Associations between MHC class I variation and blood pathogen prevalence in caracal

Alicia Ly

March 2017

Submitted in partial fulfilment of the requirements for the degree of Masters of Science in Conservation Biology

Supervisor:
Dr. Jacqueline M. Bishop

Percy FitzPatrick Institute of African Ornithology
University of Cape Town
Rondebosch, 7701 South Africa

Table of Contents

List of Figures

List of Tables

Acknowledgements

I wish to express sincere thanks to my supervisor, Jacqueline M. Bishop, for all of the support and assistance throughout this project. Your unwavering encouragement was greatly appreciated and incredibly inspiring.

Thank you to Storme Viljoen for guiding me throughout my lab work and contributing to the pathogen data used in this study.

I would like to acknowledge the support of the Urban Caracal Project. Thank you to Laurel Serieys, Joleen Broadfield, and all of the project volunteers for contributing to the collection of the samples used in this study and introducing me to caracal conservation.

I appreciate the contributions of Arjun Amar to the statistical analyses used in this study.

Finally, thank you to my classmates of CB2016 for your encouragement throughout the past year.

Abstract

Adaptive genetic variability is vital to long-term species survival, as it presents the potential for evolutionary adaptive responses to environmental change. Genes of the Major Histocompatibility Complex (MHC) trigger the vertebrate adaptive immune response to pathogens through the recognition and presentation of foreign peptides, and thus provide an informative genetic marker for studying the adaptive potential of species with respect to disease. MHC class I loci mediate the immune system's recognition of intracellular pathogens, including protozoans, viruses and bacteria, and the high levels of genetic diversity reported at these loci is thought to be primarily the result of pathogen-mediated selection. Although variation within a number of MHC genes has been described in many felid species, the relationship between MHC diversity and pathogen prevalence within wild felid populations has rarely been examined previously. Because many wild felids are highly sensitive to the effects of anthropogenic-induced environmental changes, evolutionarily relevant adaptive genetic variation is particularly important for their conservation. This study explored the relationship between MHC class I exon 2 and exon 3 diversity and a number of tick-borne blood pathogens in a population of caracal (*Caracal caracal*) in the urban landscape of Cape Town, South Africa. The characterization of MHC class I diversity identified 38 and 45 unique, putatively functional alleles in the population, from exon 2 and exon 3 respectively. The influence of MHC allelic diversity and specific alleles, together with sex, age class, home range size and urban cover within home range, on individual pathogen prevalence was assessed using Generalized Linear Models. A positive relationship between *Babesia* infection and multiple measures of MHC diversity, including nucleotide diversity, average number of nucleotide differences and number of exon 3 alleles, was detected. Additionally, the presence of two specific alleles was significantly correlated with *Babesia* infection. High levels of infection by a number of tick-borne pathogens were observed in the population, suggesting environmental factors also have an important influence on pathogen prevalence. However, home range and urban cover within home range, as well as sex and age, were not significantly associated with *Babesia* infection. This study provides the first assessment of the relationship between MHC class I adaptive genetic variation and individual pathogen prevalence in caracal.

Keywords: major histocompatibility complex; MHC class I; caracal; pathogen-mediated selection; *Babesia*

Introduction

Adaptive genetic variation in a changing world

Anthropogenic activities including urbanization, habitat alteration, and overexploitation, are increasingly impacting all levels of biological diversity, across the globe (Chapin et al. 2000; Crutzen & Stoermer 2000; Garner et al. 2005). In recent decades rapid human population growth, together with the associated expansion of urban areas, has accelerated the rate of environmental modification and subsequent loss of biodiversity (Vitousek et al. 1997; Ceballos et al. 2015). Habitat loss and fragmentation (Fahrig 2003), climate change (Thomas et al. 2004; Urban 2015), and shifting patterns of disease (Altizer et al. 2003) are among the effects of these anthropogenic activities, and pose significant extinction risks to many species. Anthropogenic-induced changes to natural populations also contribute to the loss of genetic diversity, an evolutionarily and ecologically important trait that can be monitored at the level of individuals, populations and species (Garner et al. 2005). In a changing world, the capacity for species to adapt is essential for their persistence (Lande & Shannon 1996; Meyers & Bull 2002; Boutin & Lane 2014; Carlson et al. 2014). All evolutionarily adaptive responses to environmental fluctuations are ultimately dependent on genetic variation (Burger & Lynch 1995; Lacy 1997; Hoffmann & Sgrò 2011; Hansen et al. 2012). Thus, genetic variability, and its conservation, is vital to adaptation and long-term species survival (Lande & Shannon 1996; Lacy 1997; Frankham 2005; Pertoldi et al. 2007).

Populations that experience rapid decline in numbers often undergo a concomitant loss of genetic variation, as quantified by measures of heterozygosity, general polymorphism and allelic diversity (Lande & Shannon 1996; Lacy 1997; Frankham 2005; Pertoldi et al. 2007). Losses of genetic variability can result from decreased gene flow, increased inbreeding and/or increased genetic drift through population fragmentation and increased genetic substructure (Stevens et al. 1997; Frankham 2005). The effects of such loss vary depending on the type of variation lost: neutral or adaptive (van Tienderen et al. 2002; Holderegger et al 2006). Variation within neutral (non-coding) gene regions does not necessarily cause direct fitness effects, whereas the loss of adaptive genetic variation, i.e. variation under selection, can directly affect individual fitness. Reduced adaptive genetic variation can contribute to short-term decline in individual fitness (Roelke et al. 1993), as well as long-term decrease in a population's adaptive potential (O'Brien et al. 1985; Lande & Shannon 1996; Lacy 1997; Meyers & Bull 2002; Frankham 2005), and thereby contribute to the general decline in the

size of a population. For example, the vulnerability of genetically homogeneous populations to widespread pathogen outbreaks is commonly recognized (Altizer et al. 2003; King & Lively 2012). Genetically similar individuals are more likely to transmit infections between themselves, while genetically dissimilar individuals are more likely to have resistant genotypes that can stop pathogens from multiplying, reproducing and/or spreading throughout a population (King & Lively 2012). Thus, genetic diversity plays a key role in individual disease resistance and in buffering populations against epidemics (Altizer et al. 2003). The maintenance of adaptive genetic variation is thus critical for species conservation in a rapidly changing world undergoing habitat fragmentation, urban development, and shifting disease patterns (Lande & Shannon 1996; Lacy 1997; Sommer 2005).

For the past two decades quantification of genetic variation at neutral markers has generally been the dominant focus of ecological studies (Sommer 2005; Holderegger et al. 2006). Because natural selection does not act upon these neutral markers, it is challenging to find causal associations between measures of neutral genetic variation and population and individual fitness however (Bekessy et al. 2003; Holderegger et al. 2006). Frequently utilized neutral gene markers include those within non-coding single nucleotide polymorphisms, together with microsatellite and minisatellite loci (van Tienderen et al. 2002; Holderegger et al. 2006). While these markers provide valuable information for population genetic, phylogeographic and phylogenetic reconstructions, they are limited for understanding selection pressures and adaptive potential (van Tienderen et al. 2002). Because coding markers are under selection (purifying, balancing and/or positive), variation within these regions reflects recent and historic adaptive and evolutionary processes, e.g. pathogen-mediated selection on genes of the immune system (van Tienderen et al. 2002; Sommer 2005). Coding-regions that are under positive selection can also reveal detectable changes over shorter timescales than neutral markers (van Tienderen et al. 2002). In vertebrates, the assemblage of genes that form the adaptive immune system, known as the Major Histocompatibility Complex (MHC) is amongst the most diverse in the vertebrate genome, and provides researchers with a highly informative region for the study of adaptive potential in declining and small populations (Sommer 2005; Spurgin & Richardson 2010; Blanchong et al. 2016).

Adaptive immune diversity of the MHC and its association with pathogens

The adaptive immune system plays a crucial role in vertebrate defence against pathogen-mediated diseases, and thereby the long-term persistence of populations and species (Spielman et al. 2004; Sommer 2005). As a subsystem of the overall immune system, the adaptive immune system specializes in differentiating 'self' from 'non-self' (foreign) molecules, and provides a mechanism for immunological memory of infections (Flajnik & Kasahara 2001). The MHC consists of a collection of closely linked genes that contribute to the adaptive immune system's recognition of foreign pathogens (Simpson 1988; Edwards & Hedrick 1998; Flajnik & Kasahara 2001). MHC genes encode glycoprotein receptors, which bind and present peptides to T-lymphocytes on the surface of nucleated cells. When these receptors are bound to foreign pathogen-derived peptides, the adaptive immune system is signalled to initiate an appropriate immune response (Simpson 1988). MHC genes are divided into two main types, class I and class II, which mediate the immune system's recognition and destruction of intracellular and extracellular pathogens respectively (Edwards & Hedrick 1998; Flajnik & Kasahara 2001). Pathogen and self-derived peptides are presented to the immune system by the receptor's functionally important, peptide binding region (PBR); the PBR of class I cell surface receptors is encoded by two separate exons (exon 2 and exon 3), while a single exon, exon 2, encodes the PBR of class II receptors.

The MHC consists of multiple loci (polygenic) and multiple alleles (polymorphic) and is considered the most diverse gene family in vertebrates (Edwards & Hedrick 1998). The evolutionary pattern observed in the MHC is explained by the birth-and-death evolution model, which proposes that repeated gene duplication creates new genes, some of which remain in the genome, whereas others are deleted or inactivated through deleterious mutations (Nei & Rooney 2005). MHC alleles are bi-parentally inherited and co-dominantly expressed, generating high variation in number of alleles and level of heterozygosity within individuals (Simpson 1988). High diversity of MHC allele sequences within most natural populations is interpreted as an adaptation to resist infection by a wide range of pathogens, driven by pathogen-mediated (positive) selection (Simpson 1988; Yuhki & O'Brien 1990). Evidence supporting the influence of MHC loci on vulnerability to infectious diseases through pathogen-mediated selection has been reported for many species, such as the great reed warbler (*Acrocephalus arundinaceus*; Westerdahl et al. 2005), Atlantic salmon (*Salmo salar*; Lohm et al. 2002; Kjøglum et al. 2006), three-spined sticklebacks (*Gasterosteus*

aculeatus; Wegner et al. 2003a), pythons (*Liasis fuscus*; Madsen & Ujvari 2006), Soay sheep (*Ovis aries*; Paterson et al. 1998), mouse lemurs (*Microcebus murinus*; Schad et al. 2005), and raccoons (*Procyon lotor*; Ruiz-López et al. 2014).

Two nonexclusive mechanisms have been proposed as drivers of this pathogen-mediated selection: heterozygote advantage and rare-allele advantage (Hedrick 2002; De Boer et al. 2004). Heterozygote advantage (also described as the overdominance model) proposes that heterozygous individuals have the ability to recognize a wider variation of pathogen peptides than homozygous individuals (Hughes & Nei 1988; Yuhki & O'Brien 1990). Recognition of many pathogens promotes increased resistance to diseases and potentially higher relative fitness, thus creating an advantage for heterozygous individuals. A selective advantage for heterozygosity with new mutations suggests nonsynonymous codon substitution rates will exceed synonymous ones. This nucleotide substitution pattern was first observed within regions of MHC class I in humans and mice (Hughes & Nei 1988). In addition, many experimental studies report a significant influence of MHC heterozygosity on individual pathogen prevalence in wild populations (Wegner et al. 2003a; Westerdahl et al. 2005; Madsen & Ujvari 2006). However, mathematical models and stochastic simulations demonstrate that heterozygote advantage alone does not sufficiently explain the high levels of MHC polymorphism observed in vertebrates, providing support for the influence of additional selection pressures on the evolution of MHC diversity (De Boer et al. 2004).

Another suggested mechanism of pathogen-mediated selection at the MHC occurs through rare-allele advantage via the mechanism of a co-evolutionary arms race (termed the Red Queen hypothesis; Van Valen 1977; Dawkins & Krebs 1979) between pathogens and their host's defences (Ladle 1992). In the presence of rapidly evolving host immune systems, pathogens themselves experience strong selective pressure to continuously evolve in order to infect common host genotypes; this establishes an advantage for unique host genotype mutations. Through this mechanism (also termed negative frequency-dependent selection; Takahata & Nei 1990), specific alleles may be selected *against* when they occur frequently, and selected *for* when they are rare (Ladle 1992; Borghans et al. 2004). Because host-pathogen interactions are significantly influenced by local environmental conditions the advantage of specific alleles is thought to vary across both spatial and temporal scales (Hedrick 2002; Zhang & He 2013). The influence of one or more specific alleles on pathogen resistance or susceptibility has been established for a number of MHC loci, in species such as

Soay sheep (Paterson et al. 1998), pythons (Madsen & Ujvari 2006), mouse lemurs (Schad et al. 2005) and house sparrows (*Passer domesticus*; Bonneaud et al. 2006; Loiseau et al. 2008).

Reproductive mechanisms have also been proposed as complementary processes for upholding diversity at the MHC (Potts & Wakeland 1993; Penn & Potts 1999). For example, negative assortative mating, cued by odour, is thought to foster increased mating between individuals with different genotypes, which then encourages heterozygosity in offspring (Penn & Potts 1999). It is likely that multiple selection pressures act on MHC simultaneously (De Boer et al. 2004) and an intermediate level of MHC diversity may be optimal for individuals (Nowak et al. 1992; Wegner et al. 2003b). Due to its maintenance of high diversity and role in the adaptive immune system, MHC is highly likely to be critical to species' adaptive responses to fluctuating and unpredictable pathogen environments (Altizer et al. 2003). Consequently, it has become the model system for studying the adaptive potential of species with respect to disease (Sommer 2005; Spurgin & Richardson 2010; Blanchong et al. 2016).

Adaptation to shifting disease patterns

Pathogen communities are one of the most important drivers of change within natural populations, influencing both population size and genetic composition over time and space (Altizer et al. 2003). Pathogens act as selective forces on host species, promoting the evolution of a wide range of cellular defence mechanisms and ultimately immunological resistance when exposed to a variety of pathogens (Altizer et al. 2003; Altizer et al. 2006). Immunogenetic diversity in host populations therefore contributes to an essential buffering mechanism against widespread disease outbreaks (Altizer et al. 2003). Changes in environmental conditions, e.g. habitat fragmentation and climate change, caused by anthropogenic activities, e.g. urbanization and globalization, can directly impact pathogen communities and patterns of disease outbreak (Daszak et al. 2000; Altizer et al. 2006). A number of recent studies suggest that new routes of exposure, driven by anthropogenic encroachment into wildlife areas and range-shifts induced by climate change (Harvell et al. 2002; Hoffmann & Sgrò 2011), are subjecting host species to new pathogen environments (Altizer et al. 2003; Daszak et al. 2000). In fact, climate change is predicted to elicit large-scale changes in behaviour, movement and phenology of both hosts and pathogens (Altizer et al. 2013). Yet such shifts in pathogen environments, as well as their drivers, remain relatively understudied in wild populations (Spielman et al. 2004). This is of concern given that many

wild populations are threatened by multiple pressures simultaneously (e.g. restricted ranges, habitat fragmentation and overexploitation), and thus particularly vulnerable to infectious diseases (Murray et al. 1999). In such cases, adaptive genetic diversity and potential for evolutionary response will become paramount for the persistence of wild animal populations.

Mesocarnivores as indicators of trophic stability and related species trends

Due to their natural occurrence at low densities and large range requirements, carnivores are especially susceptible to the negative genetic effects, and subsequent disease vulnerability, of anthropogenic activities (O'Brien 1985; Inskip & Zimmermann 2009). Barriers to dispersal, such as roads and urban development, limit gene flow in species that naturally maintain large territories (Riley et al. 2006; Serieys et al. 2014). Furthermore, declining carnivore populations are extremely vulnerable to the negative effects of inbreeding promoted by home range contractions that cause naturally outbreeding populations to become isolated and inbred (O'Brien et al. 1985; Roelke et al. 1993). In addition to impacting the genetic makeup of carnivore populations, these pressures can directly influence the patterns of exposure to pathogens and frequency of human-carnivore conflicts (Murray et al. 1999; Inskip & Zimmermann 2009).

As human populations increasingly encroach on wild areas, the large range and dietary requirements of apex carnivores have led to increased conflict with humans in recent decades (Inskip & Zimmermann 2009). Persecutions by humans, in addition to habitat loss, has driven the decline and often complete removal of apex predators from many environments, especially in non-protected areas (Inskip & Zimmermann 2009; Prugh et al. 2009). One ecological outcome of the removal of apex predators is the increase of mesopredators, termed "mesopredator release", where populations of medium-sized predators flourish in the absence of historical apex predators (Soulé et al. 1988; Prugh et al. 2009). Mesopredators are predators from middle trophic levels that typically prey on smaller-sized animals. Because many mesopredator species exploit a broad prey resource base, they have reached high densities in and around many urban areas (Inskip & Zimmermann 2009; Prugh et al. 2009). Growing mesopredator populations increasingly affect the abundance of species at lower trophic levels. As a result, mesopredators are gaining greater importance in the ecosystems in which they occur and may be critical to maintaining trophic stability in ecosystems where large apex predators no longer occur (Prugh et al. 2009). However, as mesopredator populations grow, they face challenges related to fragmented and isolated habitat patches,

and exposure to the novel pressures of urban environments, including novel pathogens, disease outbreaks, poisons and conflict with humans (Prugh et al. 2009).

Infectious diseases are particularly threatening to urban mesocarnivore populations of conservation concern that have been left vulnerable by other factors such as range restriction and over-exploitation (Murray et al. 1999). In such cases, otherwise minor diseases can potentially devastate already declining populations. Additionally, due to their large home ranges, many carnivores traverse urban-rural boundaries on a regular basis, increasing their exposure to pathogen threats from both environments (Gehrt et al. 2010). Because many urban environments are associated with particularly high tick loads, tick-borne blood pathogens are a significant threat to urban and peri-urban mesocarnivores (Gehrt et al. 2010; Bevins et al. 2012; Rizzoli et al. 2014; Mackenstedt et al. 2015). A single tick can carry multiple disease-causing agents simultaneously and these pathogens can be transmitted by a wide variety of host types (Gilbert et al. 2001). Indeed, many pathogens found in wild populations have been transmitted from domestic animal hosts (Murray et al. 1999). Thus, mesocarnivores with large home ranges and/or large urban area cover within home ranges may have a higher risk of exposure to ticks and tick-borne pathogens (Riley et al. 2004; Gehrt et al. 2010; Rizzoli et al. 2014).

In the face of contemporary ecological pressures, including infectious diseases, the detection of genetic and population trends for mesocarnivores can provide valuable indicators of related trends in many other species that are similarly threatened. One such mesocarnivore is the caracal (*Caracal caracal*), a medium-sized felid distributed across Africa, the Middle East and southwestern Asia (Avenant & Nel 1998; Macdonald & Loveridge 2010). Due to its wide geographical distribution, caracal within different environments face a range of threats with varying intensities, causing local population trends to differ. Caracal are mainly solitary and typically occupy large territories; males maintain territories that can be up to three times larger than female territories (Avenant & Nel 1998). Like other carnivores, their wide home ranges leave caracal vulnerable to the effects of anthropogenic activities that limit, modify and isolate their natural movement; these effects include loss of genetic diversity, high levels of mortality via road deaths and persecution, together with exposure to rodenticide poisons and emerging diseases (Serieys et al. 2014). Indeed, a recent review of worldwide patterns in human-felid conflict identified human conflicts as the most significant threat to regional and local caracal conservation throughout its range (Inskip & Zimmermann 2009). In the face of

such a variety of threats, knowledge of genetic trends and adaptive potential could be valuable to inform conservation and management strategies for caracal and other emerging mesopredators (Lacy 1997).

MHC variation and pathogens in urban caracal

Population trends in patterns of adaptive genetic variation have been studied in many carnivore and mesopredator species. Numerous studies have quantified MHC variation within natural populations of wild felids, including characterization of MHC class II loci in species such as cheetah (*Acinonyx jubatus*; Castro-Prieto et al. 2011a), Bengal tiger (*Panthera tigris*; Pokorny et al. 2010) and Eurasian lynx (*Lynx lynx*; Wang et al. 2009). However, few studies of felid populations have characterised variation in MHC class I loci, and studies that have explored the relationship between MHC variation and pathogen infection are often limited by small sample size (Castro-Prieto et al. 2011a; 2011b). For example, a study by Addie et al. (2004) failed to detect a significant association between MHC class II loci and feline infectious peritonitis in domestic cats (*Felis catus*), but referred to small sample size as a possible explanation for this. Other studies that have suggested an association between felid MHC variation and disease susceptibility have only used indirect measurements (e.g. allograft rejection experiments in captive cheetah; O'Brien et al. 1985). Research that directly explores the influence of MHC gene diversity on pathogen prevalence will clearly contribute to a more nuanced understanding of disease vulnerability in the modified environments urban carnivores inhabit.

This study examines the relationship between MHC class I genetic variability and intracellular tick-borne pathogen prevalence within an urban population of caracal in the city of Cape Town, South Africa. Pathogen prevalence varies substantially between caracal populations across South Africa, likely influenced by the varying land uses and environmental pressures surrounding distinct populations (S. Viljoen, unpublished data from ongoing MSc study). Caracal from Cape Town navigate a heavily fragmented and transformed urban environment, in which there are barriers to gene flow and numerous urban-associated, disease-transmitting species, like squirrels, pigeons and domestic pets. The prevalence of tick-borne pathogens among caracal in Cape Town is significantly higher than that of caracals in both natural protected areas and agricultural landscapes (S. Viljoen, unpublished data). These tick-borne blood pathogens include *Hepatozoon felis*, *Babesia felis*, *Babesia leo*, *Babesia venatorum* and *Anaplasma* species, which cause a range of diseases,

including hepatozoonosis, babesiosis and anaplasmosis (Averbeck et al. 1990; Penzhorn et al. 2001; Penzhorn et al. 2004). These diseases significantly affect individual fitness by means of immune suppression, fever, appetite-loss, fatigue and anaemia (Schoeman et al. 2001; Bosman et al. 2007; Kumar et al. 2008).

Genetic variation at exon 2 and exon 3 of MHC class I loci together with the prevalence of a number of important blood pathogens was examined in 26 caracal individuals occurring in and around the metropole of the city of Cape Town. DNA and amino acid sequence variation at exon 2 and exon 3 was characterised and a Generalized Linear Model framework was used to explore the influence of (i) allelic richness, (ii) sequence diversity and (iii) specific alleles on individual pathogen prevalence. The impacts of urban land use and home range size (calculated from GPS collar data), as well as sex and age class, were also investigated as contributing factors to the role of MHC gene diversity on patterns of pathogen infection. Understanding all potential drivers of pathogen infection can provide valuable insight into the adaptive potential of populations in rapidly changing landscapes. This study provides the first assessment of caracal adaptive immunogenetic variation and its influence on pathogen prevalence.

Methods

Sample collection

This study contributes to an ongoing program of research by the Urban Caracal Project (urbancaracal.org) at the University of Cape Town (UCT), and makes use of caracal blood samples collected during field sampling and collaring of individuals in Cape Town, South Africa. Whole blood samples (10-30 ml) were drawn from the saphenous vein of anaesthetised, live-trapped animals that were ear-tagged and radio-collared. Samples were collected from 26 individuals (Table 1), from December 2014 to October 2016, and stored at -80°C before processing. All sample collection from captured animals was conducted under the supervision of a qualified veterinarian, with ethical approval and appropriate permits from the University of Cape Town Science Faculty Animal Ethics Committee (2014/V20/LS), SANParks (2014-2017/CRC/2014-017) and Cape Nature (AAA007-00147-0056).

Individual radio-collars were programmed to collect GPS locations during a period ranging from three weeks to six months. Locations were recorded with two different fix schedules; beginning at three-hour intervals throughout the 24-hour cycle, followed by 20-minute intervals every 9-10th day, for 24-36 hours. Home ranges were calculated based on three-hour fixes by L. Serieys, for 21 study individuals (Table 1), using LoCoH (Getz et al. 2007). A percentage estimate of urban land use within each home range polygon was calculated by referencing the most recent available urban land use boundary data (City of Cape Town, 2012). Home range calculations were not obtainable for the remaining five study individuals due to dispersal or relocation out of the study area, or mortality less than one month after capture.

Table 1. Summary of the study caracal individuals from Cape Town, South Africa. Sex, age class (Ad = adult, Sub = sub-adult, Kit = kitten), home range size (km^2), urban area use within home range (km^2) and pathogen presence/absence data is reported for each individual. Pathogens were identified to genus. A dash indicates lack of data.

Individual	Sex	Age	Home range	Urban area	*Hepatozoon*	*Babesia*	*Anaplasma*	*Mycoplasma*
C1	M	Ad	76	3.87	0	1	1	0
C2	F	Ad	3.5	-	0	1	1	0
C3	F	Ad	10	0.44	0	1	1	0
C4	M	Ad	105	7.30	0	1	1	0
C5	-	Kit	-	-	1	0	1	0
C6	M	Ad	76	12.10	0	1	1	0
C7	M	Sub	22	10.91	0	1	1	0
C8	M	Sub	6	1.30	0	1	1	0
C9	F	Ad	19	2.81	0	1	1	0
C10	M	Ad	55	6.95	0	1	1	0
C11	M	Sub	22	2.63	0	1	1	0
C12	M	Ad	105	27.02	0	1	1	0
C13	F	Ad	20	3.48	0	1	1	0
C16	M	Ad	82	12.88	0	1	1	0
C17	M	Ad	50	9.08	0	1	1	0
C18	M	Sub	202	2.22	0	0	1	0
C19	F	Ad	10	9.79	0	1	1	0
C20	M	Ad	14	10.55	0	1	0	1
C21	M	Sub	35	8.12	0	1	1	0
C22	F	Sub	7	0.29	0	1	1	0
C24	F	Ad	30	0.02	0	1	1	0
C25	M	Ad	102	0	0	1	1	0
C26	M	Ad	-	-	0	0	1	0
C27	F	Ad	-	-	0	1	1	0
C28	F	Ad	-	-	0	1	1	0
C29	M	Ad	-	-	0	1	0	0

DNA extraction and PCR detection of pathogens

PCR optimisation and initial pathogen screening for 13 of the 26 study individuals was carried out by S. Viljoen (UCT) in 2015-2016. The remaining 13 individuals were screened as part of this study. Genomic DNA was extracted from whole blood samples using a DNeasy Blood and Tissue Kit (Qiagen, Germany) and stored at 4°C during processing. Using primers that target gene regions from a selection of tick-borne pathogens, multiple PCRs were performed with genomic DNA from each individual, displayed in Table 2. All PCRs were carried out on an Applied Biosystems 2720 Thermal Cycler (Applied Biosystems, USA). Reactions using RLB (Gubbels et al. 1999) and BTF (Jefferies et al. 2007) primer sets

consisted of 0.5 µl of forward and reverse 10 µM primers, 4 µl extracted DNA, 12.5 µl Dream Taq Green PCR Master Mix (ThermoScientific ™, South Africa) and 7.5 µl ultrapure water (ddH_2O). Reactions with the Ehr primer set (Bekker et al. 2002; Matjila et al. 2008) contained 0.5 µl of each 10 µM primer, 2 µl extracted DNA, 12.5 µl Dream Taq Green PCR Master Mix and 9.5 µl ddH_2O. All amplified products were visualized using 1% agarose gels, under UV light. Samples that amplified a band of the expected size (based on the primers' target) were considered positive. Positive samples were sequenced at the Central Analytical Facility, University of Stellenbosch, South Africa.

Table 2. Primers used for detection of tick-borne pathogens, through PCR screening. PCR protocols used for each primer set are detailed.

Target	Primer name	5'-3' sequences	PCR protocols		
			No. of cycles	Temp (°C)	Time
Babesia/Theileria 18S rRNA (Gubbels et al. 1999)	RLB-F	GAC ACA GGG AGG TAG TGA CAA G	1	94 94	10 min 20 s
	RLB-R	CTA AGA ATT TCA CCT CTG ACA GT	2	67 72	30 s 30 s
			2	Repeat at 65, 63, 61, 59, 57	20, 30, 30 s
			1	72	7 min
Babesia/Theileria/ Hepatozoon 18S rRNA (Jefferies et al. 2007)	BTF1	GGC TCA TTA CAA CAG TTA TAG	1	94 58 72	3 min 1 min 2 min
	BTR2	GGA CTA CGA CGG TAT CTG ATC G	35	94 58 72	30 s 20 s 30 s
			1	72	7 min
Ehrlichia/Anaplasma 16S rRNA (Bekker et al. 2002; Matjila et al. 2008)	Ehr-F	GGA ATT CAG AGT TGG ATC MTG GYT CAG	1	95	2 min
	Ehr-R	CGG GAT CCC GAG TTT GCC GGG ACT TYT TCT	35	95 61 72	30 s 30 s 30 s
			1	72	10 min

Samples indicating the presence of multiple infections from related species (with multiple peaks at identical positions in the sequence chromatogram) were cloned using the pGEM®-T Easy Vector System I (Promega, Madison, WI, USA) and JM109 *Escherichia coli* cells (Promega, Madison, WI, USA). Blue-white screening was used to identify positive clones on LIAX (Luria broth, 100 mM IPTG, 100mg/mL Ampicillin, 3% X-gal in dimethylformamide) agar plates. Five to ten positive colonies, and one negative colony, were selected for PCR

amplification with M13 primers designed to amplify the target insertion section of the vector. Reactions with the M13 forward (5'-GTA AAA CGA CGG CCA GT- 3') and reverse (5'-CAG GAA ACA GCT ATG AC-3') primers followed cycling conditions of: 95°C for 2 min, 35 cycles of 95°C for 30 s, 50°C for 20 s, 72°C for 40 s, and 72°C for 10 min. Each colony reaction contained 1 µl of 0.1 pmol/µl forward and reverse primers, 10 µl of Dream Taq Green PCR Master Mix (ThermoScientific ™, South Africa) and ddH_20, to a total volume of 20 µl. Again, successful amplifications were sent offsite to be sequenced.

Screening of MHC class I variation

Genomic DNA extracted from whole blood samples was also used to amplify exon 2 and exon 3 of MHC class I loci, using felid-specific primers from Smith & Hoffman (2001). Primers were designed to the adjacent intron regions (Figure 1) and PCR products were approximately 800 base pairs: exon 2 forward primer (5'-GGGAGCCCCGCTTCATCG-3') and exon 3 reverse primer (5'-CGTCTCCTTCCCCATGTCC-3') (Smith & Hoffman 2001). PCR conditions were as follows: initial denaturing at 94°C for 3 min, followed by 40 cycles of 1 min denaturation at 94°C, 2 min at 64°C, 2 min extension at 72°C, and ending with a final extension for 7 min at 72°C. PCR reactions contained 1 µl of diluted (1/10) genomic DNA, 0.5 µl of forward and reverse primers (at 50 pmol/µl), 12 µl of Dream Taq PCR Master Mix (ThermoScientific ™, South Africa) and ddH_20, to a total volume of 25 µl. Following visualization by gel electrophoresis on 1% agarose gels, positive amplifications were excised and cleaned up with the Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA).

Because gene duplication and extensive allelic diversity characterize MHC class I and class II loci, individual variation was determined using cloning and sequencing. PCR products were cloned using the pGEM®-T Easy Vector System I (Promega, Madison, WI, USA) with JM109 *E. coli* cells (Promega, Madison, WI, USA); and blue-white screening was used to identify positive clones on LIAX agar plates. Twelve to fifteen positive colonies were selected for colony PCR amplification and sequencing with plasmid specific primers (M13 forward: 5'-GTA AAA CGA CGG CCA GT- 3' and M13 reverse: 5'-CAG GAA ACA GCT ATG AC-3'), following the protocols previously described.

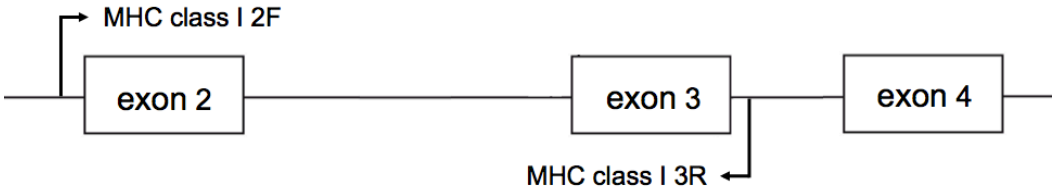


Figure 1. Schematic representation of a section of an MHC class I gene. Arrows indicate the location of the primers used in this study. The exon 2 and exon 3 coding sequences include the antigen-binding region of MHC class I molecules.

Assessing MHC class I diversity and testing selection

Sequence chromatograms were checked and manually edited in BioEdit v.7.2.5 (Hall 1999). Sequences were trimmed to include the coding regions of MHC class I exon 2 and exon 3, which were identified by alignment with published caracal sequences (Genbank accessions: AF283072, AF283078, AF283089, AF283095; Smith & Hoffman 2001). Sequences were aligned using the Clustal W alignment algorithm (Chenna et al. 2003) implemented in BioEdit, using default options. The number of unique sequences for each individual and the population was calculated using DNAsp v5.10 (Librado & Rozas 2009). To avoid overestimation of allelic diversity due to rare sequencing errors only alleles found in more than one individual and with at least two nucleotide differences were counted as unique, presumably functional alleles. Alleles were considered pseudo genes if the same premature stop codon(s) or disrupted reading frame(s) occurred within a sequence of more than one individual. These sequences were excluded from further analyses. Additional measures of diversity (gene diversity, nucleotide diversity and average number of nucleotide differences) were calculated for each individual and the population's allele repertoire using DNAsp v5.10 (Librado & Rozas 2009).

The phylogenetic placement of the caracal MHC class I sequences isolated in this study was determined using the neighbour-joining method as implemented in MEGA v6.0 using the Tamura-Nei model (Kumar et al. 2016). Branch support was evaluated by 1,000 bootstrap replicates. Additional exon 2 and exon 3 sequences available for a number carnivore species were sourced from Genbank (Appendix A and B) and aligned using the Clustal W algorithm (Chenna et al. 2003) in BioEdit using default options.

Amino acid sequences were deduced and functionally significant amino acids within the peptide binding region were identified through alignment with published domestic cat MHC

class I exon 2 and exon 3 sequences (Accessions: AF283068, AF283075, AF283085 AF283092; Smith & Hoffman 2001) and caracal MHC class I exon 2 and exon 3 sequences (Accessions: AF283072, AF283078, AF283095, AF283089; Smith & Hoffman 2001). To test for positive selection in the final data set for each exon, the ratio of non-synonymous to synonymous substitutions in the amino acid alignments was calculated. A higher ratio of non-synonymous (dN) substitutions relative to synonymous (dS) ones provides evidence for positive selection, i.e. $\omega = dN/dS > 1$. These ratios were compared for the entire MHC class I exon 2 and exon 3 sequences, the PBRs only, and the non-PBRs only. Calculations were made with a modified (Jukes-Cantor) Nei-Gojobori method (Nei & Gojobori 1986) in MEGA v6.0 (Kumar et al. 2016), using 1,000 bootstrap replicates.

Statistical Analysis

The relationship between MHC class I diversity and pathogen prevalence was analysed using a Generalized Linear Model (GLM) framework in R v3.3.2 (R Development Core Team 2016). Due to a lack of sufficient variation in individual pathogen prevalence across the population, as well as limited sample size, analyses were conducted using a univariate approach. Binomial GLMs were used to explore the effect of (i) the number of MHC class I alleles and (ii) diversity within alleles on the presence/absence of *Babesia* in individual caracal. Because the population's allele frequency distribution for both exons showed one or two very common alleles and numerous rare alleles, the analysis for the role of specific alleles was limited to three alleles in order to permit the detection of statistically significant results. The GLMs implemented explored the association between the presence/absence of these alleles and *Babesia* infection. Additional GLMs were used to analyse the effect of sex, age, home range size and urban area use within home ranges on individual *Babesia* prevalence.

Results

Characterization of MHC class I exon 2 and exon 3 diversity

From the 312 clone sequences examined, 41 unique alleles were identified in the MHC class I exon 2 (Figure S1). Three non-functional alleles, considered pseudogenes, were excluded from further analysis due to the presence of stop codons in the amino acid translations (Figure 2). Thus, 38 unique, putatively functional exon 2 alleles were identified in the population sample of 26 individuals. Each individual had between two and six functional alleles (mean = 3.2 ± 1.3 SD), suggesting a minimum of three functional loci present. Measures of genetic diversity within the allele repertoire of exon 2, as well as exon 3, are displayed in Table 3.

The assessment of exon 3 sequences revealed 65 unique alleles in the population (Figure S2), eight of which included stop codons in the amino acid translations (Figure 3). An additional 12 sequences had a nucleotide deletion that resulted in a frame shift mutation, and likely represents a pseudogene. A BLAST analysis identified high affinity to a pseudogene sequence isolated from a tissue expression study in domestic cat (Accession: KC763055; Holmes et al. 2013). Further analysis of cDNA is, however, required to confirm this, as a number of non-classical MHC loci are also known in mammals (Pratheek et al. 2014). These sequences were also removed from further analysis, resulting in a total of 45 unique, putatively functional alleles at exon 3. Between two and seven putatively functional alleles (mean = 4.2 ± 1.8 SD) were detected in each individual from the sample population (Table 3), suggesting at least four putatively functional loci based on the MHC exon 3 sequences in caracal.

Table 3. Summary of MHC class I exon 2 and exon 3 diversity measures* for the caracal sample population (n = 26) from Cape Town, South Africa.

	D	A	Pseudo	No. of putatively functional alleles per individual	Gene diversity	Pi	Avg. no. of nucleotide differences	Avg. no. of amino acid differences	Amino acid divergence (%)
Exon 2	41	38	Yes	2 - 6	0.99	0.06	11.4	6.2	10.2
Exon 3	65	45	Yes	2 - 7	1.0	0.07	14.5	8.7	11.9

*D = no. of detected unique alleles; A = no. of unique, putatively functional alleles; Pseudo = presence of pseudogenes; Pi = nucleotide diversity; and amino acid divergence = average number of amino acid codon differences/total number of amino acid codons.

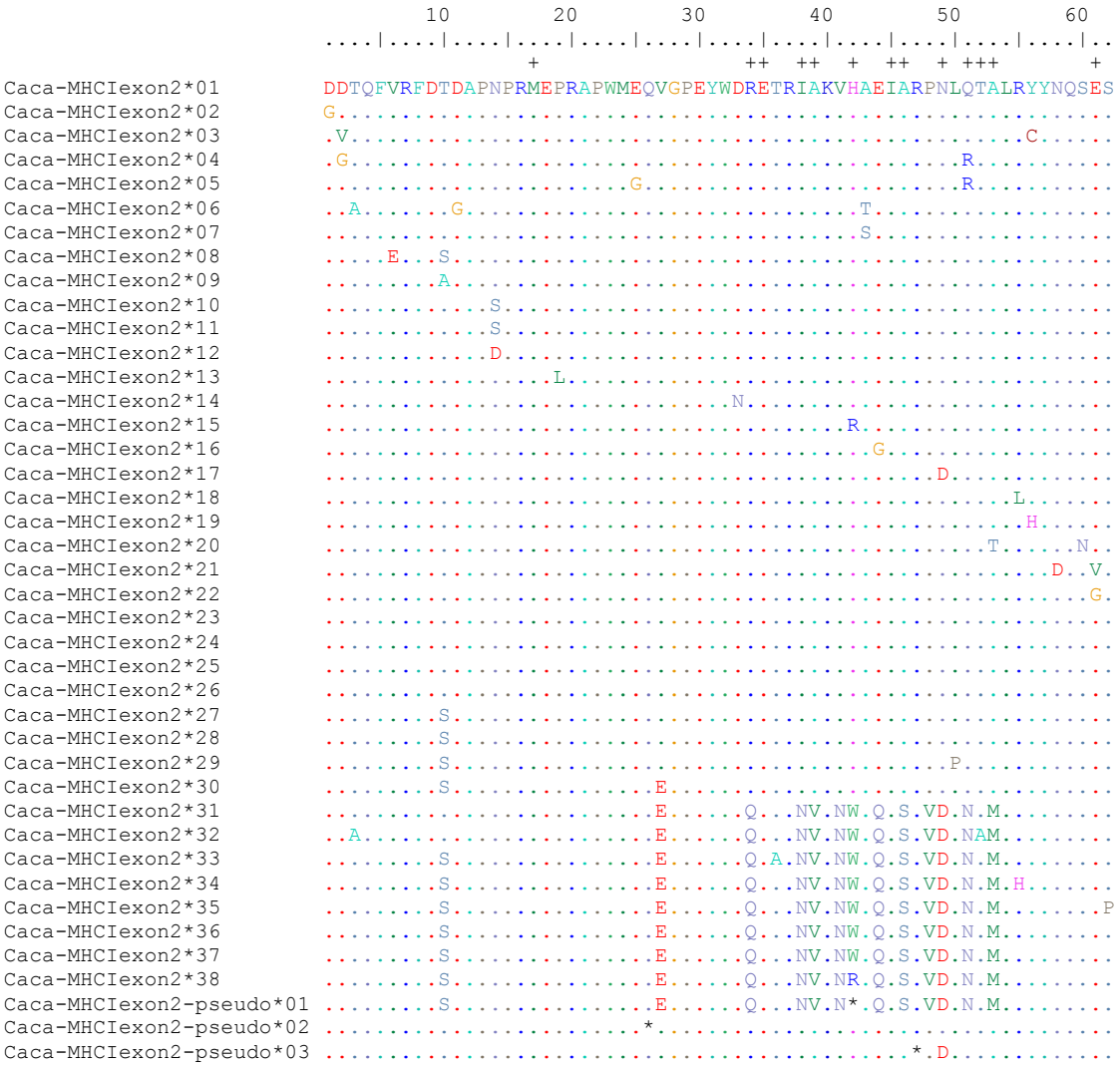


Figure 2. Amino acid translations of the unique MHC class I exon 2 alleles detected in this study. Dots and single letters designate nucleotide bases identical to and different from the first allele respectively. Asterisks represent stop codons. Sites with a + indicate previously reported functionally important amino acids within the peptide binding region of MHC class I receptors (Yuhki & O'Brien 1994; Smith & Hoffman 2001).

Figure 3. Amino acid translations of the unique MHC class I exon 3 alleles detected in this study. Dots and single letters designate nucleotide bases identical to and different from the first allele respectively. Asterisks represent stop codons. Sites with a + indicate previously reported amino acids within the MHC class I peptide binding region (Yuhki & O'Brien 1994; Smith & Hoffman 2001).

Phylogenetic relationships

The final class I sequence alignment for each exon together with published sequences from other carnivores, including felids, are reported in Appendix A and B. For each exon a sub-set of representative sequences were used to construct phylogenetic trees, comprising 19 exon 2 alleles (Figure 4) and 21 exon 3 alleles (Figure 5). Two clades characterized the evolutionary relationships of both exons. Clade one, in both analyses, is comprised of highly similar sequences from caracal and a number of related felid species, including domestic cat, Geoffroy's cat (*Leopardus geoffroyi*), sand cat (*Felis margarita*), golden cat (*Catopuma temminckii*), margay (*Leopardus wiedii*) and Pallas's cat (*Otocolobus manul*). Clade two includes sequences from caracal and related species, but with substantially longer branch lengths, reflecting their higher overall levels of nucleotide variation.

Positive selection at exon 2 and exon 3 of caracal MHC class I loci

The positive selection test, modelled on the theory that positive selection enhances the rate of nonsynonymous codon substitutions compared with that of synonymous ones, revealed evidence for positive selection within the PBR codons of exon 3 (Table 4). In fact, the comparison of only PBR exon 3 codons displayed over twice as many nonsynonymous substitutions as synonymous ones (ω = dN/dS = 2.373), although the difference was not statistically significant (one sided z-test; Z = 1.179 p = 0.120). Evidence for positive selection was not detected in the full sequences or non-PBR codons in both exon 2 and exon 3. Additionally, the PBR codons in exon 2 did not show an excess of nonsynonymous changes (ω = dN/dS = 0.846).

Table 4. Mean genetic distances* of synonymous (dS) and nonsynonymous (dN) substitutions in putatively functional MHC class I exon 2 and exon 3 alleles in caracal.

	dS ± SD	dN ± SD	ω = dN/dS	p
Exon 2				
All codons	0.074 ± 0.068	0.065 ± 0.068	0.878	1.000
PBR only	0.227 ± 0.299	0.192 ± 0.247	0.846	1.000
Non-PBR only	0.053 ± 0.042	0.038 ± 0.032	0.717	1.000
Exon 3				
All codons	0.097 ± 0.095	0.065 ± 0.058	0.670	1.000
PBR only	0.075 ± 0.091	0.178 ± 0.201	2.373	0.120
Non-PBR only	0.100 ± 0.099	0.050 ± 0.040	0.500	1.000

*Calculated by the Nei & Gojobori (1986) method with Jukes-Cantor corrections. The p-value of the Z test is reported, with a null hypothesis dN/dS ≤ 1.

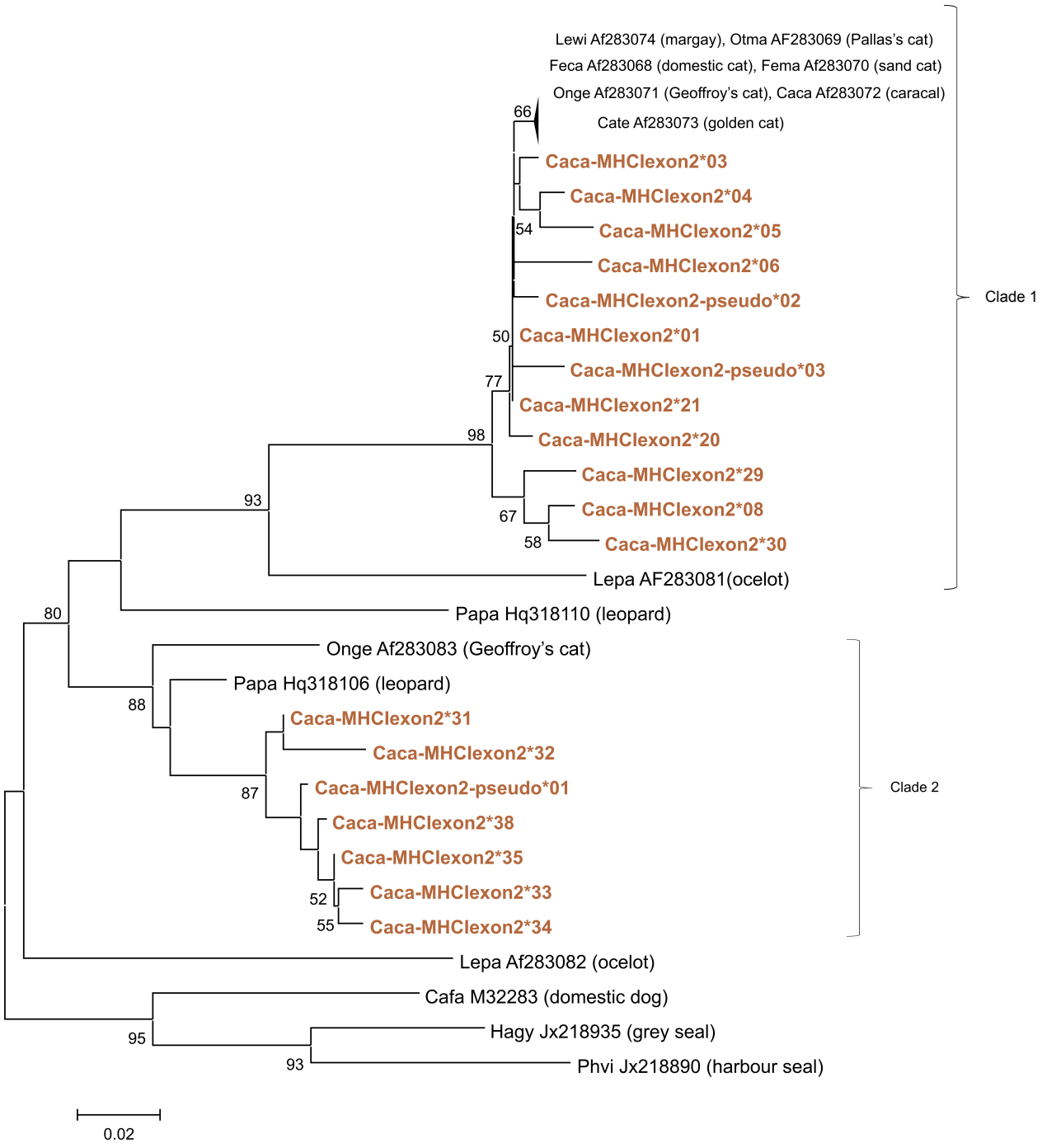


Figure 4. Phylogenetic placement of caracal MHC class I exon 2 alleles and related species using neighbour-joining tree construction and rooted with domestic dog and two seal species. Bootstrap support values >50% are indicated.

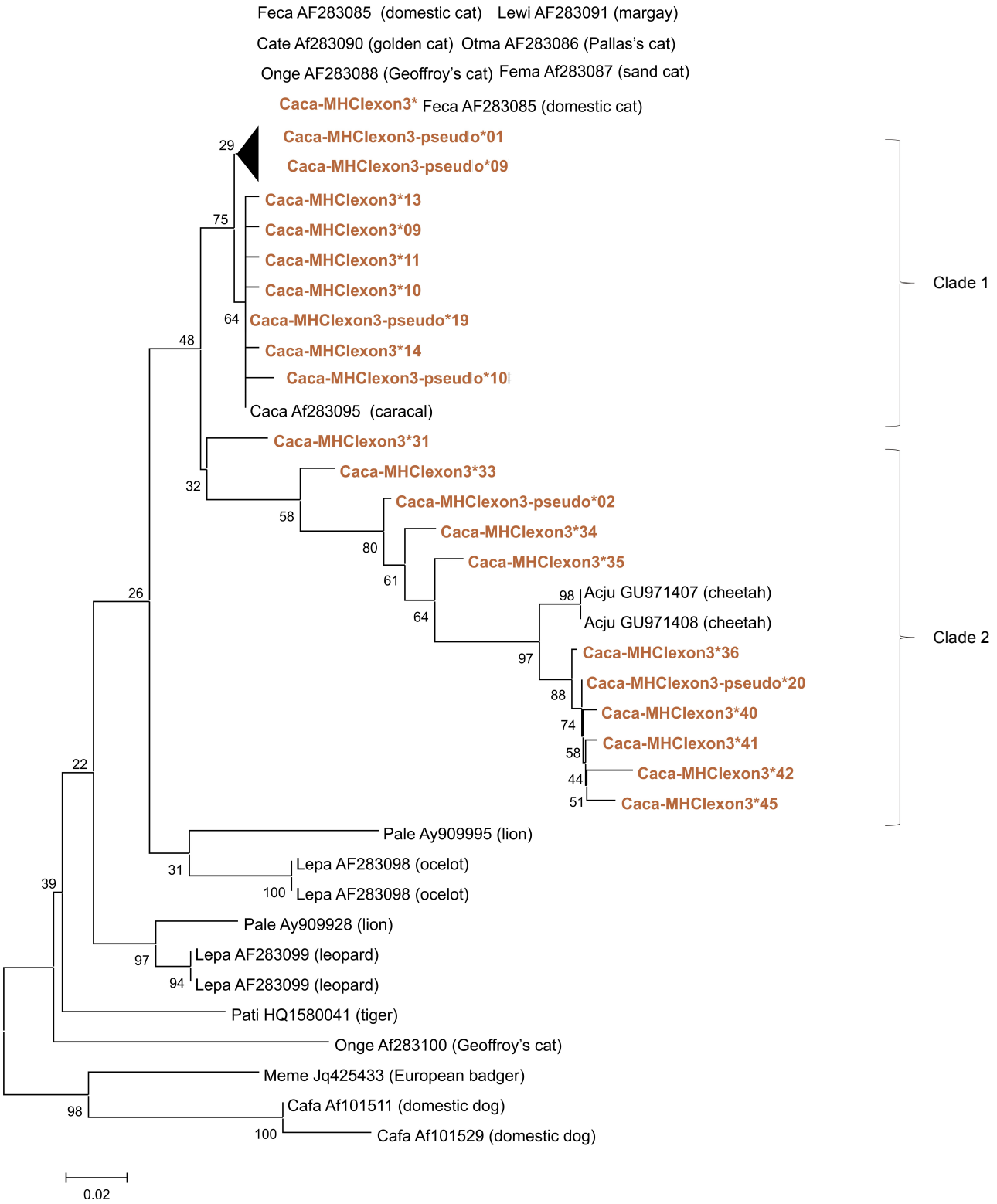


Figure 5. Phylogenetic placement of caracal MHC class I exon 3 alleles and related species using neighbour-joining tree construction and rooted with domestic dog and European badger. Bootstrap support values >50% are indicated.

Influence of exon 2 alleles and diversity on Babesia prevalence

Individuals free of *Babesia* infection (n = 3) had 2-3 MHC class I exon 2 alleles, while individuals that were infected (n = 23) had between two and six. This difference between the numbers of alleles of infected (mean = 3.3 ± 1.3 SD) and uninfected (mean = 2.3 ± 0.6 SD) individuals was not statistically significant (Table 5). Gene diversity within each individual's allele repertoire also did not have a significant effect on whether the individual was infected or uninfected with *Babesia* (χ^2 = 1.089, p = 0.297). However, both nucleotide diversity and average number of nucleotide differences in individuals' exon 2 alleles were found to have a significant effect on *Babesia* infection (χ^2 = 5.628, p = 0.018; χ^2 = 6.471, p = 0.011 respectively). On average, infected individuals had a higher nucleotide diversity (0.040 ± 0.025) and average number of nucleotide differences (7.5 ± 4.7 SD) than uninfected individuals.

Influence of exon 3 alleles and diversity on Babesia prevalence

All genetic diversity measures of MHC class I exon 3 had significant associations with *Babesia* prevalence (Table 5). On average, infected individuals possessed nearly double the number of exon 3 alleles (mean = 4.4 ± 1.8 SD) that uninfected individuals had (mean = 2.3 ± 1.2 SD). In addition, higher gene diversity, nucleotide diversity and average number of nucleotide differences were associated with *Babesia* infection. In fact, the latter two diversity measures showed highly significant differences between infected and uninfected individuals (p = 0.003).

Specific MHC class I alleles and pathogen prevalence

At both exon 2 and exon 3, one allele was found to be extremely common throughout the population, while the majority of alleles were very rare. The most common allele at exon 2, Caca-MHC1-exon2*01, was detected in every individual of the study population; and at exon 3, Caca-MHC1-exon3*01 occurred in 25 out of 26 individuals. Three alleles, Caca-MHC1-exon2*36, Caca-MHC1-exon3*12, and Caca-MHC1-exon3*37, occurred at intermediate frequencies in the population (0.54, 0.23 and 0.46 respectively) and were tested for their effect on *Babesia* prevalence. All other alleles isolated in this study occurred in less than 8% (one or two individuals) of the population. A significant correlation between two specific alleles and presence of *Babesia* was observed. Probability of a *Babesia* infection for individuals carrying Caca-MHC1-exon2*36 (χ^2 = 5.101, p = 0.024) or Caca-MHC1-

exon3*37 (χ^2 = 4.048, p = 0.044) was significantly higher than those that did not carry the alleles (Figure 6). These alleles were not exclusively inherited together, as only nine individuals carried both within the same sequence (frequency = 0.35). Interestingly, when analysed as full sequences i.e. exon 2 and exon 3 for each individual sequence, the probability of *Babesia* infection for individuals carrying a sequence with both Caca-MHC1-exon2*36 and Caca-MHC1-exon3*37 was not significantly different from individuals without both alleles (χ^2 = 2.753, p = 0.097). The difference between the probability of infection for individuals with the Caca-MHC1-exon3*12 allele and those without it was also not statistically significant (χ^2 = 1.688, p = 0.194).

Table 5. Genetic diversity variables for individuals that were infected (n = 23) and uninfected (n = 3) with *Babesia*. Measures of home range and urban area cover in home range for infected (n = 20) and uninfected individuals (n = 1) are also included. Mean values, standard deviations and p-values* are displayed for each variable.

	Babesia		Analysis	
Model variable	Infected (mean ± SD)	Uninfected (mean ± SD)	p	Significance
Exon 2				
No. of alleles	3.3 ± 1.3	2.3 ± 0.6	0.118	-
Gene diversity	0.652 ± 0.167	0.530 ± 0.342	0.297	-
Nucleotide diversity	0.040 ± 0.025	0.008 ± 0.003	0.018	*
Avg. no. of nucleotide differences	7.539 ± 4.728	1.111 ± 0.631	0.011	*
Exon 3				
No. of alleles	4.4 ± 1.8	2.3 ± 1.2	0.038	*
Gene diversity	0.746 ± 0.137	0.465 ± 0.303	0.018	*
Nucleotide diversity	0.034 ± 0.025	0.003 ± 0.002	0.003	**
Avg. no. of nucleotide differences	7.419 ± 5.535	0.611 ± 0.419	0.003	**
Home range	50.1 ± 49.4	202	0.274	-
Urban area cover in home range	6.6 ± 6.5	2.2	0.369	-

*p < 0.05; **p < 0.01; - non-significant

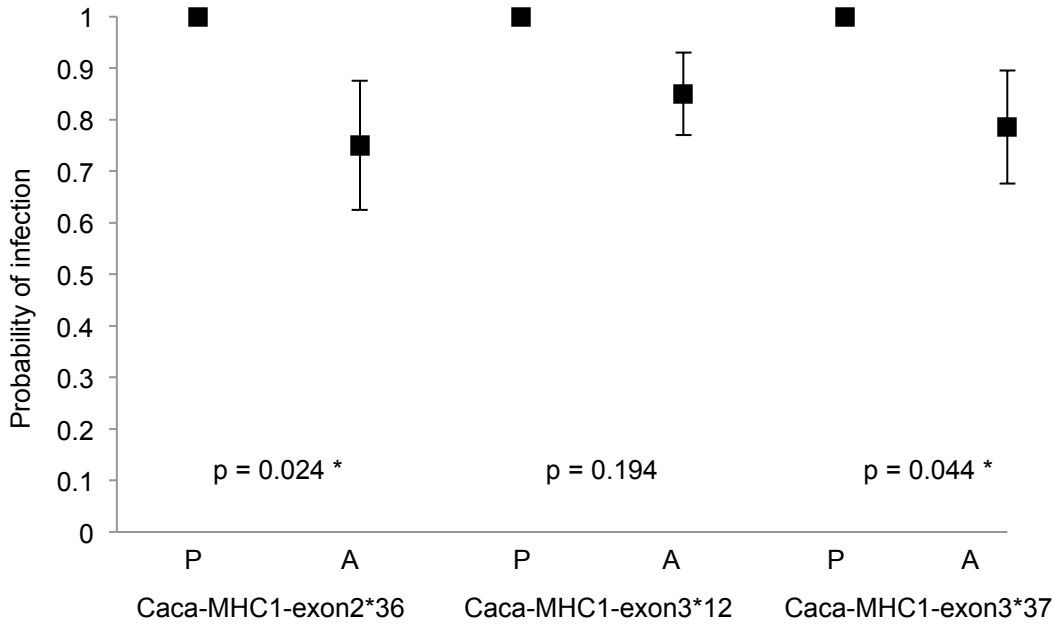


Figure 6. Probability of infection (mean ± SE) by *Babesia* in caracal that carry the following specific alleles: Caca-MHC1-exon2*36, Caca-MHC1-exon3*12, Caca-MHC1-exon3*37. P = presence of allele, A = absence of allele. p-values and significance codes (*p < 0.05) are shown.

Sex, age, home range, urban area cover, and pathogen prevalence

To investigate factors that influence infection rates beyond adaptive genetic variation, sex and age class were included in additional GLMs. The sex of individuals did not demonstrate a significant influence on infection with *Babesia*. While the probability of *Babesia* infection was higher for males (probability = 1.0) than for females (probability = 0.875), the difference was not significant (χ^2 = 1.882, p = 0.170). The age class analysis categorized individuals as adults or sub-adults, excluding one kitten. The probability of infection was 0.95 for adults and 0.83 for sub-adults; although again, the results were not statistically significant (χ^2 = 0.696, p = 0.404).

Home range size (mean = 50.1 km^2 ± 49.4 SD) and cover of urban areas within home ranges (mean = 6.59 km^2 ± 6.49 SD) was quite variable between the 21 individuals for which data was available. However, the distribution of detected pathogens was highly similar between these individuals, with only one individual free of *Babesia* infection, one free of *Anaplasma* infection, one with a *Mycoplasma* infection and zero with a *Hepatozoon* infection. This lack of variability of infections and limited sample size may have contributed to the lack of statistically significant findings. Neither home range (χ^2 = 1.197, p = 0.274) nor extent of urban area within home ranges (χ^2 = 0.740, p = 0.390) had a significant association with pathogen prevalence.

Discussion

MHC class I diversity and selection trends in the caracal

The importance of the MHC in population and individual susceptibility to pathogens makes it valuable to include in studies of adaptive potential (Sommer 2005; Spurgin & Richardson 2010; Blanchong et al. 2016). Although MHC diversity has been characterized in many felid species, diversity at class I loci remains understudied. This study provides the first comprehensive evaluation of MHC class I diversity in the caracal. Through examining diversity at the exons involved in binding and presenting foreign peptides to the immune system, and its association with individual *Babesia* prevalence, a number of interesting findings were revealed. Out of 312 clone sequences isolated from 26 individuals, 38 and 45 unique, putatively functional alleles were identified from exon 2 and exon 3 respectively. A number of alleles with differing nucleotide sequences contained only synonymous substitutions, which do not cause changes when translated into amino acids, resulting in the detection of 31 (exon 2) and 43 (exon 3) unique amino acid sequences. The number of alleles observed per individual suggests the presence of a minimum of four functional class I loci in the caracal. Because the number of clones sequenced per individual was limited to 12-15, a conservative approach was used for quantifying the number of unique alleles; these results may therefore slightly underestimate the true number of MHC class I alleles and loci present in the species. Nevertheless, these findings do correspond to estimates reported in previously published literature on felid MHC class I loci; Yuhki et al. (2008) reported three in the domestic cat (*felis catus*), Castro-Prieto et al. (2011a) identified three in the cheetah (*Acinonyx jubatus*), Sachdev et al. (2005) identified five in the Asiatic lion (*Panthera leo persica*), and Pokorny et al. (2010) reported four in the Bengal tiger (*Panthera tigris tigris*).

Phylogenetic analyses of caracal class I sequences revealed (i) close evolutionary affinity with published class I felid sequences and (ii) the presence of two clades, both of which comprise caracal sequences together with sequences from other felid species. These two clades are characterised by distinct amino acid profiles, which can be seen very clearly in Figure 4 and Figure 5. In both exon analyses, the first clade is comprised of highly similar sequences from the caracal and related species. This finding is in agreement with Smith and Hoffman's (2001) analysis of MHC class I sequences isolated from three different felid lineages, where the authors report an allele shared by all three lineages analysed in their study (pantherine, domestic cat and ocelot), and part of which was also isolated in this study (Caca-

MHC1-exon3*01). A more detailed analysis of felid MHC class I genes is clearly required to gain a more comprehensive understanding of evolutionary patterns of diversification and possible gene conversion events.

Despite exon 2 and exon 3 together encoding the functionally important peptide binding sites of MHC class I, few published studies on felids have examined both exons. Those that have studied both generally observed higher polymorphism in exon 2 than in exon 3 (e.g. domestic cat, Yuhki & O'Brien 1994; Asiatic lion, Sachdev et al. 2005; and Bengal tiger, Pokorny et al. 2010). However, the present study detected higher diversity at exon 3 in caracal, particularly within the peptide binding sites. In addition, evidence for positive selection was only identified in the functionally important amino acid residues of the exon 3 PBR. Although this result was not statistically significant, it does suggest some degree of positive selection likely acts on the functionally important codons in caracal MHC class I exon 3, and that variation at exon 3 may be more important for caracal MHC class I function than studies of related species have suggested. The ratio of non-synonymous to synonymous substitutions in exon 2 was highly similar when analysed for all codons and the PBR codons only, and both ratios were relatively high ($\omega > 0.8$). Nevertheless, no evidence of positive selection was detected in these regions. Interestingly, these results are very different from those of the Bengal tiger, within which highly significant evidence for selection in the exon 2 PBR was observed (Pokorny et al. 2010).

Association between high MHC class I diversity measures and Babesia prevalence

Pathogen screening of the population revealed that caracal of Cape Town have high levels of infection by a number of tick-borne blood pathogens. No individual tested was free of any infections. Many individuals (n = 23) tested positive for multiple infections, a pattern that has been observed in other studies of pathogen prevalence in South African felid populations (Penzhorn et al. 2006; Bosman et al. 2007). In addition, there was little variation in the pathogen prevalence across the population, limiting the extent to which a GLM framework could be used to explore the influence of MHC variation, and other interacting factors, on infection by certain pathogens. Despite this, multiple measures for MHC class I diversity were significantly associated with the prevalence of *Babesia*. Infection with *Babesia* was linked with higher exon 2 diversity (nucleotide diversity and average number of nucleotide differences) and exon 3 diversity (number of alleles, gene diversity, nucleotide diversity and average number of nucleotide differences) relative to that of uninfected individuals. Although

low diversity is often assumed to contribute to pathogen susceptibility, there are a number of possible explanations for an association between high diversity and pathogen infection. Firstly, several studies have argued that maximal numbers of MHC allelic diversity does not result in the highest individual fitness levels (Nowak et al. 1992; Wegner et al 2003). This is due to the interaction of MHC molecules and T-cell receptors, which drives an immunological compromise between the numbers of expressed MHC molecules and T-cell receptors. The expression of too many MHC molecules causes an overall loss of T-cell receptors through the thymic selection process, which can lead to the presentation of too many self-derived peptides, leading to inhibition of overall pathogen recognition efficiency (Lawlor et al. 1990; Nowak et al. 1992; Wegner et al 2003). The pressure of opposing selective forces (thymic and positive) acting on MHC likely favours an intermediate level of individual MHC diversity (Nowak et al. 1992). Indeed, experimental evidence has demonstrated that an intermediate level of MHC diversity can correspond to the highest levels of fitness in the face of multiple parasites (Wegner et al. 2003b). Thymic selection is likely to contribute to an optimum number of MHC class I alleles in natural populations, but very little is known about it in non-model organisms.

Association between specific MHC class I alleles and Babesia prevalence

Another point to consider is the effect of specific alleles or specific allele combinations on pathogen susceptibility, which may be more important than overall levels of MHC diversity. Numerous studies have found evidence for the influence of one or more specific MHC alleles on both pathogen resistance and pathogen susceptibility (Schad et al. 2005; Bonneaud et al. 2006; Loiseau et al. 2008). It is possible that expressing more total alleles increases the chance of having a specific allele that is associated with higher probability of infection. In this study, alleles Caca-MHC1-exon2*36 and Caca-MHC1-exon3*37 were both linked to a significant increase in the probability of *Babesia* infection. These alleles were found in approximately 54% (n = 14) and 46% (n = 12) of the population sample respectively. These sequences are not exclusively inherited together i.e. only nine individuals carried both alleles simultaneously. When full sequences were analysed for each individual, there was no significant effect of carrying both alleles. The majority of unique alleles detected in this study were rare, found in only one or two individuals, and were at frequencies that were too low for statistical analysis of their effects on immunocompetence. A wide variation in the distribution of alleles was also shown in the domestic cat MHC Class II (Addie et al. 2004). Additional

sampling of caracal could help determine whether any of these rare alleles do relate to pathogen resistance.

An alternative explanation for the observed positive relationship between high MHC diversity and specific alleles with pathogen infection is that high variation and/or specific alleles might contribute to individuals surviving infection. In such a case, caracal that have a large number of alleles, or specifically carry alleles Caca-MHC1-exon2*36 or Caca-MHC1-exon3*37, may be better at tolerating infections; whereas caracal that do not carry the alleles more often succumb to infections and their lethal effects. Unfortunately, an analysis of cellular indicators of immune response and stress was not part of this study. However, future research that includes cellular measures of health might reveal interesting associations between the MHC and infection tolerance. This type of analysis could be especially informative in Cape Town, where the caracal population appears to support a large number of individuals with multiple infections of tick-borne pathogens.

Multiple drivers of pathogen prevalence in wild populations

It is important to remember that a range of variables interacts to influence pathogen prevalence in wild populations; these include genetic diversity, environmental conditions and population history. The influence of sex and age class was therefore also investigated, but showed no significant effect on *Babesia* prevalence in the population. This suggests that sub-adult males do not have an increased risk of pathogen exposure despite frequently being pushed to smaller, suboptimal edges of habitat by larger, territorial adults (L. Serieys, unpublished data). Testing for infection by a range of additional pathogens, and in more individuals, would clearly contribute to a more thorough exploration of the role of age and gender on patterns of individual infection.

Exposure to urban and agricultural landscapes versus natural habitat influences the probability of exposure to ticks and thereby infection by tick-borne pathogens in host populations (Bevins et al. 2012; Rizzoli et al. 2014; Mackenstedt et al. 2015). The lack of individual variation in infections within Cape Town's caracal population, yet distinct infection trend compared with populations in natural and agricultural areas, suggests that the environment is likely to be an important factor in determining pathogen prevalence for caracal in South Africa. Unfortunately, there is scant information about the tick loads in the environments across these regions. In addition, seasonal changes in tick loads together with

differing tolerance levels by host species make it difficult to assess the overall disease pressure on populations (Mackenstedt et al. 2015). Despite this limitation, studies have found that use of urban areas can increase the risk of disease exposure for carnivore populations (Riley et al. 2004). With this in mind, home range size together with the proportion of urban cover in an individual's home range were used as ecological proxies to test for any relationship between pathogen prevalence and landscape. No significant effect of these factors was detected, although a larger sample size is required to better assess the role of landscape in patterns of infection.

While a larger sample size might overcome some of the limitations of the current dataset, the results also highlight the need for future studies to examine pathogen burden and MHC variation over time. Because wild populations are exposed to frequent environmental fluctuations, host-pathogen interactions, as well as subsequent selection pressure, significantly vary with time (Hedrick 2002). Although long-term studies of the MHC in wild populations are scarce, fluctuating selection has been observed within the Soay sheep (*Ovis aries*) MHC class II, during a 13-year study (Charbonnel & Pemberton 2005). Additional long-term studies of adaptive genetic variation will contribute to a better understanding of the patterns of selection pressure for wild populations in rapidly transforming landscapes that present the immune system with many novel stressors.

Conclusion

This study assessed MHC class I variation in an urban population of caracal, and provides the first evaluation of the relationship between class I diversity and specific alleles on *Babesia* prevalence in a wild felid population. The adaptive genetic variation at the MHC provides an ideal subject for better understanding population susceptibility to pathogens in wild populations. Examination of this relationship in transformed, and particularly urban landscapes, is especially important in the face of rapid environmental change. Future studies that employ a similar approach will contribute to a more nuanced understanding of the ways adaptive genetic variation contributes to disease susceptibility, and will become increasingly valuable to conservation in a changing world.

Literature Cited

Addie DD, Kennedy LJ, Ryvar R, Willoughby K, Gaskell RM, Ollier WER, ... Radford AD. 2004. Feline leucocyte antigen class II polymorphism and susceptibility to feline infectious peritonitis. Journal of Feline Medicine and Surgery 6:59-62.

Altizer S, Harvell D, Friedle E. 2003. Rapid evolutionary dynamics and disease threats to biodiversity. Trends in Ecology & Evolution 18:589-596.

Altizer S, Dobson A, Hosseini P, Hudson P, Pascual M, Rohani P. 2006. Seasonality and the dynamics of infectious diseases. Ecology Letters 9:467-484.

Altizer S, Ostfeld RS, Johnson PT, Kutz S, Harvell CD. 2013. Climate change and infectious diseases: from evidence to a predictive framework. Science 341:514-519.

Avenant NL & Nel JA. 1998. Home-range use, activity, and density of caracal in relation to prey density. African Journal of Ecology 36:347-359.

Averbeck GA, Bjork KE, Packer C, Herbst L. 1990. Prevalence of hematozoans in lions (*Panthera leo*) and cheetah (*Acinonyx jubatus*) in Serengeti National Park and Ngorongoro Crater, Tanzania. Journal of Wildlife Diseases 26:392-394.

Bekessy SA, Ennos RA, Burgman MA, Newton AC, Ades PK. 2003. Neutral DNA markers fail to detect genetic divergence in an ecologically important trait. Biological Conservation 110:267-275.

Bekker CP, de Vos S, Taoufik A, Sparagano OA, Jongejan F. 2002. Simultaneous detection of *Anaplasma* and *Ehrlichia* species in ruminants and detection of *Ehrlichiaruminantium* in *Amblyomma variegatum* ticks by reverse line blot hybridization. Veterinary Microbiology 89:223-238.

Bevins SN, Carver S, Boydston EE, Lyren LM, Alldredge M, Logan KA, ... Salman M. 2012. Three pathogens in sympatric populations of pumas, bobcats, and domestic cats: implications for infectious disease transmission. PLoS ONE 7:e31403.

Blanchong JA, Robinson SJ, Samuel MD, Foster JT. 2016. Application of genetics and genomics to wildlife epidemiology. The Journal of Wildlife Management 80:593-608.

Bonneaud C, Pérez-Tris J, Federici P, Chastel O, Sorci G. 2006. Major histocompatibility alleles associated with local resistance to malaria in a passerine. Evolution 60:383-389.

Borghans JA, Beltman JB, De Boer RJ. 2004. MHC polymorphism under host-pathogen coevolution. Immunogenetics 55:732-739.

Bosman AM, Venter EH, Penzhorn BL. 2007. Occurrence of *Babesia felis* and *Babesia leo* in various wild felid species and domestic cats in Southern Africa, based on reverse line blot analysis. Veterinary Parasitology 144:33-38.

Boutin S & Lane JE. 2014. Climate change and mammals: evolutionary versus plastic responses. Evolutionary Applications 7:29-41.

Burger R & Lynch M. 1995. Evolution and extinction in a changing environment: a quantitative-genetic analysis. Evolution 49:151-163.

Carlson SM, Cunningham CJ, Westley PA. 2014. Evolutionary rescue in a changing world. Trends in Ecology & Evolution 29:521-530.

Castro-Prieto A, Wachter B, Sommer S. 2011a. Cheetah paradigm revisited: MHC diversity in the world's largest free-ranging population. Molecular Biology and Evolution 28:1455-1468.

Castro-Prieto A, Wachter B, Melzheimer J, Thalwitzer S, Sommer S. 2011b. Diversity and Evolutionary Patterns of Immune Genes in Free-Ranging Namibian Leopards (*Panthera pardus pardus*). Journal of Heredity 102:653-665.

Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM. 2015. Accelerated modern human-induced species losses: Entering the sixth mass extinction. Science Advances 1:e1400253.

Chapin III FS, Zavaleta ES, Eviner VT, Naylor RL, Vitousek PM, Reynolds HL, Hooper DU, Lavorel S, Sala OE, Hobbie SE, Mack MC. 2000. Consequences of changing biodiversity. Nature 405:234-242.

Charbonnel N & Pemberton J. 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.

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Research 31:3497-3500.

Crutzen PJ & Stoermer EF. 2000. The 'Anthropocene'. Global Change Newsletter 41:17–18.

Daszak P, Cunningham AA, Hyatt AD. 2000. Emerging infectious diseases of wildlife-threats to biodiversity and human health. Science 287:443-449.

Dawkins R & Krebs JR. 1979. Arms races between and within species. Proceedings of the Royal Society of London B: Biological Sciences 205:489-511.

De Boer RJ, Borghans JA, Van Boven M, Keşmir C, Weissing FJ. 2004. Heterozygote advantage fails to explain the high degree of polymorphism of the MHC. Immunogenetics 55:725-731.

Edwards SV & Hedrick PW. 1998. Evolution and ecology of MHC molecules: from genomics to sexual selection. Trends in Ecology & Evolution 13:305-311.

Fahrig L. 2003. Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and Systematics 34:487-515.

Flajnik MF & Kasahara M. 2001. Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. Immunity 15:351-362.

Frankham R. 2005. Genetics and extinction. Biological Conservation 126:131-140.

Garner A, Rachlow JL, Hicks JF. 2005. Patterns of genetic diversity and its loss in mammalian populations. Conservation Biology 19:1215-1221.

Gehrt SD, Riley SP, Cypher BL. 2010. Urban carnivores: ecology, conflict, and conservation. Johns Hopkins University Press, Baltimore, Maryland.

Getz WM, Fortmann-Roe S, Cross PC, Lyons AJ, Ryan SJ, Wilmers CC. 2007. LoCoH: nonparameteric kernel methods for constructing home ranges and utilization distributions. PloS ONE 2:e207.

Gilbert L, Norman R, Laurenson KM, Reid HW, Hudson PJ. 2001. Disease persistence and apparent competition in a three-host community: an empirical and analytical study of large-scale, wild populations. Journal of Animal Ecology 70:1053-1061.

Gubbels JM, De Vos AP, Van der Weide M, Viseras J, Schouls LM, De Vries E, Jongejan F. 1999. Simultaneous detection of bovine, *Theileria* and *Babesia* species by reverse line blot hybridization. Journal of Clinical Microbiology 37:1782-1789.

Hall T. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95–98.

Hansen MM, Olivieri I, Waller DM, Nielsen EE. 2012. Monitoring adaptive genetic responses to environmental change. Molecular Ecology 21:1311-1329.

Harvell CD, Mitchell CE, Ward JR, Altizer S, Dobson AP, Ostfeld RS, Samuel MD. 2002. Climate warming and disease risks for terrestrial and marine biota. Science 296:2158-2162.

Hedrick PW. 2002. Pathogen resistance and genetic variation at MHC loci. Evolution 56:1902-1908.

Hoffmann AA & Sgrò CM. 2011. Climate change and evolutionary adaptation. Nature 470:479-485.

Holderegger R, Kamm U, Gugerli F. 2006. Adaptive vs. neutral genetic diversity: implications for landscape genetics. Landscape Ecology 21:797-807.

Holmes JC, Holmer SG, Ross P, Buntzman AS, Frelinger JA, Hess PR. 2013. Polymorphisms and tissue expression of the feline leukocyte antigen class I loci FLAI-E, FLAI-H, and FLAI-K. Immunogenetics 65:675-689.

Hughes AL & Nei M. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335:167-170.

Inskip C & Zimmermann A. 2009. Human-felid conflict: a review of patterns and priorities worldwide. Oryx 43:18-34.

Jefferies R, Ryan UM, Irwin PJ. 2007. PCR-RFLP for the detection and differentiation of the canine piroplasm species and its use with filter paper-based technologies. Veterinary Parasitology 144:20-27.

King KC & Lively CM. 2012. Does genetic diversity limit disease spread in natural host populations?. Heredity 109:199-203.

Kjøglum S, Larsen S, Bakke HG, Grimholt U. 2006. How specific MHC class I and class II combinations affect disease resistance against infectious salmon anaemia in Atlantic salmon (*Salmo salar*). Fish & Shellfish Immunology 21:431-441.

Kumar M, Shekhar P, Haque S, Mahto D. 2008. Feline babesiosis. Veterinary World 1:120-121.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33:1870-1874.

Lacy RC. 1997. Importance of genetic variation to the viability of mammalian populations. Journal of Mammalogy 78:320-335.

Ladle RJ. 1992. Parasites and sex: catching the Red Queen. Trends in Ecology & Evolution 7:405-408.

Lande R & Shannon S. 1996. The role of genetic variation in adaptation and population persistence in a changing environment. Evolution 50:434-437.

Lawlor DA, Zemmour J, Ennis PD, Parham P. 1990. Evolution of class-I MHC genes and proteins: from natural selection to thymic selection. Annual Review of Immunology 8:23-63.

Librado P & Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Lohm J, Grahn M, Langefors Å, Andersen Ø, Storset A, von Schantz T. 2002. Experimental evidence for major histocompatibility complex-allele-specific resistance to a bacterial infection. Proceedings of the Royal Society of London B: Biological Sciences 269:2029-2033.

Loiseau C, Zoorob R, Garnier S, Birard J, Federici P, Julliard R, Sorci G. 2008. Antagonistic effects of a MHC class I allele on malaria-infected house sparrows. Ecology Letters 11:258-265.

Macdonald D & Loveridge A. 2010. The biology and conservation of wild felids. Volume 2. Oxford University Press, New York.

Mackenstedt U, Jenkins D, Romig T. 2015. The role of wildlife in the transmission of parasitic zoonoses in peri-urban and urban areas. International Journal for Parasitology: Parasites and Wildlife 4:71-79.

Madsen T & Ujvari B. 2006. MHC class I variation associates with parasite resistance and longevity in tropical pythons. Journal of Evolutionary Biology 19:1973-1978.

Matjila PT, Leisewitz AL, Jongejan F, Penzhorn BL. 2008. Molecular detection of tick-borne protozoal and ehrlichial infections in domestic dogs in South Africa. Veterinary Parasitology 155:152-157.

Meyers LA & Bull JJ. 2002. Fighting change with change: adaptive variation in an uncertain world. Trends in Ecology & Evolution 17:551-557.

Murray DL, Kapke CA, Evermann JF, Fuller TK. 1999. Infectious disease and the conservation of free-ranging large carnivores. Animal Conservation 2:241-254.

Nei M & Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Molecular Biology and Evolution 3:418-426.

Nowak MA, Tarczy-Hornoch K, Austyn JM. 1992. The optimal number of major histocompatibility complex molecules in an individual. Proceedings of the National Academy of Sciences 89:10896–10899.

O'Brien SJ, Roelke ME, Marker L, Newman A, Winkler CA, Meltzer D, ... Wildt DE. 1985. Genetic basis for species vulnerability in the cheetah. Science 227:1428-1434.

Paterson S, Wilson K, Pemberton JM. 1998. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (*Ovis aries L.*). Proceedings of the National Academy of Sciences 95:3714-3719.

Penn DJ & Potts WK. 1999. The evolution of mating preferences and major histocompatibility complex genes. The American Naturalist 153:145-164.

Penzhorn BL, Kjemtrup AM, López-Rebollar LM, Conrad PA. 2001. *Babesia leo* n. sp. from lions in the Kruger National Park, South Africa, and its relation to other small piroplasms. Journal of Parasitology 87:681-685.

Penzhorn, BL, Schoeman T, Jacobson LS. 2004. Feline babesiosis in South Africa: a review. Annals of the New York Academy of Sciences 1026:183-186.

Pertoldi C, Bijlsma R, Loeschcke V. 2007. Conservation genetics in a globally changing environment: present problems, paradoxes and future challenges. Biodiversity and Conservation 16:4147-4163.

Pokorny I, Sharma R, Goyal SP, Mishra S, Tiedemann R. 2010. MHC class I and MHC class II DRB gene variability in wild and captive Bengal tigers (*Panthera tigris tigris*). Immunogenetics 62:667-679.

Potts WK & Wakeland EK. 1993. Evolution of MHC genetic diversity: a tale of incest, pestilence and sexual preference. Trends in Genetics 9:408-412.

Pratheek BM, Nayak TK, Sahoo SS, Mohanty PK, Chattopadhyay S, Chakraborty NG. 2014. Mammalian non-classical major histocompatibility complex I and its receptors: important contexts of gene, evolution, and immunity. Indian Journal of Human Genetics 20:129-141.

Prugh LR, Stoner CJ, Epps CW, Bean WT, Ripple WJ, Laliberte AS, Brashares JS. 2009. The rise of the mesopredator. Bioscience 59:779-791.

R Development Core Team. 2016. R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna.

Riley SP, Foley J, Chomel B. 2004. Exposure to feline and canine pathogens in bobcats and gray foxes in urban and rural zones of a national park in California. Journal of Wildlife Diseases 40:11-22.

Riley SP, Pollinger JP, Sauvajot RM, York EC, Bromley C, Fuller TK, Wayne RK. 2006. Fast-track: A southern California freeway is a physical and social barrier to gene flow in carnivores. Molecular Ecology 15:1733-1741.

Rizzoli A, Silaghi C, Obiegala A, Rudolf I, Hubálek Z, Földvári G, Plantard O, Vayssier-Taussat M, Bonnet S, Špitalská E and Kazimírová M. 2014. *Ixodes ricinus* and its transmitted pathogens in urban and peri-urban areas in Europe: new hazards and relevance for public health. Frontiers in Public Health 2:251.

Roelke ME, Martenson JS, O'Brien SJ. 1993. The consequences of demographic reduction and genetic depletion in the endangered Florida panther. Current Biology 3:340-350.

Nei M & Rooney AP. 2005. Concerted and birth-and-death evolution of multigene families. Annual Review of Genetics 39:121-152.

Ruiz-López MJ, Monello RJ, Schuttler SG, Lance SL, Gompper ME, Eggert LS. 2014. Major Histocompatibility Complex, demographic, and environmental predictors of antibody presence in a free-ranging mammal. Infection, Genetics and Evolution 28:317-327.

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

Schad J, Ganzhorn JU, Sommer S. 2005. Parasite burden and constitution of major histocompatibility complex in the Malagasy mouse lemur, *Microcebus murinus*. Evolution 59:439-450.

Serieys LE, Lea A, Pollinger JP, Riley SP, Wayne RK. 2014. Disease and freeways drive genetic change in urban bobcat populations. Evolutionary Applications 8:75-92.

Schoeman T, Lobetti RG, Jacobson LS, Penzhorn BL. 2001. Feline babesiosis: signalment, clinical pathology and concurrent infections. Journal of the South African Veterinary Association 72:4-11.

Simpson E. 1988. Function of the MHC. Immunology 64:27-30.

Smith MM & Hoffman SM. 2001. A class I MHC locus compared among felid species. Mammalian Genome 12:394-396.

Sommer S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Frontiers in Zoology 2:16.

Soulé ME, Bolger DT, Alberts AC, Wrights J, Sorice M, Hill S. 1988. Reconstructed dynamics of rapid extinctions of chaparral-requiring birds in urban habitat islands. Conservation Biology 2:75-92.

Spielman D, Brook BW, Briscoe DA, Frankham R. 2004. Does inbreeding and loss of genetic diversity decrease disease resistance? Conservation Genetics 5:439-448.

Spurgin LG & Richardson DS. 2010. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proceedings of the Royal Society of London B: Biological Sciences 277:979-988.

Stevens L, Yan G, Pray LA. 1997. Consequences of inbreeding on invertebrate host susceptibility to parasitic infection. Evolution 51:2032-2039.

Takahata N & M Nei. 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics 124:967-978.

Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BF, De Siqueira MF, Grainger A, Hannah L, Hughes L. 2004. Extinction risk from climate change. Nature 427:145-148.

Urban MC. 2015. Accelerating extinction risk from climate change. Science 348:571-573.

Van Tienderen PH, de Haan AA, Van der Linden CG, Vosman B. 2002. Biodiversity assessment using markers for ecologically important traits. Trends in Ecology & Evolution 17:577-582.

Van Valen L. 1977. The red queen. The American Naturalist 111:809-810.

Vitousek PM, Mooney HA, Lubchenco J, Melillo JM. 1997. Human domination of Earth's ecosystems. Science 277:494-499.

Wang X, Wei K, Zhang Z, Xu X, Zhang W, Shen F, ... Yue B. 2009. Major histocompatibility complex Class II DRB exon 2 diversity of the Eurasian lynx (*Lynx lynx*) in China. Journal of Natural History 43:245-257.

Wegner KM, Reusch TBH, Kalbe M. 2003a. Multiple parasites are driving major histocompatibility complex polymorphism in the wild. Journal of Evolutionary Biology 16:224-232.

Wegner KM, Kalbe M, Kurtz J, Reusch TBH, Milinski M. 2003b. Parasite selection for immunogenetic optimality. Science 301:1343.

Westerdahl H, Waldenström J, Hansson B, Hasselquist D, von Schantz T, Bensch S. 2005. Associations between malaria and MHC genes in a migratory songbird. Proceedings of the Royal Society of London B: Biological Sciences 272:1511-1518.

Yuhki N & O'Brien SJ. 1990. DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history. Proceedings of the National Academy of Sciences 87:836-840.

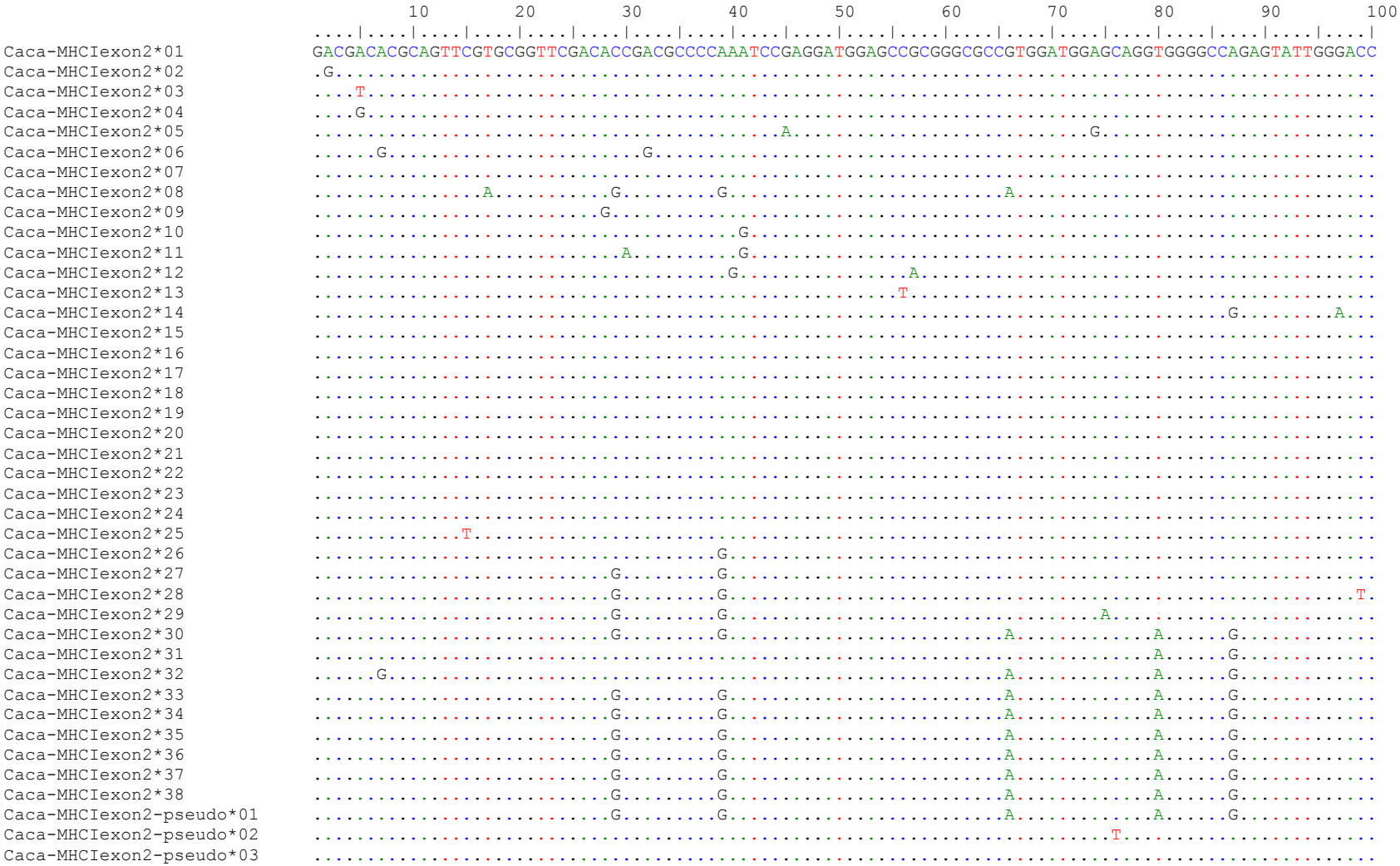
Yuhki N & O'Brien SJ. 1994. Exchanges of short polymorphic DNA segments predating speciation in feline major histocompatibility complex class I genes. Journal of Molecular Evolution 39:22–33.

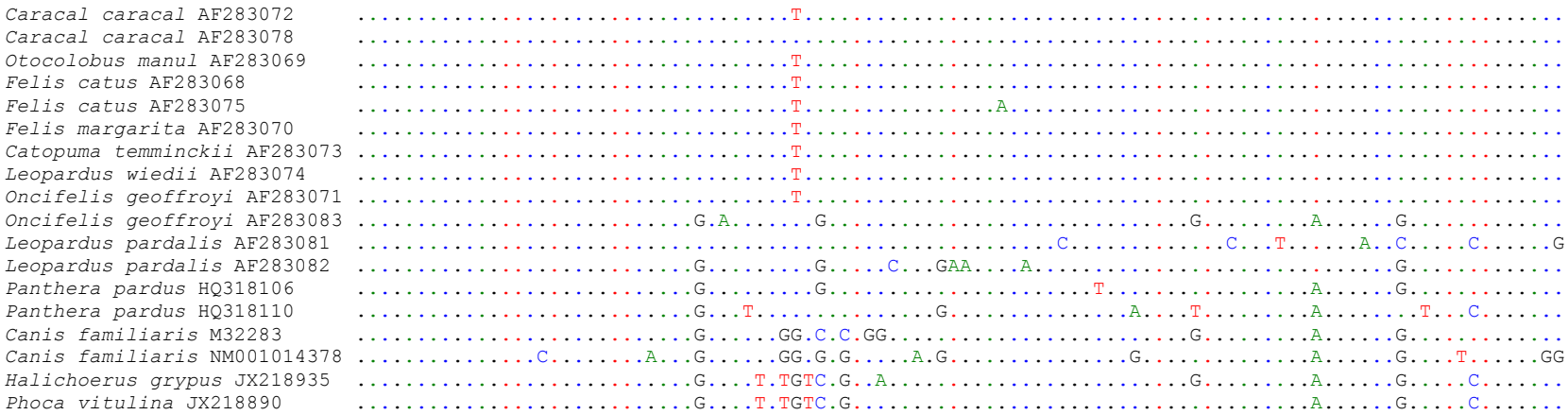
Yuhki N, Mullikin JC, Beck T, Stephens R, O'Brien SJ. 2008. Sequences, annotation and single nucleotide polymorphism of the major histocompatibility complex in the domestic cat. PLoS ONE 3:e2674.

Zhang M, & He H. 2013. Parasite-mediated selection of major histocompatibility complex variability in wild brandt's voles (*Lasiopodomys brandtii*) from Inner Mongolia, China. BMC Evolutionary Biology 13:149.

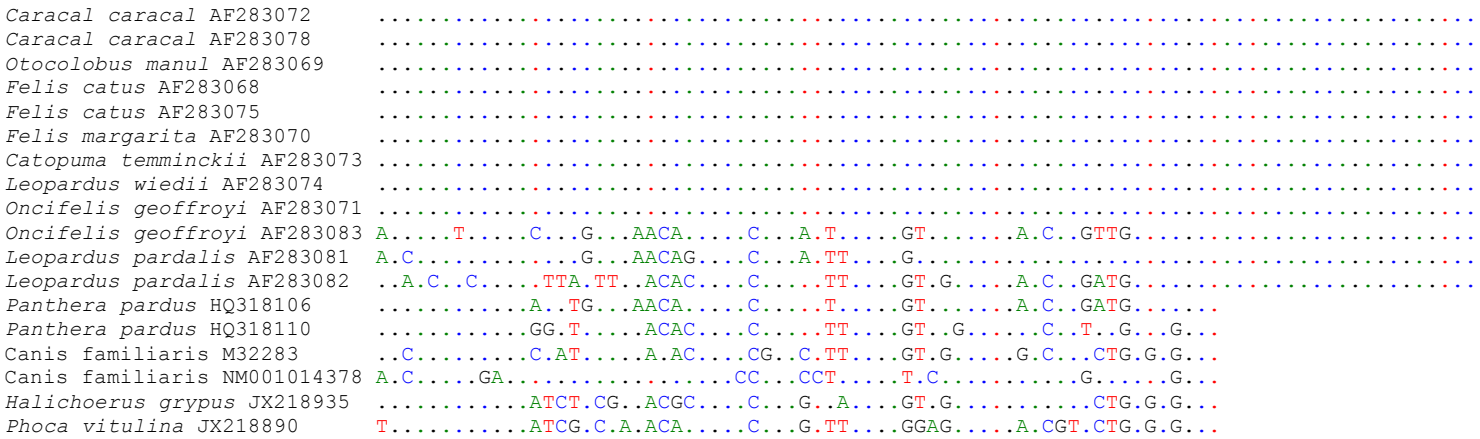
Appendix A

Figure S1. Alignment of the unique MHC class I exon 2 nucleotide sequences detected in this study, together with published sequences from related species used in phylogenetic analysis. Dots indicate shared identity with the first allele.




Caracal caracal AF283072
Caracal caracal AF283078
Otocolobus manul AF283069
Felis catus AF283068
Felis catus AF283075
Felis margarita AF283070
Catopuma temminckii AF283073
Leopardus wiedii AF283074
Oncifelis geoffroyi AF283071
Oncifelis geoffroyi AF283083
Leopardus pardalis AF283081
Leopardus pardalis AF283082
Panthera pardus HQ318106
Panthera pardus HQ318110
Canis familiaris M32283
Canis familiaris NM001014378
Halichoerus grypus JX218935
Phoca vitulina JX218890


110 120 130 140 150 160 170 180
Caca-MHCIexon2*01 GGGAGACGCGGATCGCCAAGGTGCACGCAGAGATTGCCCGACCGAACCTGCAGACCGCCCTCCGCTACTACAACCAGAGCGAGTCC
Caca-MHCIexon2*02
Caca-MHCIexon2*03
Caca-MHCIexon2*04
Caca-MHCIexon2*05
Caca-MHCIexon2*06
Caca-MHCIexon2*07
Caca-MHCIexon2*08
Caca-MHCIexon2*09
Caca-MHCIexon2*10
Caca-MHCIexon2*11
Caca-MHCIexon2*12
Caca-MHCIexon2*13
Caca-MHCIexon2*14
Caca-MHCIexon2*15
Caca-MHCIexon2*16
Caca-MHCIexon2*17
Caca-MHCIexon2*18
Caca-MHCIexon2*19
Caca-MHCIexon2*20
Caca-MHCIexon2*21
Caca-MHCIexon2*22
Caca-MHCIexon2*23
Caca-MHCIexon2*24
Caca-MHCIexon2*25
Caca-MHCIexon2*26
Caca-MHCIexon2*27
Caca-MHCIexon2*28
Caca-MHCIexon2*29
Caca-MHCIexon2*30
Caca-MHCIexon2*31
Caca-MHCIexon2*32
Caca-MHCIexon2*33
Caca-MHCIexon2*34
Caca-MHCIexon2*35
Caca-MHCIexon2*36
Caca-MHCIexon2*37
Caca-MHCIexon2*38
Caca-MHCIexon2-pseudo*01
Caca-MHCIexon2-pseudo*02
Caca-MHCIexon2-pseudo*03

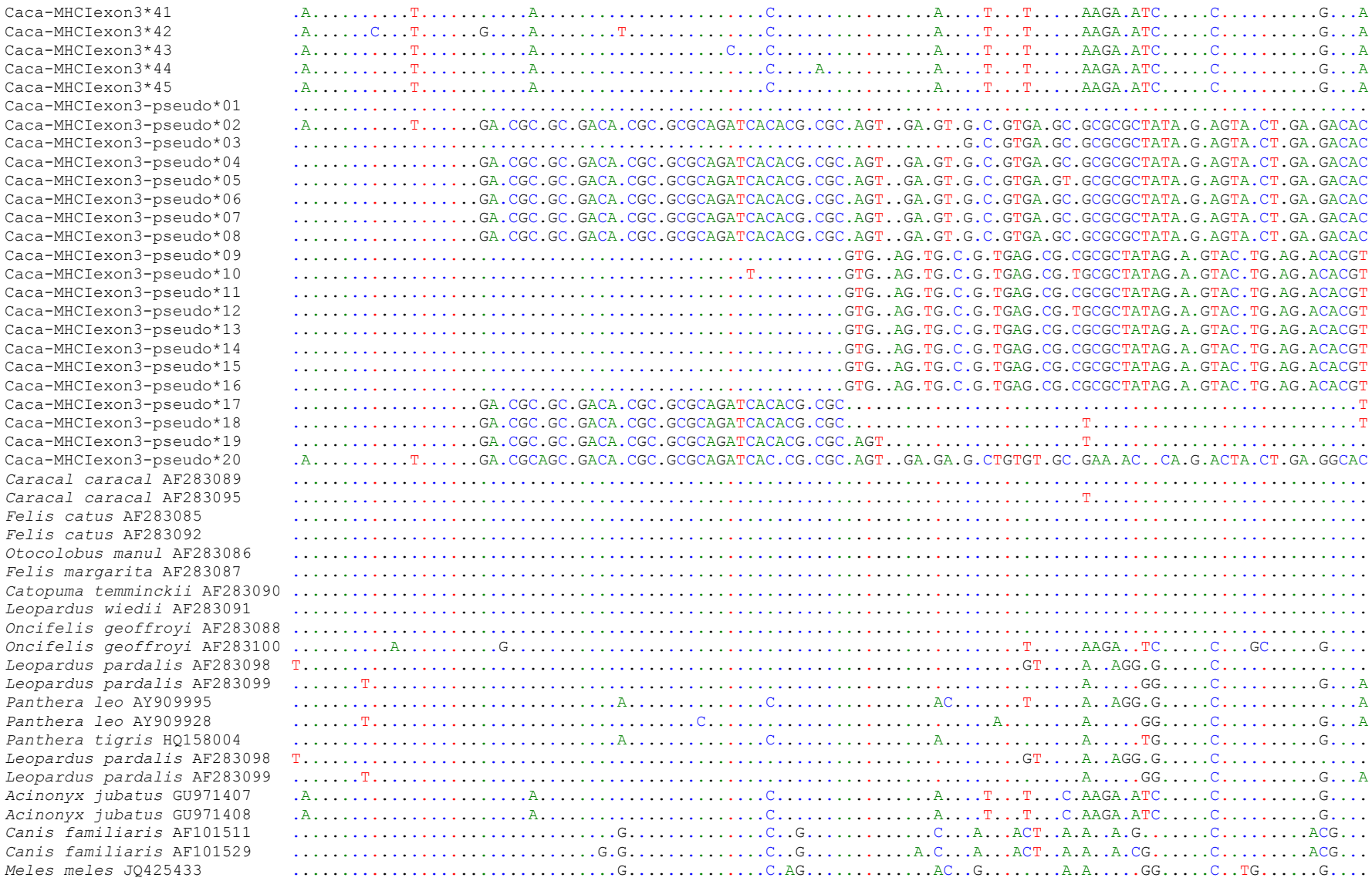

Caracal caracal AF283072
Caracal caracal AF283078
Otocolobus manul AF283069
Felis catus AF283068
Felis catus AF283075
Felis margarita AF283070
Catopuma temminckii AF283073
Leopardus wiedii AF283074
Oncifelis geoffroyi AF283071
Oncifelis geoffroyi AF283083
Leopardus pardalis AF283081
Leopardus pardalis AF283082
Panthera pardus HQ318106
Panthera pardus HQ318110
Canis familiaris M32283
Canis familiaris NM001014378
Halichoerus grypus JX218935
Phoca vitulina JX218890

Appendix B

Figure S2. Alignment of the unique MHC class I exon 3 nucleotide sequences detected in this study, together with published sequences from related species used in phylogenetic analysis. Dots indicate shared identity with the first allele.


Caca-MHCIexon3*41
Caca-MHCIexon3*42
Caca-MHCIexon3*43
Caca-MHCIexon3*44
Caca-MHCIexon3*45
Caca-MHCIexon3-pseudo*01
Caca-MHCIexon3-pseudo*02
Caca-MHCIexon3-pseudo*03
Caca-MHCIexon3-pseudo*04
Caca-MHCIexon3-pseudo*05
Caca-MHCIexon3-pseudo*06
Caca-MHCIexon3-pseudo*07
Caca-MHCIexon3-pseudo*08
Caca-MHCIexon3-pseudo*09
Caca-MHCIexon3-pseudo*10
Caca-MHCIexon3-pseudo*11
Caca-MHCIexon3-pseudo*12
Caca-MHCIexon3-pseudo*13
Caca-MHCIexon3-pseudo*14
Caca-MHCIexon3-pseudo*15
Caca-MHCIexon3-pseudo*16
Caca-MHCIexon3-pseudo*17
Caca-MHCIexon3-pseudo*18
Caca-MHCIexon3-pseudo*19
Caca-MHCIexon3-pseudo*20
Caracal caracal AF283089
Caracal caracal AF283095
Felis catus AF283085
Felis catus AF283092
Otocolobus manul AF283086
Felis margarita AF283087
Catopuma temminckii AF283090
Leopardus wiedii AF283091
Oncifelis geoffroyi AF283088
Oncifelis geoffroyi AF283100
Leopardus pardalis AF283098
Leopardus pardalis AF283099
Panthera leo AY909995
Panthera leo AY909928
Panthera tigris HQ158004
Leopardus pardalis AF283098
Leopardus pardalis AF283099
Acinonyx jubatus GU971407
Acinonyx jubatus GU971408
Canis familiaris AF101511
Canis familiaris AF101529
Meles meles JQ425433


120 130 140 150 160 170 180 190 200 210
Caca-MHCIexon3*01 CGAGGACCTGCGCTCCTGGACCGCGGCGGACACCGCGGCGCAGATCACACGCCGCAAGTGGGAGGTGGCCGGTGAGGCGGCGCGCTATAGGAAGTACCTGGAGGACACG
Caca-MHCIexon3*02
Caca-MHCIexon3*03
Caca-MHCIexon3*04
Caca-MHCIexon3*05
Caca-MHCIexon3*06
Caca-MHCIexon3*07
Caca-MHCIexon3*08
Caca-MHCIexon3*09
Caca-MHCIexon3*10
Caca-MHCIexon3*11
Caca-MHCIexon3*12
Caca-MHCIexon3*13
Caca-MHCIexon3*14
Caca-MHCIexon3*15
Caca-MHCIexon3*16
Caca-MHCIexon3*17
Caca-MHCIexon3*18
Caca-MHCIexon3*19
Caca-MHCIexon3*20
Caca-MHCIexon3*21
Caca-MHCIexon3*22
Caca-MHCIexon3*23
Caca-MHCIexon3*24
Caca-MHCIexon3*25
Caca-MHCIexon3*26
Caca-MHCIexon3*27
Caca-MHCIexon3*28
Caca-MHCIexon3*29
Caca-MHCIexon3*30
Caca-MHCIexon3*31
Caca-MHCIexon3*32
Caca-MHCIexon3*33
Caca-MHCIexon3*34
Caca-MHCIexon3*35
Caca-MHCIexon3*36
Caca-MHCIexon3*37
Caca-MHCIexon3*38
Caca-MHCIexon3*39
Caca-MHCIexon3*40


Caca-MHCIexon3*41
Caca-MHCIexon3*42
Caca-MHCIexon3*43
Caca-MHCIexon3*44
Caca-MHCIexon3*45
Caca-MHCIexon3-pseudo*01
Caca-MHCIexon3-pseudo*02
Caca-MHCIexon3-pseudo*03
Caca-MHCIexon3-pseudo*04
Caca-MHCIexon3-pseudo*05
Caca-MHCIexon3-pseudo*06
Caca-MHCIexon3-pseudo*07
Caca-MHCIexon3-pseudo*08
Caca-MHCIexon3-pseudo*09
Caca-MHCIexon3-pseudo*10
Caca-MHCIexon3-pseudo*11
Caca-MHCIexon3-pseudo*12
Caca-MHCIexon3-pseudo*13
Caca-MHCIexon3-pseudo*14
Caca-MHCIexon3-pseudo*15
Caca-MHCIexon3-pseudo*16
Caca-MHCIexon3-pseudo*17
Caca-MHCIexon3-pseudo*18
Caca-MHCIexon3-pseudo*19
Caca-MHCIexon3-pseudo*20
Caracal caracal AF283089
Caracal caracal AF283095
Felis catus AF283085
Felis catus AF283092
Otocolobus manul AF283086
Felis margarita AF283087
Catopuma temminckii AF283090
Leopardus wiedii AF283091
Oncifelis geoffroyi AF283088
Oncifelis geoffroyi AF283100
Leopardus pardalis AF283098
Leopardus pardalis AF283099
Panthera leo AY909995
Panthera leo AY909928
Panthera tigris HQ158004
Leopardus pardalis AF283098
Leopardus pardalis AF283099
Acinonyx jubatus GU971407
Acinonyx jubatus GU971408
Canis familiaris AF101511
Canis familiaris AF101529
Meles meles JQ425433

Appendix C

Table S1. Summary of MHC class I exon 2 and exon 3 diversity measures for each caracal individual included in this study. Diversity measures were calculated using DNAsp v5.10 (Librado & Rozas 2009).

Individual	Exon 2				Exon 3			
	No. of alleles	Gene diversity	Nucleotide diversity	Average no. of nucleotide differences	No. of alleles	Gene diversity	Nucleotide diversity	Average no. of nucleotide differences
C1	4	0.89	0.06	11.3	6	0.85	0.06	13.1
C2	3	0.55	0.04	8.2	5	0.85	0.06	13.0
C3	3	0.76	0.05	8.6	7	0.77	0.05	10.5
C4	6	0.68	0.05	8.9	3	0.77	0.03	6.0
C5	3	0.85	0.01	1.8	1	0.17	0.00	0.2
C6	2	0.49	0.07	12.6	4	0.73	0.07	15.1
C7	2	0.30	0.04	7.9	7	0.67	0.02	5.2
C8	4	0.88	0.01	1.4	4	0.44	0.00	0.6
C9	2	0.49	0.07	12.6	6	0.94	0.04	8.7
C10	5	0.80	0.07	13.5	6	0.88	0.07	15.7
C11	3	0.68	0.03	5.2	3	0.68	0.03	5.7
C12	3	0.82	0.05	8.7	3	0.74	0.04	9.6
C13	3	0.44	0.05	8.3	2	0.73	0.01	1.1
C16	3	0.58	0.01	1.2	3	0.73	0.00	0.9
C17	2	0.46	0.00	0.5	7	0.91	0.01	1.8
C18	2	0.17	0.01	0.8	3	0.46	0.00	0.7
C19	2	0.46	0.03	4.7	3	0.58	0.03	6.9
C20	2	0.65	0.04	7.7	3	0.74	0.04	9.7
C21	3	0.80	0.07	12.8	7	0.89	0.07	16.2
C22	4	0.80	0.07	12.8	5	0.73	0.04	9.6
C24	6	0.88	0.08	15.1	3	0.76	0.07	15.5
C25	2	0.68	0.00	0.8	1	0.58	0.00	0.8
C26	2	0.58	0.00	0.7	3	0.77	0.01	1.0
C27	3	0.58	0.01	1.0	3	0.85	0.01	1.3
C28	6	0.77	0.05	8.7	5	0.91	0.01	3.0
C29	4	0.58	0.01	1.0	5	0.46	0.00	0.7