Statistics **7**:434-455.

Bryan, G. T., K.-S. Wu, L. Farrall, Y. Jia, H. P. Hershey, S. A. McAdams, K. N. Faulk, G. K. Donaldson, R. Tarchini, and B. Valent. 2000. A Single Amino Acid Difference Distinguishes Resistant and Susceptible Alleles of the Rice Blast Resistance Gene Pi-ta. Plant Cell **12**:2033-2046.

Bryja, J. C., N. Berthier, K. Galan, M. Cosson, J. F. . 2007. Density-related changes in selection pattern for major histocompatibility complex genes in fluctuating populations of voles. Molecular Ecology **16**:5084-5097.

Burrows, R. 1995. Demographic Changes and Social Consequences in Wild Dogs, 1964-1992. Pages 400-420 in A. R. E. Sinclair, Arcese, P., editor. Serengeti II: Dynamics, Management and Conservation of an Ecosystem. University of Chicago Press, Chicago.

Burrows, R., H. Hofer, and M. East. 1995. Population Dynamics, Intervention and Survival in African Wild Dogs (*Lycaon pictus*). Proceedings of the Royal Society B: Biological Sciences **262**:235-245.

Burrows, R., H. Hofer, and M. L. East. 1994. Demography, Extinction and Intervention in a Small Population: the Case of the Serengeti Wild Dogs. Proceedings of the Royal Society B: Biological Sciences **256**:281-292.

Campos, J. L., D. Posada, and P. Morán. 2006. Genetic variation at MHC, mitochondrial and microsatellite loci in isolated populations of Brown trout (*Salmo trutta*). Conservation Genetics **7**:515-530.

Carroll, L. S., D. J. Penn, and W. K. Potts. 2002. Discrimination of MHC-derived odors by untrained mice is consistent with divergence in peptide-binding region residues. Proceedings of the National Academy of Sciences **99**:2187-2192.

Caughley, G. 1994. Directions in cosnervation biology. Journal of Animal Ecology **63**:215-244.

Charbonnel, N., and J. Pemberton. 2005. A long-term genetic survey of an ungulate population reveals balancing selection acting on MHC through spatial and temporal fluctuations in selection. Heredity **95**:377-388.

Charlesworth, B. 2009. Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. Nature Reviews Genetics **10**:195-205.

Charlesworth, B., and D. Charlesworth. 1999. The genetic basis of inbreeding depression. Genetical Research **74**:329-340.

Childes, S. L. 1988. The past history, present status and distribution of the hunting dog, Lycaon pictus, in Zimbabwe. Biological Conservation **44**:301-316.

Cornuet, J. M., and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics **144**:2001-2014.

Courchamp, F., T. Clutton-Brock, and B. Grenfell. 2000. Multipack dynamics and the Allee effect in the African wild dog, *Lycaon pictus*. Animal Conservation **3**:277-285.

Courchamp, F., and D. W. Macdonald. 2001. Crucial importance of pack size in the African wild dog *Lycaon pictus*. Animal Conservation **4**:169-174.

Crandall, K. A. 2009. A multifaceted approach to species conservation. Animal Conservation **12**:105-106.

Crawford, N. G. 2009. SMOGD: Software for the Measurement of Genetic Diversity. Molecular Ecology Resources **10**:556-557.

Creel, S., and N. M. Creel. 1998. Six ecological factors that may limit African wild dogs, *Lycaon pictus*. Animal Conservation **1**:1-9.

Creel, S., and N. M. Creel 2002. The African wild dog: behavior, ecology and conservation. Princeton University Press, Princeton, NJ.

Creel, S., N. M. Creel, M. G. L. Mills, and S. L. Monfort. 1997. Rank and reproduction in cooperatively breeding African wild dogs: Behavioral and endocrine correlates. Behavioral Ecology **8**:298-306.

Creel, S., M. G. L. Mills, and J. W. McNutt. 2004. African wild dogs. Pages 337-352 in D. W. Macdonald, and C. Sillero-Zubiri, editors. Biology and Conservation of Wild Canids. Oxford University Press, Oxford.

Crow, J. F., and M. Kimura 1970. An Introduction to Population Genetics Theory. Burgess Publishing, Minneapolis.

Cutrera, A., and E. Lacey. 2007. Trans-species polymorphism and evidence of selection on class II MHC loci in tuco-tucos (Rodentia: Ctenomyidae). Immunogenetics **59**:937-948.

Daltry, J. C., Q. Bloxam, G. Cooper, M. L. Day, J. Hartley, M. Henry, K. Lindsay, and B. E. Smith. 2001. Five years of conserving the world's rarest snake, the Antiguan racer *Alsophis antiguae*. Oryx **35**:119-127.

Daszak, P., A. A. Cunningham, and A. D. Hyatt. 2000. Emerging Infectious Diseases of Wildlife - Threats to Biodiversity and Human Health. Science **287**:443-449.

Demandt, M. 2010. Temporal changes in genetic diversity of isolated populations of perch and roach. Conservation Genetics **11**:249-255.

Di Rienzo, A., A. C. Peterson, J. C. Garza, A. M. Valdes, M. Slatkin, and N. B. Freimer. 1994. Mutational processes of simple-sequence repeat loci in human populations. Proceedings of the National Academy of Sciences of the United States of America **91**:3166-3170.

Ditchkoff, S. S., S. R. Hoofer, R. L. Lochmiller, R. E. Masters, and R. A. Van Den Bussche. 2005. MHC-DRB evolution provides insight into parasite resistance in white-tailed deer. The Southwestern Naturalist **50**:57-64.

Doherty, P. C., and R. M. Zinkernagel. 1975. A biological role for the major histocompatibility anitgens. The Lancet **305**:1406-1409.

Earl, D. A. 2009. Stucture Harvester v.0.5.

Edmands, S. 2007. Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Molecular Ecology **16**:463-475.

Egea, R., S. Casillas, and A. Barbadilla. 2008. Standard and generalized McDonald-Kreitman test: a website to detect selection by comparing different classes of DNA sites. Nucleic Acids Research **36**:W157-162.

Eggert, F., W. Muller-rucholtz, and R. Fertsi. 1998. Olfactory cues associated with the major histocompatibility complex. Genetica **104**:191-197.

Ejsmond, M., and J. Radwan. 2009. MHC diversity in bottlenecked populations: a simulation model. Conservation Genetics **In press**.

Ekblom, R., S. A. Sæther, P. Jacobsson, P. Fiske, T. Sahlman, M. Grahn, J. A. Kålås, and J. Höglund. 2007. Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). Molecular Ecology **16**:1439-1451.

Ellstrand, N. C., and D. R. Elam. 1993. Population Genetic Consequences of Small Population-Size - Implications for Plant Conservation. Annual Review of Ecology and Systematics **24**:217-242.

Ennos, R. A., N. R. Cowie, C. J. Legg, and C. Sydes. 1997. Which measures of genetic variation are relevatnt in plan conservation? A case study of *Primula scotica* Pages 73-79 in T. J. Crawford, J. Spencer, D. Setevens, M. B. Usher, T. E. Tew, and J. Warren, editors. The role of genetics in conserving small populations. JNCC, Peterborough.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology **14**:2611-2620.

Excoffier, L., Laval, G., Schneider, S. 2006. Arlequin 3.11; An Integrated Software Package for Population Genetics Data Analysis, Geneva, Switzerland: Genetics and Biometry Laborator, University of Geneva.

Francisco, L., Langston, AA, Mellersh, CS, Neal, CL, and Ostrander, EA. 1996. A class of highly polymorphic tetranucleotide repeats for canine genetic Mammalian Genome **7**:359-362.

François, B., and L.-M. Nicolas. 2002. The estimation of population differentiation with microsatellite markers. Molecular Ecology **11**:155-165.

Frankham, R. 1995a. Effective population size/adult population size ratios in wildlife: a review. Genetics Research **66**:95-107.

Frankham, R. 1995b. Inbreeding and Extinction: A Threshold Effect. Conservation Biology **9**:792-799.

Frankham, R. 1996. Relationship of Genetic Variation to Population Size in Wildlife. Conservation Biology **10**:1500-1508.

Frankham, R. 2007. Genetic adaptation to captivity in species conservation programs. Molecular Ecology **0**:???-???

Frankham, R. 2008. Genetic adaptation to captivity in species conservation programs. Molecular Ecology **17**:325-333.

Frankham, R., J. D. Ballou, and D. A. Briscoe 2002. Introduction to conservation genetics. Cambridge University Press.

Frantzen, M. A. J., J. W. H. Ferguson, and M. S. de Villiers. 2001. The conservation role of captive African wild dogs (*Lycaon pictus*). Biological Conservation **100**:253-260.

Freckleton, R. P. 2009. The seven deadly sins of comparative analysis. Journal of Evolutionary Biology **22**:1367-1375.

Fuller, T. K., Kat, P.W., Bulger, J.B., Maddock, A.H., Ginsberg, J.R., Burrows, R., McNutt, J.W., & Mills, M.G.L. . 1992. Population dynamics of African wild dogs in D. R. McCullough, and H. Barrett, editors. Wildlife 2001: Populations. Elsevier Science Publishers, London.

Fuller, T. K., M. G. L. Mills, M. Borner, K. Laurenson, and P. W. Kat. 1992. Long distance dispersal by African wild dogs in East and South Africa. Journal of African Zoology **106**:535-537.

Furlong, R. F., and Z. Yang. 2008. Diversifying and Purifying Selection in the Peptide Binding Region of DRB in Mammals. Journal of Molecular Evolution **66**:384-394.

Fyumagwa, D., and H. Wiik. 2001. TAWRI Wildlife Veterinary Programme Annual Report.

Garrigan, D., and P. W. Hedrick. 2003. Perspective: Detecting adaptive molecular polymorphism: Lessons from the MHC. Evolution **57**:1707-1722.

Garza, J. C., and E. G. Williamson. 2001. Detection of reduction in population size using data from microsatellite loci. Molecular Ecology **10**:305-318.

Gascoyne, S. C., M. K. Laurenson, S. Lelo, and M. Borner. 1993. Rabies in African wild dogs (*Lycaon pictus*) in the Serengeti region, Tanzania. Journal Of Wildlife Diseases **29**:396-402.

Gautschi, B., G. Jacob, J. J. Negro, J. A. Godoy, J. P. Müller, and B. Schmid. 2003. Analysis of relatedness and determination of the source of founders in the captive bearded vulture, *Gypaetus barbatus*, population. Conservation Genetics **4**:479-490.

Gebremedhin, B., G. F. Ficetola, S. Naderi, H. R. Rezaei, C. Maudet, D. Rioux, G. Luikart, F. Ø, W. Thuiller, and P. Taberlet. 2009. Combining genetic and ecological data to assess the conservation status of the endangered Ethiopian walia ibex. Animal Conservation **12**:89-100.

Gebremedhin, B., Ficetola, G. F., Naderi, S., Rezaei, H. R., Maudet, C., Rioux, D., Flagstad, G. Luikart Ø., Thuiller, W., Taberlet, P. . 2009. Frontiers in identifying conservation units: from neutral markers to adaptive genetic variation. Animal Conservation **12**:107-109.

Ginsberg, J., G. Mace, and S. Albon. 1995. Local Extinction in a Small and Declining Population: Wild Dogs in the Serengeti. Proceedings of the Royal Society B: Biological Sciences **262**:221-228.

Girman, D., J, M. G. L. Mills, E. Geffen, and R. Wayne, K. 1997. A molecular genetic analysis of social structure, dispersal, and interpack relationships of the African wild dog (*Lycaon pictus*). Behavioral Ecology and Sociobiology **40**:187-198.

Girman, D. J., P. W. Kat, M. G. L. Mills, J. R. Ginsberg, M. Borner, V. Wilson, J. H. Fanshawe, C. Fitzgibbon, L. M. Lau, and R. K. Wayne. 1993. Molecular Genetic

and Morphological Analyses of the African Wild Dog (*Lycaon pictus*). Journal of Heredity **84**:450-459.

Girman, D. J., C. Vila, E. Geffen, S. Creel, M. G. L. Mills, J. W. McNutt, J. Ginsberg, P. W. Kat, K. H. Mamiya, and R. K. Wayne. 2001. Patterns of population subdivision, gene flow and genetic variability in the African wild dog (*Lycaon pictus*). Molecular Ecology **10**:1703-1723.

Glatson, A. R. 2001. Relevance of stubook data to the successful captive management of grey mouse lemurs. International Journal of Primatology **22**:57-69.

Goodall, J. 2003. Problems faced by wild and captive chimpanzees: finding solutions. Pages 145-151 in S. J. Armstrong, and R. G. Botzler, editors. The animal ethics reader. Routledge, London.

Goossens, B. t., L. s. Chikhi, M. Ancrenaz, I. Lackman-Ancrenaz, P. Andau, and M. W. Bruford. 2006. Genetic Signature of Anthropogenic Population Collapse in Orang-utans. PLoS Biology **4**:e25.

Gottelli, D., J. Marino, C. Sillero-Zubiri, and S. M. Funk. 2004. The effect of the last glacial age on speciation and population genetic structure of the endangered Ethiopian wolf (*Canis simensis*). Molecular Ecology **13**:2275-2286.

Gottelli, D., C. Sillero-Zubiri, G. D. Applebaum, M. S. Roy, D. J. Girman, J. Garcia-Moreno, E. A. Ostrander, and R. K. Wayne. 1994. Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simesis*. Molecular Ecology **3**:301-312.

Goudet, J. 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. Journal of Heredity **86**:485-486.

Groombridge, J. J., C. G. Jones, M. W. Bruford, and R. A. Nichols. 2000. Conservation biology: *Ghost alleles* of the Mauritius kestrel. Nature **403**:616-616.

Gusset, M., S. J. Ryan, M. Hofmeyr, G. van Dyk, H. T. Davies-Mostert, J. A. Graf, C. Owen, M. Szykman, D. W. Macdonald, S. L. Monfort, D. E. Wildt, A. H. Maddock, M. G. L. Mills, R. Slotow, and M. J. Somers. 2008. Efforts going to the dogs? Evaluating attempts to re-introduce endangered wild dogs in South Africa. Journal of Applied Ecology **45**:100-108.

Gusset, M., R. Slotow, and M. J. Somers. 2006. Divided we fail: the importance of social integration for the re-introduction of endangered African wild dogs (*Lycaon pictus*). Journal of Zoology **270**:502-511.

Gutierrez-Espeleta, G. A., P. W. Hedrick, S. T. Kalinowski, D. Garrigan, and W. M. Boyce. 2001. Is the decline of desert bighorn sheep from infectious disease the result of low MHC variation? Heredity **86**:439-450.

Guyon, R., T. D. Lorentzen, C. Hitte, L. Kim, E. Cadieu, H. G. Parker, P. Quignon, J. K. Lowe, C. Renier, B. Gelfenbeyn, F. Vignaux, H. B. DeFrance, S. Gloux, G. G. Mahairas, C. André, F. Galibert, and E. A. Ostrander. 2003. A 1-Mb resolution radiation hybrid map of the canine genome. Proceedings of the National Academy of Sciences of the United States of America **100**:5296-5301.

Hartl, D. L., and A. G. Clark 2007. Principles of Population Genetics. Sinauer Associates, Inc.

Hartstone-Rose, A., L. Werdelin, D. J. De Ruiter, L. R. Berger, and S. E. Churchill. 2010. The Plio-pleistocene ancestor of wild dogs, *Lycaon sekowein* sp. Journal of Paleontology **84**:299-308.

Harvey, P. H., and M. D. Pagel 1991. The comparative method in evolutionary biology. Oxford University Press, Oxford.

Hasegawa, M., H. Kishino, and T.-a. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution **22**:160-174.

Hawkins, C. E., C. Baars, H. Hesterman, G. J. Hocking, M. E. Jones, B. Lazenby, D. Mann, N. Mooney, D. Pemberton, S. Pyecroft, M. Restani, and J. Wiersma. 2006. Emerging disease and population decline of an island endemic, the Tasmanian devil *Sarcophilus harrisii*. Biological Conservation **131**:307-324.

Haydon, D. T., D. A. Randall, L. Matthews, D. L. Knobel, L. A. Tallents, M. B. Gravenor, S. D. Williams, J. P. Pollinger, S. Cleaveland, M. E. J. Woolhouse, C. Sillero-Zubiri, J. Marino, D. W. Macdonald, and M. K. Laurenson. 2006. Low-coverage vaccination strategies for the conservation of endangered species. Nature **443**:692-695.

Hayney, M. S., G. A. Poland, R. M. Jacobson, D. Rabe, D. J. Schaid, S. J. Jacobsen, and J. J. Lipsky. 1998. Relationship of HLA-DQAI alleles and humoral antibody following measles vaccination. International Journal of Infectious Diseases **2**:143-146.

Hedrick, P. W. 2001. Conservation genetics: where are we now? Trends in Ecology & Evolution **16**:629-636.

Hedrick, P. W. 2003. The major histocompability complex (MHC) in declining populations: an example of adaptive varation. Pages 97-113 in W. V. Holt, A. R. Pickard, J. C. Rodger, and D. E. Wildt, editors. Reproduction Science and Integrated Conservation. Cambridge University Press.

Hedrick, P. W., T. J. Kim, and K. M. Parker. 2001a. Parasite resistance and genetic variation in the endangered Gila topminnow. Animal Conservation **4**:103-109.

Hedrick, P. W., R. N. Lee, and C. Buchanan. 2003. Canine parvovirus enteritis, canine distemper, and major histocompatibility complex genetic variation in Mexican wolves. Journal Of Wildlife Diseases **39**:909-913.

Hedrick, P. W., R. N. Lee, and D. Garrigan. 2002. Major histocompatibility complex variation in red wolves: evidence for common ancestry with coyotes and balancing selection. Molecular Ecology **11**:1905-1913.

Hedrick, P. W., R. N. Lee, and K. M. Parker. 2000. Major histocompatibility complex (MHC) variation in the endangered Mexican wolf and related canids. Heredity **85**:617-624.

Hedrick, P. W., P. S. Miller, E. Geffen, and R. Wayne. 1997. Genetic evaluation of the three captive mexican wolf lineages. Zoo Biology **16**:47-69.

Hedrick, P. W., and K. M. Parker. 1998. MHC variation in the endangered Gila topminnow. Evolution **52**:194-199.

Hedrick, P. W., K. M. Parker, and R. N. Lee. 2001b. Using microsatellite and MHC variation to identify species, ESUs, and MUs in the endangered Sonoran topminnow. Molecular Ecology **10**:1399-1412.

Hedrick, P. W., K. M. Parker, E. L. Miller, and P. S. Miller. 1999. Major Histocompatibility Complex Variation in the Endangered Przewalski's Horse. Genetics **152**:1701-1710.

Heller, R., E. D. Lorenzen, J. B. A. Okello, C. Masembe, and H. R. Siegismund. 2008. Mid-Holocene decline in African buffalos inferred from Bayesian coalescent-based analyses of microsatellites and mitochondrial DNA. Molecular Ecology **17**:4845-4858.

Henry, P., P. Miquelle, T. Sugimoto, D. R. McCullough, A. Caccone, and M. A. Russello. 2009. *In situ* population structure and *ex situ* representation of the endangered Amur tiger. Molecular Ecology **18**:3173-3184.

Hewitt, G. 2004. The structure of biodiversity - insights from molecular phylogeography. Frontiers in Zoology **1**:4.

Higgins, K., and M. Lynch. 2001. Metapopulation extinction caused by mutation accumulation. Proceedings of the National Academy of Sciences **98**:2928-2933.

Hill, A. V. S. 1991. HLA associations with malaria in Africa: some implications for MHC evolution. Pages 403-419 in J. Klein, and D. Klein, editors. Molecular evolution of the major histocompatability complex. Springer, Berlin.

Hill, A. V. S. 1998. The immunogenetics of human infectious diseases. Annual Review of Immunology **16**:593-617.

Hill, W. G. 1981. Estimation of effective population size from data on linkage disequilibrium. Genetical Research **38**:209-216.

HLA-Database. 2010. IMGT HLA Database.

Hofmeyr, M., J. Bingham, E. P. Lane, A. Ide, and L. Nel. 2000. Rabies in African wild dogs (*Lycaon pictus*) in the Madikwe Game Reserve, South Africa. Veterinary Record **146**:50-52.

Hoglund, J. 2009. Evolutionary Conservation Genetics. Oxford University Press, New York.

Holsinger, K. E., and B. S. Weir. 2009. Genetics in geographically structured populations: defining, estimating and interpreting FST. Nature Reviews Genetics **10**:639-650.

Howard, J., P. E. Marinari, and D. E. Wildt. 2003. Black-footed ferret: model for assisted reproductive technologies contributing to *in situ* conservation in W. Holdt, A. R.

Pickard, J. C. Rodger, and D. E. Wildt, editors. Reproductive science and integrated conservation. Cambridge University Press, Cambridge.

Hubisz, M. J., Falush, D., Stephens, M., Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources **9**:1322-1332.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics **17**:754-755.

Hurst, J. L., C. E. Payne, C. M. Nevison, A. D. Marie, R. E. Humphries, D. H. L. Robertson, A. Cavaggioni, and R. J. Beynon. 2001. Individual recognition in mice mediated by major urinary proteins. Nature **414**:631-634.

ISIS. 2010. International Species Information System. ISIS.

IUCN. 2010. IUCN Red List of Threatened Species.

IUCN/SSC. 2008. Regional conservation strategy for the cheetah and wild dog in eastern Africa. IUCN, Gland.

IUCN/SSC. 2009. Regional conservation strategy for the cheetah and wild dog in southern Africa. IUCN, Gland.

Jeffery, K. J. M., and C. R. M. Bangham. 2000. Do infectious diseases drive MHC diversity? Microbes and Infection **2**:1335-1341.

Johnsen, B. O., and A. J. Jensen. 1986. Infestations of Atlantic salmon *Salmo salar* by *Gyrodactylus salaris* in Norwegian rivers. Journal of Fish Biology **29**:233-241.

Jones, A. G., and W. R. Ardren. 2003. Methods of parentage analysis in natural populations. Molecular Ecology **12**:2511-2523.

Jones, M. E., A. Cockburn, R. Hamede, C. Hawkins, H. Hesterman, S. Lachish, D. Mann, H. McCallum, and D. Pemberton. 2008. Life-history change in disease-ravaged Tasmanian devil populations. Proceedings of the National Academy of Sciences **105**:10023-10027.

Jost, L. 2008. Gst and its relatives do not measure differentiation. Molecular Ecology **17**:4015-4026.

Kalinowski, S. T. 2005. hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Molecular Ecology Notes **5**:187-189.

Kat, P., K. Alexander, J. Smith, and L. Munson. 1995. Rabies and African Wild Dogs in Kenya. Proceedings of the Royal Society B: Biological Sciences **262**:229-233.

Keller, L. F. 1998. Inbreeding and Its Fitness Effects in an Insular Population of Song Sparrows (*Melospiza melodia*). Evolution **52**:240-250.

Keller, L. F., K. J. Jeffery, P. Arcese, M. A. Beaumont, W. M. Hochachka, J. N. M. Smith, and M. W. Bruford. 2001. Immigration and the ephemerality of a natural population bottleneck: evidence from molecular markers. Proceedings of the Royal Society of London. Series B: Biological Sciences **268**:1387-1394.

Keller, L. F., and D. M. Waller. 2002. Inbreeding effects in wild populations. Trends in Ecology & Evolution **17**:230-241.

Kennedy, L. J. 2007. 14th International HLA and Immunogenetics Workshop: Report on joint study on canine DLA diversity. Tissue Antigens **69**:269-271.

Kennedy, L. J., J. M. Angles, A. Barnes, L. E. Carmichael, A. D. Radford, W. E. R. Ollier, and G. M. Happ. 2007a. DLA-DRB1, DQA1, and DQB1 Alleles and Haplotypes in North American Gray Wolves. Journal of Heredity **98**:491-499.

Kennedy, L. J., J. M. Angles, A. Barnes, S. D. Carter, O. Francino, J. A. Gerlach, G. M. Happ, W. E. R. Ollier, W. Thomson, and J. L. Wagner. 2001. Nomenclature for factors of the dog major histocompatibility system (DLA), 2000: second report of the ISAG DLA Nomenclature Committee. Animal Genetics **32**:193-199.

Kennedy, L. J., A. Barnes, G. M. Happ, R. J. Quinnell, D. Bennett, J. M. Angles, M. J. Day, N. Carmichael, J. F. Innes, D. Isherwood, S. D. Carter, W. Thomson, and W. E. R. Ollier. 2002a. Extensive interbreed, but minimal intrabreed, variation of DLA class II alleles and haplotypes in dogs. Tissue Antigens **59**:194-204.

Kennedy, L. J., A. Barnes, G. M. Happ, R. J. Quinnell, O. Courtenay, S. D. Carter, W. E. R. Ollier, and W. Thomson. 2002b. Evidence for extensive DLA polymorphism in different dog populations. Tissue Antigens **60**:43-52.

Kennedy, L. J., A. Barnes, W. E. R. Ollier, and M. J. Day. 2006a. Association of a common dog leucocyte antigen class II haplotype with canine primary immune-mediated haemolytic anaemia. Tissue Antigens **68**:502-508.

Kennedy, L. J., S. D. Carter, A. Barnes, S. Bell, D. Bennett, B. Ollier, and W. Thomson. 1999. DLA-DRB1 polymorphisms in dogs defined by sequence-specific oligonucleotide probes (SSOP). Tissue Antigens **53**:184-189.

Kennedy, L. J., S. D. Carter, A. Barnes, S. Bell, D. Bennett, W. E. R. Ollier, and W. Thomson. 2000. DLA-DQA1 polymorphisms in dogs defined by sequence-specific oligonucleotide probes (SSOP). Tissue Antigens **55**:257-261.

Kennedy, L. J., L. J. Davison, A. Barnes, A. D. Short, N. Fretwell, C. A. Jones, A. C. Lee, W. E. R. Ollier, and B. Catchpole. 2006b. Identification of susceptibility and protective major histocompatibility complex haplotypes in canine diabetes mellitus. Tissue Antigens **68**:467-476.

Kennedy, L. J., H. J. Huson, J. Leonard, J. M. Angles, L. E. Fox, J. W. Wojciechowski, C. Yuncker, and G. M. Happ. 2006c. Association of hypothyroid disease in Doberman Pinscher dogs with a rare major histocompatibility complex DLA class II haplotype. Tissue Antigens **67**:53-56.

Kennedy, L. J., M. Lunt, A. Barnes, L. McElhinney, A. R. Fooks, D. N. Baxter, and W. E. R. Ollier. 2007b. Factors influencing the antibody response of dogs vaccinated against rabies. Vaccine **25**:8500-8507.

Kennedy, L. J., S. Quarmby, N. Fretwell, A. J. Martin, P. G. Jones, C. A. Jones, and W. E. R. Ollier. 2005. High-Resolution Characterization of the Canine DLA-DRB1 Locus Using Reference Strand-Mediated Conformational Analysis. Journal of Heredity **96**:836-842.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press.

Klein, J. 1980. Natural History of the Major Histcompatibility Complex. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore.

Klein, J., and V. Horejsi 1997. Immunology. Blackwell Science, Oxford.

Klein, J., and A. Sato. 2000. The HLA system. Part 1. New England Journal of Medicine **343**:702-709.

Klein, J., A. Sato, and N. Nikolaidis. 2007. MHC, TSP, and the Origin of Species: From Immunogenetics to Evolutionary Genetics. Annual Review of Genetics **41**:281-304.

Klein, J., Y. Satta, and C. O'hUigin. 1993. The molecular descent of the major histocompatibility complex. Annual Review of Immunology **11**:269-295.

Kosakovsky Pond, S. L., and S. D. W. Frost. 2005. Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. Molecular Biology and Evolution **22**:1208-1222.

Kosakovsky Pond, S. L., D. Posada, M. B. Gravenor, C. H. Woelk, and S. D. W. Frost. 2006. GARD: a genetic algorithm for recombination detection. Bioinformatics **22**:3096-3098.

Koutsogiannouli, E. A., K. A. Moutou, T. Sarafidou, C. Stamatis, V. Spyrou, and Z. Mamuris. 2009. Major histocompatibility complex variation at class II DQA locus in the brown hare (*Lepus europaeus*). Molecular Ecology **18**:4631-4649.

Kriener, K., C. O'HUigin, H. Tichy, and J. Klein. 2000. Convergent evolution of major histocompatibility complex molecules in humans and New World monkeys. Immunogenetics **51**:169-178.

Lacy, R. C. 1997. Importance of genetic variation to the viability of mammalian populations. Journal of Mammalogy **78**:320-335.

Lande, R. 1988. Genetics and Demography in Biological Conservation. Science **241**:1455-1460.

Landry, C., and L. Bernatchez. 2001. Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). Molecular Ecology **10**:2525-2539.

Langella, O. 1999. POPULATIONS v.1.2.30.

Lavergne, S., and J. Molofsky. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proceedings of the National Academy of Sciences **104**:3883-3888.

Lehman, N., A. Eisenhawer, K. Hansen, L. D. Mech, R. O. Peterson, P. J. P. Gogan, and R. K. Wayne. 1991. Introgression of Coyote Mitochondrial-DNA into Sympatric North-American Gray Wolf Populations. Evolution **45**:104-119.

Lindblad-Toh, K., C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe, M. Kamal, M. Clamp, J. L. Chang, E. J. Kulbokas, M. C. Zody, E. Mauceli, X. Xie, M. Breen, R. K. Wayne, E. A. Ostrander, C. P. Ponting, F. Galibert, D. R. Smith, P. J. deJong, E. Kirkness, P. Alvarez, T. Biagi, W. Brockman, J. Butler, C.-W. Chin, A. Cook, J. Cuff, M. J. Daly, D. DeCaprio, S. Gnerre, M. Grabherr, M. Kellis, M. Kleber, C. Bardeleben, L. Goodstadt, A. Heger, C. Hitte, L. Kim, K.-P. Koepfli, H. G. Parker, J. P. Pollinger, S. M. J. Searle, N. B. Sutter, R. Thomas, C. Webber, and E. S. Lander. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature **438**:803-819.

Lindsey, P. A., R. R. Alexander, J. T. du Toit, and M. G. L. Mills. 2005. The potential contribution of ecotourism to African wild dog *Lycaon pictus* conservation in South Africa. Biological Conservation **123**:339-348.

López, G., M. López-Parra, L. Fernández, C. Martínez-Granados, F. Martínez, M. L. Meli, J. M. Gil-Sánchez, N. Viqueira, M. A. Díaz-Portero, R. Cadenas, H. Lutz, A. Vargas, and M. A. Simón. 2009. Management measures to control a feline leukemia virus outbreak in the endangered Iberian lynx. Animal Conservation **23**.

Lu, Y. 2001. Roles of lineage sorting and phylogenetic relationship in the genetic diversity at the self-incompatibility locus of Solanaceae. Heredity **86**:195-205.

Luikart, G., F. W. Allendorf, J. M. Cornuet, and W. B. Sherwin. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. Journal of Heredity **89**:238-247.

Lynch, M. 1996. A quantitative genetic perpsective on conservation issues. Pages 471-501 in J. C. Avise, and J. L. Hamrick, editors. Conservation genetics: case histories from nature. Springer Verlag, New York.

Macdonald, D. W., and C. Sillero-Zubiri. 2004. Biology and Conservation of Wild Canids. Oxford University Press.

Madsen, T., Stille, B., Shine, R. 1996. Inbreeding depression in an insolated population of adders *Vipera berus*. Biological Conservation **75**:113-118.

Mäkinen, H. S., J. M. Cano, and J. Merilä. 2008. Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. Molecular Ecology **17**:3565-3582.

Malcolm, J. R. 1979. Social organisation and communal rearing in African wild dogs. Harvard University.

Margan, S., H. , R. Nurthen, K. , M. Montgomery, E. , L. Woodworth, M. , E. Lowe, H. , D. Briscoe, A. , and R. Frankham. 1998. Single large or several small? Population fragmentation in the captive management of endangered species. Zoo Biology **17**:467-480.

Marsden, C. D., B. K. Mable, R. Woodroffe, G. S. A. Rasmussen, S. Cleaveland, J. W. McNutt, M. Emmanuel, R. Thomas, and L. J. Kennedy. 2009. Highly Endangered African Wild Dogs (*Lycaon pictus*) Lack Variation at the Major Histocompatibility Complex. Journal of Heredity **100**:S54-65.

Marshall, T. C., and J. A. Spalton. 2000. Simultaneous inbreeding and outbreeding depression in reintroduced Arabian oryx. Animal Conservation **3**:241-248.

Maugham, R. C. F. 1914. Wild game in Zambesia. John Murray, London.

McCarthy, A. J., M.-A. Shaw, and S. J. Goodman. 2007. Pathogen evolution and disease emergence in carnivores. Proceedings of the Royal Society B: Biological Sciences **274**:3165-3174.

McNutt, J. W. 1996. Sex-biased dispersal in African wild dogs, *Lycaon pictus*. Animal Behaviour **52**:1067-1077.

McNutt, J. W., M. G. L. Mills, K. McCreery, G. Rasmussen, R. Robbins, and R. Woodroffe. 2008. *Lycaon pictus*. IUCN 2008. 2008 IUCN Red List of Threatened Species.

Mech, L. D., and L. Boitani. 2008. Canis lupus. In: IUCN Red List of Threatened Species.

Mee, E. T., A. Badhan, J. Karl, R. W. Wiseman, K. Cutler, L. A. Knapp, N. Almond, D. H. O'Connor, and N. J. Rose. 2009. MHC haplotype frequencies in a UK breeding colony of Mauritian cynomolgus macaques mirror those found in a distinct population from the same geographic origin. Journal of Medical Primatology **38**:1-14.

Michel, C., L. Bernatchez, and J. Behrmann-Godel. 2009. Diversity and evolution of MHII [beta] genes in a non-model percid species -The Eurasian perch (*Perca fluviatilis* L.). Molecular Immunology **46**:3399-3410.

Mikko, S., and L. Andersson. 1995. Low major histocompatibility complex class II diversity in European and North American moose. Proceedings of the National Academy of Sciences of the United States of America **92**:4259-4263.

Milinski, M. 2006. The Major Histocompatibility Complex, Sexual Selection, and Mate Choice. Annual Review of Ecology, Evolution, and Systematics **37**:159-186.

Miller, H. C., and D. M. Lambert. 2004. Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). Molecular Ecology **13**:3709-3721.

Miller, K. M., K. H. Kaukinen, T. D. Beacham, and R. E. Withler. 2001. Geographic heterogeneity in natural selection on an MHC locus in sockeye salmon. Genetica **111**:237-257.

Mills, M. G. L., Ellis Susie, Woodroffe Rosie, Maddock Anthony, Stander Phillip, Pole Alistair, Rasmussen Greg, Fletcher Pat, Bruford Mike, Wildt David, M. David, and S. Ulysses. 1998. Population and Habitat Viability Assessment African Wild Dog (*Lycaon pictus*) in Southern Africa. Final Workshop Report. IUCN/SSC Conservation Breeding Specialist Group, Apple Valley, MN.

Mills, M. G. L., and M. L. Gorman. 1997. Factors Affecting the Density and Distribution of Wild Dogs in the Kruger National Park. Conservation Biology **11**:1397-1406.

Minitab, I. 2007. Minitab 15 Statistical Software.

Moran, P. 2002. Current conservation genetics: building an ecological approach to the synthesis of molecular and quantitative genetic methods. Ecology of Freshwater Fish **11**:30-55.

Morell, V. 1995. Wildlife Biology: Dogfight Erupts Over Animal Studies in the Serengeti. Science **270**:1302-1303.

Muirhead, C. A. 2001. Consequences of population structure on genes under balancing selection. Evolution **55**:1532-1541.

Munguia-Vega, A., Y. Esquer-Garrigos, L. Rojas-Bracho, R. Vazquez-Juarez, A. Castro-Prieto, and S. Flores-Ramirez. 2007. Genetic drift vs. natural selection in a long-term small isolated population: major histocompatibility complex class II variation in the Gulf of California endemic porpoise (*Phocoena sinus*). Molecular Ecology **16**:4051-4065.

Neff, M. W., K. W. Broman, C. S. Mellersh, K. Ray, G. M. Acland, G. D. Aguirre, J. S. Ziegle, E. A. Ostrander, and J. Rine. 1999. A Second-Generation Genetic Linkage Map of the Domestic Dog, *Canis familiaris*. Genetics **151**:803-820.

Novembre, J., and M. Stephens. 2008. Interpreting principal component analyses of spatial population genetic variation. Nature Genetics **40**:646-649.

Nozawa, M., Y. Suzuki, and M. Nei. 2009. Reliabilities of identifying positive selection by the branch-site and the site-prediction methods. Proceedings of the National Academy of Sciences **106**:6700-6705.

Nsubuga, A., J. Holzman, L. Chemnick, and O. Ryder. 2010. The cryptic genetic structure of the North American captive gorilla population. Conservation Genetics **11**:161-172.

O'Brien, S. J. 1994. A Role for Molecular Genetics in Biological Conservation. Proceedings of the National Academy of Sciences **91**:5748-5755.

O'Brien, S. J., and J. F. Evermann. 1988. Interactive influence of infectious disease and genetic diversity in natural populations. Trends in Ecology & Evolution **3**:254-259.

Okello, J. B. A., G. Wittemyer, H. B. Rasmussen, P. Arctander, S. Nyakaana, I. Douglas-Hamilton, and H. R. Siegismund. 2008. Effective population size dynamics reveal impacts of historic climatic events and recent anthropogenic pressure in African elephants. Molecular Ecology **17**:3788-3799.

Oliver, M., K., , X. Lambin, T. Cornulier, and S. B. Piertney. 2009a. Spatio-temporal variation in the strength and mode of selection acting on major histocompatibility complex diversity in water vole (*Arvicola terrestris*) metapopulations. Molecular Ecology **18**:80-92.

Oliver, M. K., S. Telfer, and S. B. Piertney. 2009b. Major histocompatibility complex (MHC) heterozygote superiority to natural multi-parasite infections in the water

vole (*Arvicola terrestris*). Proceedings of the Royal Society B: Biological Sciences **276**:1119-1128.

Otting, N., N. de Groot, G. Doxiadis, and R. Bontrop. 2002. Extensive Mhc-DQB variation in humans and non-human primate species. Immunogenetics **54**:230-239.

Packer, C., R. Hilborn, A. Mosser, B. Kissui, M. Borner, G. Hopcraft, J. Wilmshurst, S. Mduma, and A. R. E. Sinclair. 2005. Ecological Change, Group Territoriality, and Population Dynamics in Serengeti Lions. Science **307**:390-393.

Paetkau, D., R. Slade, M. Burden, and A. Estoup. 2004. Direct, real-time estimation of migration rate using assignment methods: a simulation-based exploration of accuracy and power Molecular Ecology **13**:55-65.

Palsboll, P. J., M. Berube, and F. W. Allendorf. 2007. Identification of management units using population genetic data. Trends in Ecology & Evolution **22**:11-16.

Paradis, E. 2006. Analysis of Phylogenetics and Evolution with R. Springer, New York.

Peakall, R., and P. E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes **6**:288-295.

Peel, D., J. R. Ovenden, and S. L. Peel. 2004. NeEstimator: software for estimating effective population size, Version 1.3. Queensland Government, Department of Primary Industries and Fisheries.

Penn, D., and W. Potts. 1998. MHC-disassortative mating preferences reversed by cross-fostering. Proceedings Of The Royal Society Of London Series B-Biological Sciences **265**:1299-1306.

Perini, F. A., C. A. M. Russo, and C. G. Schrago. 2009. The evolution of South American endemic canids: a history of rapid diversification and morphological parallelism. Journal of Evolutionary Biology **23**:311-322.

Piertney, S. B., and M. K. Oliver. 2006. The evolutionary ecology of the major histocompatibility complex. Heredity **96**:7-21.

Piry, S., A. Alapetite, J. M. Cornuet, D. Paetkau, L. Baudouin, and A. Estoup. 2004. GENECLASS2: A Software for Genetic Assignment and First-Generation Migrant Detection. Journal of Heredity **95**:536-539.

Piry, S., G. Luikart, and J. M. Cornuet. 1999. Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. Journal of Heredity **90**:502-503.

Pole, A. 2000. The behaviour and ecology of African wild dogs (Lycaon pictus) in an environment with reduced competitor density Department of Zoology. University of Aberdeen, Aberdeen.

Pollack, J. P., J. D. Ballou, and R. C. Lacy. 2001. Population management 2000 v1.213. Chicago Zoological Society, Brookfield, IL.

Pond, S. L. K., S. D. W. Frost, and S. V. Muse. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics **21**:676-679.

Poon, A. F. Y., S. D. W. Frost, and S. L. K. Pond. 2009. Detecting Signatures of Selection from DNA Sequences Using Datamonkey. Pages 163-183. Bioinformatics for DNA Sequence Analysis.

Posada, D., and K. A. Crandall. 1998. Modeltest testing the model of DNA substitution. Bioinformatics **14**:817-818.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics **155**:945-959.

Pudovkin, A. I., D. V. Zaykin, and D. Hedgecock. 1996. On the potential for estimating the effective number of breeders from heterozygote excess in progeny. Genetics **144**:383-387.

Qiagen. 2006. Purification of total DNA from nails, hair, or feathers using the DNeasy® Blood & Tissue Kit.

Quinnell, R. J., L. J. Kennedy, A. Barnes, O. Courtenay, C. Dye, L. M. Garcez, M. A. Shaw, S. D. Carter, W. Thomson, and W. E. R. Ollier. 2003. Susceptibility to visceral leishmaniasis in the domestic dog is associated with MHC class II polymorphism. Immunogenetics **55**:23-28.

Radwan, J., A. Biedrzycka, and W. Babik. 2010. Does reduced MHC diversity decrease viability of vertebrate populations? Biological Conservation **143**:537-544.

Radwan, J., A. Kawalko, J. M. Wojcik, and W. Babik. 2007. MHC-DRB3 variation in a free-living population of the European bison, *Bison bonasus*. Molecular Ecology **16**:531-540.

Ralls, K., J. D. Ballou, B. A. Rideout, and R. Frankham. 2000. Genetic management of chondrodystrophy in California condors. Animal Conservation **3**:145-153.

Ramirez, O., L. Altet, C. Enseñat, C. Vilà, A. Sanchez, and A. Ruiz. 2006. Genetic assessment of the Iberian wolf *Canis lupus signatus* captive breeding program. Conservation Genetics **V7**:861-878.

Randall, D., J. Pollinger, K. Argaw, D. Macdonald, and R. Wayne. 2010. Fine-scale genetic structure in Ethiopian wolves imposed by sociality, migration, and population bottlenecks. Conservation Genetics **11**:89-101.

Randall, D., S. D. Williams, I. V. Kuzmin, C. E. Rupprecht, L. A. Tallents, Z. Tefera, K. Argaw, F. Shiferaw, D. L. Knobel, C. Sillero-Zubiri, and M. K. Laurenson. 2004. Rabies in endangered Ethiopian wolves. Emerging Infectious Diseases **10**:2214-2217.

Randall, D. A., J. Marino, D. T. Haydon, C. Sillero-Zubiri, D. L. Knobel, L. A. Tallents, D. W. Macdonald, and M. K. Laurenson. 2006. An integrated disease management strategy for the control of rabies in Ethiopian wolves. Biological Conservation **131**:151-162.

Randall, D. A., J. P. Pollinger, D. W. Macdonald, D. T. Haydon, and R. K. Wayne. 2007. Fine-scale genetic structure imposed by sociality and population bottlenecks among Ethiopian wolves in the Bale Mountains. Submitted.

Rannala, B., and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences USA **94**:9197-9201.

Rasmussen, G. 2010. International trade in Painted dogs.

Raymond, M., and F. Rousset. 1995. GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. Journal of Heredity **86**:248-249.

Reed, D. H., and R. Frankham. 2003. Correlation between fitness and genetic diversity. Conservation Biology **17**:230-237.

Reznick, D. N., F. H. Shaw, F. H. Rodd, and R. G. Shaw. 1997. Evaluation of the Rate of Evolution in Natural Populations of Guppies (*Poecilia reticulata*). Science **275**:1934-1937.

Richman, A. 2000. Evolution of balanced genetic polymorphism. Molecular Ecology **9**:1953-1963.

Riley, S. P. D., J. Pollinger, P., R. Sauvajot, M., E. York, C., C. Bromley, T. Fuller, K., and R. Wayne, K. . 2006. A southern California freeway is a physical and social barrier to gene flow in carnivores. Molecular Ecology **15**:1733-1741.

Roca, A. L., N. Georgiadis, and S. J. O'Brien. 2005. Cytonuclear genomic dissociation in African elephant species. Nature Genetics **37**:96-100.

Roelke, M. E., J. S. Martenson, and S. J. O'Brien. 1993. The consequences of demographic reduction and genetic depletion in the endangered Florida panther. Current Biology **3**:340-350.

Rozas, J., and R. Rozas. 1995. DnaSP, DNA sequence polymorphism: an interactive program for estimating population genetics parameters from DNA sequence data. Computer Applications in the Biosciences **11**:621-625.

Saccheri, I., M. Kuussaari, M. Kankare, P. Vikman, W. Fortelius, and I. Hanski. 1998. Inbreeding and extinction in a butterfly metapopulation. Nature **392**:491-494.

Sachdev, M., R. Sankaranarayanan, P. Reddanna, K. Thangaraj, and L. Singh. 2005. Major histocompatibility complex class I polymorphism in Asiatic lions. Tissue Antigens **66**:9-18.

Sanchez-Mazas, A., S. Djoulah, M. Busson, I. L. de Gouville, J. C. Poirier, C. Dehay, D. Charron, L. Excoffier, S. Schneider, A. Langaney, J. Dausset, and J. Hors. 2000. A linkage disequilibrium map of the MHC region based on the analysis of 14loci haplotypes in 50 French families. European Journal of Human Genetics **8**:33-41.

Satta, Y., C. O'HUigin, N. Takahata, and J. Klein. 1994. Intensity of natural selection at the major histocompatibility complex locus. Proceedings of the National Academy of Sciences **91**:7184-7188.

Scheffler, K., D. P. Martin, and C. Seoighe. 2006. Robust inference of positive selection from recombining coding sequences. Bioinformatics **2006**:2493-2499.

Schierup, M. H., X. Vekemans, and D. Charlesworth. 2000. The effect of subdivision on variation at multi-allelic loci under balancing selection. Genetical Research **76**:51-62.

Schmidt, P. S., E. A. Serrão, G. A. Pearson, C. Riginos, P. D. Rawson, T. J. Hilbish, S. H. Brawley, G. C. Trussell, E. Carrington, D. S. Wethey, J. W. Grahame, F. Bonhomme, and D. M. Rand. 2008. Ecological genetics in the North Atlantic: environmental gradients and adaptation at specific loci. Ecology **89**:S91-S107.

Schwartz, M. K., D. A. Tallmon, and G. Luikart. 1998. Review of DNA-based census and effective population size estimators. Animal Conservation **1**:293-299.

Schwensow, N., M. Eberle, and S. Sommer. 2010. Are there unbiquitous parasite-driven MHC selection mechanisms in gray mouse lemurs? International Journal of Primatology **In press**.

Scott, J. 1991. Painted wolves: Wild dogs of the Serengeti-Mara. Hamish Hamilton, London.

Seddon, J. M., and H. Ellegren. 2002. MHC class II genes in European wolves: a comparison with dogs Immunogenetics **54**:490-500.

Seddon, J. M., and H. Ellegren. 2004. A temporal analysis shows major histocompatibility complex loci in the Scandinavian wolf population are consistent with neutral evolution. Proceedings of the Royal Society of London Series B-Biological Sciences **271**:2283-2291.

Shen, F., ., Z. Zhihe, H. E. Wei, Y. U. E. Bisong, Z. Anju, Z. Liang, H. O. U. Rong, W. Chengdong, and W. Toshi. 2009. Microsatellite variability reveals the necessity for genetic input from wild giant pandas (*Ailuropoda melanoleuca*) into the captive population. Molecular Ecology **18**:1061-1070.

Siddle, H. V., C. Sanderson, and K. Belov. 2007. Characterization of major histocompatibility complex class I and class II genes from the Tasmanian devil (*Sarcophilus harrisii*). Immunogenetics **59**:753-760.

Sillero-Zubiri, C., M. Hoffman, and D. W. Macdonald, editors. 2004. Canids: Foxes, Wolves, Jackals and Dogs. Status Survey and Action Plan. IUCN, Gland, Switzerland.

Slade, R. W., and H. I. McCallum. 1992. Overdominant vs. Frequency-Dependent Selection at MHC Loci. Genetics **132**:861-862.

Slev, P. R., A. C. Nelson, and W. K. Potts. 2006. Sensory neurons with MHC-like peptide binding properties: disease consequences. Current Opinion in Immunology **18**:608-616.

Smith, K. F., Acevedo-Whitehouse K., Pedersen, A. B. 2009. The role of infectious diseases in biological conservation. Animal Conservation **12**:1-12.

Smith, S., K. Belov, and J. Hughes. 2010. MHC screening for marsupial conservation: extremely low levels of class II diversity indicate population vulnerability for an endangered Australian marsupial. Conservation Genetics **11**:269-278.

Smouse, P. E., and R. Peakall. 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity **82**:561-573.

Sokal, R. R., and F. J. Rohlf 1994. Biometry: The Principles and Practices of Statistics in Biological Research. W. H. Freeman.

Sommer, S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Frontiers in Zoology **2**:16.

Sommer, S., D. Schwab, and J. U. Ganzhorn. 2002. MHC diversity of endemic Malagasy rodents in relation to geographic range and social system. Behavioral Ecology and Sociobiology **51**:214-221.

Spurgin, L. G., and D. S. Richardson. 2010. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proceedings of the Royal Society B: Biological Sciences **277**:979-988.

St. Sauver, J. L., N. Dhiman, I. G. Ovsyannikova, R. M. Jacobson, R. A. Vierkant, V. S. Pankratz, S. J. Jacobsen, and G. A. Poland. 2005. Extinction of the Human Leukocyte Antigen Homozygosity Effect After Two Doses of the Measles-Mumps-Rubella Vaccine. Human Immunology **66**:788-798.

Stockwell, C. A., A. P. Hendry, and M. T. Kinnison. 2003. Contemporary evolution meets conservation biology. Trends in Ecology & Evolution **18**:94-101.

Stokstad, E. 2005. Ecology - Taking the pulse of earth's life-support systems. Science **308**:41-43.

Storz, J. F., and M. A. Beaumont. 2002. Testing for genetic evidence of population expansion and contraction: An empirical analysis of microsatellite DNA variation using a hierarchical Bayesian method. Evolution **56**:154-166.

Suárez, C., M. Patarroyo, E. Trujillo, M. Estupiñán, J. Baquero, C. Parra, and R. Rodriguez. 2006. Owl monkey MHC-DRB exon 2 reveals high similarity with several HLA-DRB lineages. Immunogenetics **58**:542-558.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. Molecular Biology and Evolution **24**:1596-1599.

Tan, P.-L., R. M. Jacobson, G. A. Poland, S. J. Jacobsen, and V. S. Pankratz. 2001. Twin studies of immunogenicity - determining the genetic contribution to vaccine failure. Vaccine **19**:2434-2439.

Templeton, A. R. 1986. Coadaptation and outbreeding depression. Pages 105-116 in M. Soule, editor. Conservation biology: The Science of Scarcity and Diversity. Sinauer Associates, Sunderland, MA.

Templeton, A. R., K. Shaw, E. Routman, and S. K. Davis. 1990. The Genetic Consequences of Habitat Fragmentation. Annals of the Missouri Botanical Garden **77**:13-27.

Thevenon, S., and D. Couvet. 2002. The impact of inbreeding depression on population survival depending on demographic parameters. Animal Conservation **5**:53-60.

Thirgood, S. J., Mduma, S.A. R., Keyyu JD., Laurenson, M. K. . 2007. The value of long-term research: conservation science and practice in Tanzania. Conservation Biology **21**:576-579.

Thom, M. D., P. Stockley, F. Jury, W. E. R. Ollier, R. J. Beynon, and J. L. Hurst. 2008. The Direct Assessment of Genetic Heterozygosity through Scent in the Mouse. Current Biology **18**:619-623.

Thomassen, R., and W. Farstad. 2009. Artificial insemination in canids: A useful tool in breeding and conservation. Theriogenology **71**:190-199.

UNEP. 2008. Proposals for amendment of the appendicies I and II for consideration by the ninth meeding of the conference of the parties. Convention on the Conservation of Migratory Species of Wild Animals, Rome.

van de Bildt, M., W, G , T. Kuiken, A. Visee, M , S. Lema, T. Fitzjohn, R, and A. Osterhaus, D, M, E. . 2002. Distemper outbreak and its effect on African wild dog conservation. Emerging infectious diseases **8**:211-213.

van Den Bussche, R. A., S. R. Hoofer, and R. L. Lochmiller. 1999. Characterization of Mhc-DRB allelic diversity in white-tailed deer (Odocoileus virginianus) provides insight into Mhc-DRB allelic evolution within Cervidae. Immunogenetics **49**:429-437.

van Heerden, J., R. Vester, M. L. Penrith, and I. Espie. 1996. Disease and mortality in captive wild dogs *Lycaon pictus*. Journal of South African Veterinary Association **67**:141-145.

van Oosterhout, C., D. A. Joyce, S. M. Cummings, J. Blais, N. J. Barson, I. W. Ramnarine, R. S. Mohammed, N. Persad, J. Cable, and J. Hey. 2006. Balancing selection, random genetic drift and genetic variation at the major histocompatibility complex in two wild populations of guppies (*Poecilia reiculata*). Evolution **60**:2562-2574.

Vassilakos, D., A. Natoli, M. Dahlheim, and A. R. Hoelzel. 2009. Balancing and Directional Selection at Exon-2 of the MHC DQB1 Locus among Populations of Odontocete Cetaceans. Molcular Biology and Evolution **26**:681-689.

Verberkmoes, W., and H. Verberkmoes. 2008. European regional African Wild Dog Stud book. Gaiapark, Kerkrade Zoo, Kerkrade.

Verginelli, F., C. Capelli, V. Coia, M. Musiani, M. Falchetti, L. Ottini, R. Palmirotta, A. Tagliacozzo, I. De Grossi Mazzorin, and R. Mariani-Costantini. 2005. Mitochondrial DNA from Prehistoric Canids Highlights Relationships Between Dogs and South-East European Wolves. Mol Biol Evol **22**:2541-2551.

Vernesi, C., M. W. Bruford, G. Bertorelle, E. Pecchioli, A. Rizzoli, and H. C. Hauffe. 2008. Where's the Conservation in Conservation Genetics? Conservation Biology **22**:802-804.

Vilà, C., P. Savolainen, and J. E. Maldonado. 1997. Multiple and ancient origins of the domestic dog. Science **276**:1687-1689.

Vilà , C., A.-K. Sundqvist, Ø. Flagstad, J. Seddon, S. Björnerfeldt, I. Kojola, A. Casulli, H. Sand, P. Wabakken, and H. Ellegren. 2003. Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. Proceedings of the Royal Society of London. Series B: Biological Sciences **270**:91-97.

vonHoldt, B. M., J. P. Pollinger, K. E. Lohmueller, E. Han, H. G. Parker, P. Quignon, J. D. Degenhardt, A. R. Boyko, D. A. Earl, A. Auton, A. Reynolds, K. Bryc, A. Brisbin, J. C. Knowles, D. S. Mosher, T. C. Spady, A. Elkahloun, E. Geffen, M. Pilot, W. Jedrzejewski, C. Greco, E. Randi, D. Bannasch, A. Wilton, J. Shearman, M. Musiani, M. Cargill, P. G. Jones, Z. Qian, W. Huang, Z.-L. Ding, Y.-p. Zhang, C. D. Bustamante, E. A. Ostrander, J. Novembre, and R. K. Wayne. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature **464**:898-902.

Wagner, J. L. 2003. Molecular Organization of the Canine Major Histocompatibility Complex. Journal of Heredity **94**:23-26.

Wagner, J. L., R. C. Burnett, S. A. DeRose, and R. Storb. 1996a. Molecular analysis and polymorphism of the DLA-DQA gene. Tissue Antigens **48**:199-204.

Wagner, J. L., R. C. Burnett, and R. Storb. 1999. Organization of the canine major histocompatibility complex: current perspectives. Journal of Heredity **90**:35-38.

Wagner, J. L., R. C. Burnett, J. D. Works, and R. Storb. 1996b. Molecular analysis of DLA-DRBB1 polymorphism. Tissue Antigens **48**:554-561.

Wagner, J. L., S. A. Derose, R. C. Burnett, and R. Storb. 1995. Nucleotide-Sequence and Polymorphism Analysis of Canine Dra Cdna Clones. Tissue Antigens **45**:284-287.

Wakeland, E. K., S. Boehme, and J. X. She. 1990. Ancestral polymorphisms of MHC class II genes: divergent allele advantage. Immunological Research **9**:115-122.

Wall, D., H. Mooney, G. Adams, G. Boxshall, A. Dobson, T. Nakashizuka, J. Seyani, C. Samper, and J. Sarukhan. 2001. An International Biodiversity Observation Year. Trends in Ecology & Evolution **16**:52-54.

Wan, Q.-H., L. Zhu, H. U. A. Wu, and S.-G. Fang. 2006. Major histocompatibility complex class II variation in the giant panda (Ailuropoda melanoleuca). Molecular Ecology **15**:2441-2450.

Waples, R. S., and M. Yokota. 2007. Temporal Estimates of Effective Population Size in Species With Overlapping Generations. Genetics **175**:219-233.

Watts, C. 1997. Capture and processing of exogenous antigens for presentation on MHC molecules. Annual Review of Immunology **15**:821-850.

Wayne, R. K., E. Geffen, D. J. Girman, K. P. Koepfli, L. Lau, M, and C. R. Marshall. 1997. Molecular systematics of the Canidae. Systematic biology **46**:622-653.

Wenink, P. W., Groen, A. F., Roelke-Parker, M. E., Prins, H. H. T. . 1998. African buffalo maintain high genetic diversity in the major histocompatibility complex in spite of historically known population bottlenecks. Molecular Ecology **7**:1315-1322.

Westerdahl, H., B. Hansson, S. Bensch, and D. Hasselquist. 2004. Between-year variation of MHC allele frequencies in great reed warblers: selection or drift? Journal of Evolutionary Biology **17**:485-492.

Whitlock, M. C., and D. E. McCaultey. 1999. Indirect measures of gene flow and migration: Fst ≠ 1(4Nm+1). Heredity **82**:117-125.

Williamson-Natesan, E. 2005. Comparison of methods for detecting bottlenecks from microsatellite loci. Conservation Genetics **6**:551-562.

Wilson, G. A., and B. Rannala. 2003. Bayesian Inference of Recent Migration Rates Using Multilocus Genotypes. Genetics **163**:1177-1191.

Wisely, S. M., D. B. McDonald, and S. W. Buskirk. 2003. Evaluation of the genetic management of the endangered black-footed ferret (*Mustela nigripes*). Zoo Biology **22**:287-298.

Woodroffe, R. 2001. Assessing the risks of intervention: immobilization, radio-collaring and vaccination of African wild dogs. Oryx **35**:234-244.

Woodroffe, R. 2010. Demography of a recovering African wild dog (*Lycaon pictus*) population. Submitted.

Woodroffe, R., K. Chapman, and E. Lemusana. 2009. Solitary breeding in an African wild dog (*Lycaon pictus*). African Journal of Ecology **47**:790-791.

Woodroffe, R., S. Cleaveland, O. Courtenay, M. K. Laurenson, and M. Artois. 2004a. Infections disease in the management and conservation of wild canids. Pages 123-142 in D. W. Macdonald, and C. Sillero-Zubiri, editors. Biology and conservation of wild canids. Oxford University Press.

Woodroffe, R., H. Davies-Mostert, J. Ginsberg, J. Graf, K. Leigh, K. McCreery, R. Robbins, G. Mills, A. Pole, G. Rasmussen, M. Somers, and M. Szykman. 2007a. Rates and causes of mortality in Endangered African wild dogs *Lycaon pictus*: lessons for management and monitoring. Oryx **41**:215-223.

Woodroffe, R., and J. Ginsberg. 1997. Past and future causes of wild dogs' population decline. Pages 58-73 in R. Woodroffe, J. Ginsberg, and D. W. Macdonald, editors. The African Wild Dog: Status Survey and Action Plan. IUCN, Gland, Switzerland.

Woodroffe, R., J. Ginsberg, and D. W. Macdonald, editors. 1997. The African Wild Dog: Status Survey and Conservation Action Plan, Gland, Switzerland. World Conservation Union.

Woodroffe, R., and J. Ginsberg, R. . 1999. Conserving the African wild dog (*Lycaon pictus*). I. Diagnosing and treating causes of decline. Oryx **33**:132-142.

Woodroffe, R., and J. R. Ginsberg. 1998. Edge Effects and the Extinction of Populations Inside Protected Areas. Science **280**:2126-2128.

Woodroffe, R., P. Lindsey, S. Romañach, A. Stein, and S. M. K. ole Ranah. 2005. Livestock predation by endangered African wild dogs (*Lycaon pictus*) in northern Kenya. Biological Conservation **124**:225-234.

Woodroffe, R., P. A. Lindsey, S. S. Romanach, and S. Ranah. 2007b. African wild dogs (*Lycaon pictus*) can subsist on small prey: Implications for conservation. Journal of Mammalogy **88**:181-193.

Woodroffe, R., J. W. McNutt, and M. G. L. Mills. 2004b. African wild dog. Pages 174-183 in C. Sillero-Zubiri, and D. W. Macdonald, editors. Foxes, Wolves, Jackals and Dogs: Status Survey and Conservation Action Plan. IUCN, Gland, Swizerland.

Wright, S. 1931. Evolution in Mendelian populations. Genetics **16**:97-159.

Xu, S., W. Ren, S. Li, F. Wei, K. Zhou, and G. Yang. 2009. Sequence Polymorphism and Evolution of Three Cetacean MHC Genes. Journal of Molecular Evolution **69**:260-275.

Yeager, M., and A. Hughes, L. 1999. Evolution of the mammalian MHC: natural selection, recombination, and convergent evolution. Immunological Reviews **167**:45-58.

Yuhki, N., T. Beck, R. Stephens, B. Neelam, and S. J. O'Brien. 2007. Comparative Genomic Structure of Human, Dog, and Cat MHC: HLA, DLA, and FLA. Journal of Heredity **98**:390-399.

Zhai, W., R. Nielsen, and M. Slatkin. 2009. An Investigation of the Statistical Power of Neutrality Tests Based on Comparative and Population Genetic Data. Molecular Biology and Evolution **26**:273-283.

Zhang, B., S. G. Fang, and Y. M. Xi. 2006. Major histocompatibility complex variation in the endangered crested ibis *Nipponia nippon* and implications for reintroduction. Biochemical Genetics **44**:113-123.

Zhang, B., M. Li, Z. Zhang, B. Goossens, L. Zhu, S. Zhang, J. Hu, M. W. Bruford, and F. Wei. 2007. Genetic Viability and Population History of the Giant Panda, Putting an End to the "Evolutionary Dead End"? Molecular Biology and Evolution **24**:1801-1810.

Zhu, L., X. Ruan, Y. Ge, Q. Wan, and S. Fang. 2007. Low major histocompatibility complex class II DQA diversity in the Giant Panda (*Ailuropoda melanoleuca*). BMC Genetics **8**:315.