
**PHYLOGEOGRAPHIC STRUCTURE
OF THE HONEY BADGER
(*MELLIVORA CAPENSIS*)**

JAMES I. RHODES

**THESIS PRESENTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE (ZOOLOGY) AT THE UNIVERSITY OF
STELLENBOSCH**



**SUPERVISOR:
CONRAD A. MATTHEE**

DECEMBER 2006

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature:

Date:

ABSTRACT

The aim of this study was to investigate the phylogeographic structuring of the honey badger, *Mellivora capensis*, a highly mobile medium sized carnivore with an extensive distribution throughout sub-Saharan Africa extending into the Middle East and India. Particular focus was placed on providing preliminary data potentially useful for the development of translocation policies for this species in southern Africa. Where possible, genetic results were also compared with current trinomial designations to determine whether subspecies status given to geographical groupings was supported by the genetic data. Mitochondrial control region sequence data was obtained for most a selection of specimen's available while nuclear microsatellite variation was determined for a subset of individuals where there were sufficient sample sizes available. Phylogeographic structuring of the maternal mitochondrial lineage was initially obscured by the co-amplification of a closely related numt. To overcome co-amplification, the numt was identified and mtDNA specific primers were designed. Mitochondrial DNA results are based on the most variable 230 bp of the control region (sequenced for 78 individuals) while five polymorphic nuclear microsatellite markers were scored (for 55 individuals). Analysis, employing both nuclear and mitochondrial data, showed that although a pattern of isolation by distance can be detected, there was evidence for the presence of phylogeographic structuring between eastern and southern Africa. This could be interpreted as due to vicariance, probably associated by rifting and climatic oscillations during the Pleistocene. Analyses support the identification of distinct management units for eastern and southern African populations although some evidence exists for secondary introgression between these two regions. Following this, we recommend that translocations between these broad geographic areas should be avoided. Within these geographic areas, because of a general pattern of isolation by distance, we recommend that individuals for translocations come from geographically proximate populations. In some instances, phylogeographic structuring is concordant with subspecies designations but additional sampling will be needed to make any firm taxonomic conclusions.

OPSOMMING

Die doelwit van hierdie studie was om die filogeografiese struktuur van die ratel, *Mellivora capensis*, 'n hoogs bewegende medium groot karnivoor met 'n wydverspreide distribusie deur sub-Sahara Afrika wat strek tot in die Midde Ooste en India, te bepaal. Spesifieke fokus is geplaas op die voorsiening van preliminêre data potensieel bruikbaar vir die ontwikkeling van verplasing strategieë vir hierdie spesie in suidelike Afrika. Genetiese resultate is, waar moontlik, vergelyk met huidige drieledige kategorieë om te bepaal of subspesies ondersteun word deur die genetiese data. Mitochondriale 'control region' DNS volgorde data was verkry vir die meeste van die monsters beskikbaar en kern mikrosatelliet variasie was bepaal vir 'n gedeelte van individue waar voldoende monster groottes beskikbaar was. Filogeografiese struktuurering van die materne mitochondriale merker was oorspronklik versteek deur die ko-amplifikasie van 'n naby verwante 'numt'. Om die ko-amplifikasie te oorkom is die 'numt' geïdentifiseer en mtDNS spesifieke voorvoeders is ontwerp. Mitochondriale DNS resultate is gebaseer op die mees veranderlike 230 bp van die 'control region' (waar die DNS volgorde vir 78 individue bepaal is) en vyf polimorfiese kern mikrosatelliet merkers (in 55 individue). Analises, wat gebruik maak van kern en mitochondriale data, toon wel 'n patroon van isolasie deur afstand, maar ook 'n duidelike sigbare filogeografiese struktuurering tussen oostelike en suidelike Afrika. Hierdie is geïnterpreteer as vikariansie, heel waarskynlik ge-assosieer deur berg verskuiwings en klimaatsveranderinge deur die Pleistocene. Analises ondersteun die identifikasie van definitiewe verkillende bestuurseenhede vir oostelike en suidelike Afrika maar sekere bewyse bestaan dat sekondêre introgressie tussen streke bestaan. Dit word aanbeveel dat translokasies tussen hierdie geografiese areas voorkom moet word. Binne geografiese areas, as gevolg van 'n algemene patroon van isolasie deur afstand, is dit aanbeveel dat individue vir verplasing van nabygeleë populasies moet wees. In sommige gevalle het filogeografiese struktuurering ooreen gestem met subspesies kategorieë, maar verdere materiaal is nodig voor definitiewe taksonomiese besluite geneem kan word.

ACKNOWLEDGEMENTS

I would like to thank a number of people that contributed to this study. First and foremost I would like to thank my supervisor Dr. Conrad Matthee for his intellectual input, understanding and advice. I am also appreciative to a number of people associated with the Evolutionary Genomics Group for their input and support, including Dr. Peter Teske, Dr. Krystal Tolley, Dr. Savel Daniels, Dr. Gavin Gous, Dr. Gauthier Dobigny, Dr. Paul Waters, Dr. Victor Rambau, Woody Coterill, Ronelle Vervey, Sandi Willows-Munro, Jane Sakwa Makokha, Keshni Gopal, and Hanneline Smit. In particular I would like to thank Dr. Bettine Jansen van Vuuren for her intellectual and emotional support.

This study would not have been possible without honey badger samples and I would like to thank those who contributed samples. Special thanks needs to go to Dr. Colleen Begg and Keith Begg who provided the majority of the samples for this study. In addition I would like to thank them for their assistance with the project and personal encouragement. I would also like to thank Woody Coterill (The Natural History Museum, Zimbabwe), Teresa Kearney (Transvaal Museum), Lloyd Wingate (Amathole Museum), Paula Jenkins (The Natural History Museum, London) and the Nairobi Museum (Kenya) for the provision of samples.

Lastly but most importantly I would like to thank my family (Mom, Dad, Deirdré, Simon, Allan, Désirée and Dylan) for their support throughout (emotional and financial). In particular I would like to thank Désirée and Dylan who sacrificed so much during the completion of this study.

To Dad

CONTENTS

DECLARATION	II
ABSTRACT	III
OPSOMMING	IV
ACKNOWLEDGEMENTS	V
LIST OF TABLES	X
LIST OF FIGURES	XI
CHAPTER 1: INTRODUCTION	1
Taxonomy	1
Synopsis of mustelid systematics.....	1
Systematic placement of the honey badger <i>Mellivora capensis</i>	3
Subspecies designations.....	4
General Biology	8
Overview	8
Conservation	12
Potential factors influencing gene flow among <i>Mellivora capensis</i> populations	14
Choice of genetic markers	17
Phylogeography	17
The mitochondria	17
Phylogeographic inferences based on genealogies	19
Nuclear markers	19
Aim	22
CHAPTER 2: MATERIALS AND METHODS	23
Genetic samples	23
DNA extraction	26
Control region sequencing	26

Control region data analysis	27
Phylogenetic analysis.....	27
Chimerism.....	28
Phylogenetic relationships relative to the numt	29
Network reconstruction.....	30
Population structure	30
Genetic diversity and isolation by distance	31
Population Demographic Parameters and Selective Neutrality	31
Microsatellite analysis	34
Microsatellite genotyping	34
Microsatellite data analysis.....	34
Population structure	34
CHAPTER 3: RESULTS	36
Mitochondrial analysis	36
Control region diversity	36
Mitochondrial diversity.....	36
Chimerism.....	37
MtDNA Phylogeography	37
Population structure	42
Population history	46
Microsatellite analysis	50
Microsatellite variation	50
Genetic variation.....	51
CHAPTER 4: DISCUSSION	56
Intraspecific variation	56
Phylogeography	58
Implications for conservation	65
Subspecies descriptions	66
Conservation Implications at a Wider scale.....	67
CHAPTER 5: CO-AMPLIFICATION OF NUMT SEQUENCES AND THE CONFIRMATION OF MTDNA CONTROL REGION SEQUENCES.....	68
Introduction.....	68
Materials and Methods.....	71
Results and Discussion.....	73

REFERENCES.....	77
APPENDIX.....	100

LIST OF TABLES

	Page
Table 1 Primers employed in this study with their respective nucleotide sequence and their origin.....	27
Table 2 Statistics generated employing groupings identified by SAMOVA. F_{ST} and Φ_{ST} values generated from mitochondrial data. Statistically significant results (generated from 1000 random permutation tests) are shown in bold.....	43
Table 3 F_{ST} and Φ_{ST} values between the honey badger populations used in this study. Figures above the diagonal represent F_{ST} values between populations and those below the diagonal represent Φ_{ST} values between populations based on mitochondrial data. Abbreviations as in Fig.2.	45
Table 4 Results of tests for selective neutrality (Fu's F_s , F_u and Li's D^* , F_u and Li's F^* and Tajima's D) and the estimated value for the population parameter τ . (τ = the time passed since the expansion). SSD: statistic testing for significant departure of observed data from that expected from a sudden expansion model.	47
Table 5 Genetic diversity indices and numbers of <i>Mellivora</i> specimens used in the population analyses.	51
Table 6 Number of alleles detected in <i>Mellivora</i> for the respective loci. For comparison the number of alleles detected in the taxa from which these loci were originally isolated are shown. The number of alleles detected in <i>Gulo gulo</i> loci (Gg-x) was determined from 16 individuals, those from <i>Martes americana</i> locus (Ma1) from 30 individuals and those at <i>Taxidea taxus</i> loci (Tt-x) from 19 individuals.	51
Table 7 F_{ST} and R_{ST} values between the honey badger populations used in this study based on the microsatellite data. Figures above the diagonal represent F_{ST} values between populations and those below the diagonal represent R_{ST} values between populations. Note that the Kenyan population was not included in the microsatellite analysis. Abbreviations as in Fig. 2.	54
Table 8 Statistics generated employing groupings identified by SAMOVA. F_{ST} and R_{ST} values generated from microsatellite data. Statistically significant results (generated from 1000 random permutation tests) are shown in bold.....	55

LIST OF FIGURES

	Page
Fig. 1 Geographic distribution of the honey badger, <i>M. capensis</i> (Map redrawn from Finn 1929; Smithers 1983; Skinner & Smithers 1990; Kingdon 1997; Baryshnikov 2000).....	9
Fig. 2 Map showing sampling localities for the honey badger including the six putative populations. Black circles indicate samples used in both mitochondrial and microsatellite analyses, samples shaded gray indicate those used for mitochondrial analysis only and the samples shaded in white represent individuals used only in the microsatellite analysis. The Indian specimen is not indicated and only included in the mtDNA analyses. Dashed lines indicate subspecies designations in Africa (modified from Vanderlaar & Ten Hwang 2003) and the solid line shows the location of the lower Zambezi River. Subspecies designations are indicated by capital letters (A-E) and population abbreviations used in text are as follows; North-western Zambia: NWZA, Kenya: KENY, Northern Mozambique: NMOZ, Southern Mozambique: SMOZ, Kalahari, South Africa: KALA and Cape, South Africa: CAPE.	25
Fig. 3 Neighbor joining phylogenetic relationships between 41 haplotypes identified in the present study. The tree is based on HKY85 distances, with an assumed proportion of invariant sites (0.647) and a gamma rate for variant sites (0.223). Bootstrap values above 60 are shown. Colours are representative of the sampling locality and correspond to the inset map.....	39
Fig. 4 Haplotype network of the 41 control region haplotypes identified in this study. Small black circles are unsampled or inferred haplotypes. Unless otherwise indicated the line connecting haplotypes represents one mutational change. Colours are representative of the sampling locality and correspond to Fig. 3.	40
Fig. 5 a.) Phylogenetic relationships of haplotypes using parsimony methods with the numt sequence to root the tree. Clades identified in previous analyses are indicated. Bootstrap values above 60 are shown. b.) Haplotype network, as above, but including the numt in the network reconstruction to indicate the position of the numt relative to the other haplotypes.....	41
Fig. 6 Relationship between genetic distance and geographic distance. D_A based on sampled populations incorporating mitochondrial data. Triangles represent comparisons between main geographical regions, namely east and southern Africa, while circles represent comparisons within main geographical regions. Squares represent comparisons between the populations and the single haplotypes found in Chad, Israel and India. Solid line represents the fitted regression line calculated for populations only.....	44
Fig. 7 Mismatch nucleotide distributions for the 41 mitochondrial DNA haplotypes. The solid line represents the observed distribution; the broken line the expected distribution under a model of growth/decline and the dotted line represents the expected distribution under a model of stasis.....	46
Fig. 8 Mismatch nucleotide distributions for: a.) clades 1 & 2; b.) clades 3 & 4; c.) clade 1 and d.) clade 4. The solid line represents the observed distribution, the broken line the expected distribution under a model of growth/decline and the dotted line represents the expected distribution under a model of stasis.	49

Fig. 9 Phylogenetic trees generated using a.) DAS distances between populations and b.) DAS distances between individuals. Both trees are based on the five microsatellite loci used in this study. Colours correspond to those used in Fig. 3.53

Fig. 10 Relationship between genetic distance and geographic distance. F_{ST} based on sampled populations incorporating microsatellite data. Diamonds represent comparisons between populations. Note that the Kenyan population was not included in the microsatellite analysis. Solid line represents the fitted regression line.55

Fig. 11 Alignment of 400bp of the presumed numt and authentic mitochondrial sequence, with ambiguities between sequences shown with a *. Both sequences were obtained from the same individual. Underlined regions on the mitochondrial sequence indicate where designed primers are situated.73

CHAPTER 1: INTRODUCTION

Taxonomy

Synopsis of mustelid systematics

The honey badger, *Mellivora capensis*, belongs to the mammalian order Carnivora, which has been thought to originate approximately 70-50 million years (My) before present (Carroll 1988). Currently the order comprises over 270 extant species (Wozencraft 1993). There is broad based consensus that the Carnivora is divided into two monophyletic groups, the Feliformia and the Caniformia (Wyss & Flynn 1993). The former includes the “cat-like” while the latter “dog-like” group, includes the family Mustelidae to which the honey badger belongs. Among the order Carnivora, the Mustelidae is the largest and most diverse family comprising 24 extant genera and 65 species (Wozencraft 1993; Nowak 1999). Members of the Mustelidae occur throughout Eurasia, Africa and America (Nowak 1999) and show a wide range of ecomorphological diversity. Species in this group are adapted to habitats ranging from deserts to completely aquatic environments.

The monophyly of the Mustelidae is supported primarily on morphological features, including the ventral closure of the suprameatal fossa, the loss of the carnassial notch on the upper 4th premolar, the loss of the upper second molar and the enlarged scent glands (Tedford 1976; Schmidt-Kittler 1981; Martin 1989; Wozencraft 1989; Bryant *et al.* 1993). More recent molecular studies, however, have indicated that this group may not be monophyletic (Ledje & Arnason 1996; Dragoo & Honeycutt 1997; Marmi *et al.* 2004). Initially, 15 mustelid subfamilies were proposed (Pocock 1921) but currently between four and seven subfamilies are recognised (Anderson 1989; Eisenberg 1989; Wozencraft 1989, 1993). The four most commonly recognized of these subfamilies are; the Mephitinae (skunks), Melinae (true badgers), Lutrinae (otters) and Mustelinae (the remaining mustelids including weasels, martens and the wolverine). Due to the distinctness of some taxa, however, additional subfamilies have also been proposed

(Anderson 1989; Simpson 1945, Wozencraft 1993). For example, Wozencraft (1989; 1993) placed the American badger, *Taxidea taxus*, in its own family the Taxiidinae because this taxon contains basicranial features absent in the Old World badgers. Likewise, the Mellivorinae represented by the monotypic honey badger, is recognized by some as an additional subfamily (see below).

Systematic placement of the honey badger *Mellivora capensis*

The honey badger was originally described by Schreber (1776) as *Viverra capensis*. Due to several unique morphological features of the honey badger the generic name was changed soon thereafter by Storr (1780), who suggested *Mellivora* derived from the Latin words “mel”, meaning “honey”, and “voro”, meaning ‘to devour’. The specific epithet remained and refers to the region, namely the Cape of Good Hope where Schreber collected the type specimen. Numerous synonyms have been given to this taxon and these are summarized by Vanderhaar and Ten Hwang (2003). The placement of the honey badger within the Mustelidae is still unresolved. This species was originally identified as a badger because of its resemblance to the European badger (*Meles meles*) pertaining to traits such as colouration, body structure and gait (Smithers 1983).. It was initially placed in the subfamily Melinae, but was later reassigned to the Mustelinae based on skull and tooth morphology (Rosevear 1974). A close evolutionary association has also been suggested between the the honey badger and the wolverine ,*Gulo gulo* (Johnstone-Scott 1981; Harrison & Bates 1991). The vacillations in taxonomy is probably indicative of the fact that the badger-like ecomorph has evolved a number of times within the Mustelidae, suggestive of convergent or parallel morphological evolution due to similar life histories (Bryant *et al.* 1993; Neal & Cheeseman 1996).

Currently the taxon is placed in the subfamily Mellivorinae, of which it is the only extant species (Wozencraft 1989). Simpson’s (1945) recognition of the Mellivorinae was not based on cladistics but rather on the ‘isolation’ of *Mellivora* (and ancestral *Eomellivora*) from other mustelids. It has been suggested that the unresolved phylogenetic relationships among members of the Mustelidae is an effect of the rapid basal radiation of this family (Koepfli & Wayne 1998).

The Mellivorinae have been recorded from the late Tertiary and Quaternary. Although the fossil records for the group are poor (Hendey 1978), it is likely that honey badgers have been present in Africa since the late Miocene (the oldest recorded fossil of this genus is

approximately 10 My old and it was recorded in the Ngorora Formation in Kenya; Bishop & Pickford 1975). This fossil, coupled with indications that the Asian *Eomellivora* and *Promellivora* are not directly related to *Mellivora*, indicates that *Mellivora* may have originated in Africa. Further support for an African origin for this genus is found in a middle Miocene mellivorine unearthed in southwestern Africa (Hendey 1978) and it is likely that *M. benfieldi*, found 7 - 3.5 My ago in southern Africa, represents the ancestral form to the extant species (Hendey 1978).

Subspecies designations

The identification of populations with independent evolutionary histories is a necessity for the successful management of biodiversity (Moritz 1994). These classifications have been based on species, subspecies and more recently on evolutionary significant units (ESUs; and the sub-categories hereof). A number of definitions for ESUs have been proposed (e.g. Ryder 1986; Waples 1991; Crandall et al. 2000). Probably the most rigorous definition of an ESU is that proposed by Vogler and DeSalle (1994), who define it as “a unit delimited by characters that diagnose clusters of individuals to the exclusion of other such clusters”. Moritz (1994) developed a definition that is dependant on the determination of reciprocal monophyly in mtDNA sequences and significant differentiation of nuclear gene allele frequencies. This definition has been used commonly in conservation genetics although other definitions have been more recently proposed (e.g. Crandall et al. 2000).

The concepts of species and subspecies are, however, problematic and even after much debate there is no consensus as to what constitutes these taxonomic entities. Currently as many as 20 species definitions are recognized, but are based on two main species concepts, namely the biological species concept (Mayr 1963) and the phylogenetic species concept (Cracraft 1983). Mayr (1963), under the biological species concept, defines species as a “group of interbreeding natural populations that are reproductively isolated from other such groups. A weakness of this definition is that the quantity of diagnosable evolutionary units may be underestimated. This has obvious conservation

implications. The phylogenetic species concept, as defined by Cracraft (1997), is defined as “the smallest population or group of populations within which there is a parental pattern of ancestry and descent which is diagnosable by unique combinations of character states.” The chief disadvantage of this species concept is that each individual organism can be defined as the smallest diagnosable population leading to a gross overestimation of the number of taxa (Avice & Wollenburg 1997). It has been argued that these two species concepts are not mutually exclusive. By retaining the sought-after properties of these two species concepts, such as reproductive isolation of the biological species concept and modifying the phylogenetic species concept to highlight the aspects of lineage sorting at macro-evolutionary scales, it has been contended that the conflict between these two species concepts can be resolved.

Subspecies delimitations are used to account for geographic differences among populations within a species (Mayr & Ashlock 1991), and have been defined as “a geographically defined aggregate of local populations which differ taxonomically from other subdivisions of the species” (Mayr 1940), or more recently as “groups of actually or potentially interbreeding populations phylogenetically distinguishable from, but reproductively compatible with, other such groups. The evidence for phylogenetic distinction must come from the concordant distributions of multiple, independent, genetically based traits” (Avice & Ball 1990). Similarly to the species concept there is no consensus definition of a subspecies and the criteria used is often open to subjective interpretation. Often then number of subspecies recognised is dependent on whether the investigators are “lumpers” or “splitters” (Simpson 1945). Furthermore the criteria used to designate subspecies have changed with time and in some instances the geographic distribution of a taxon is used as the only criteria to designate subspecies. Moreover, there is often no statistical assessment of morphological differences to test the validity of proposed subspecies (Ryder *et al.* 1988). Some systematists disregard subspecies because in a number of cases fully diagnosable taxa (“species”) have been regarded as subspecies, or because subspecies had been used to distinguish points on a cline (e.g see Cracraft 1983). In these cases proponents of this believe subspecies should be elevated to species

status or subspecies status should be taken away.” It has also been suggested that a taxon must be ranked as a full species if its members could always be distinguished, but as a subspecies if most, but not all, could be distinguished. Subsequently a subspecies has been defined as an aggregation of phenotypically similar populations of a species inhabiting a geographic subdivision of the range of the species in which at least 75% of individuals differ from all those in other populations of a species (Mayr, 1963). The important issue is that subspecies are a geographic portion of a species, not morphs co-occurring with other variants, and that they differ from each other on average, not absolutely (Groves 2004).

Irrespective of the validity of the subspecies definition used, a formally described category is recognized and is thus an important unit of conservation. These identified units have the potential to be indicative of incipient speciation and prospective ecologically relevant adaptations could be acquired during isolation (O’ Brien & Mayr 1991). Furthermore, laws regulating trade in wildlife and endangered species programs, for example the Convention on International Trade in Endangered Species (CITES), identify and protect taxa based on these species and subspecies designations (Geist 1992; Haig 1998). In addition, management programs are also generally based on the taxonomic assignment of the group under investigation, whether this is at the subspecies, species or higher taxonomic level.

The subspecies classification of *M. capensis* is ambiguous with up to 15 taxa having been described. Based on skull morphology and mantle colour, two subspecies assemblages have been recognized (Baryshnikov 2000). The first group (“*capensis*”) includes all the African subspecies as well as *M. c. wilsoni* and *M. c. pumillio* from Asia, and the second group comprising the remaining Asiatic forms (Baryshnikov 1988; 2000). A total of 10 subspecies have been described in Africa and these were based primarily on size, pelage and mantle variation (Coetzee 1971, Rosevear 1974, Baryshnikov 1988; 2000). Unfortunately traditionally recognized subspecies have often been shown not to correspond to evolutionary units, because they are often based on a limited number of

specimens, restricted geographic sampling, and morphological characters used can display extensive individual variation influenced by environmental factors (Johnson *et al.* 2001). Analogously to other large-bodied, wide-ranging species, honey badgers show great morphological variation (Hallgrimsson & Maiorana 2000), and accordingly most of the type individuals seem to represent individual variants of a highly polymorphic species. Moreover, coat colour appears to be under environmental influences (Hollister 1918; Smithers 1983). In addition, mantle colour, one of the primary distinguishing traits used, darkens with age (Rosevear 1974). It is not surprising then that the validity of many of these subspecies has been questioned (e.g. Coetzee 1971). A recent description of this species (Vanderhaar & Ten Hwang 2003) adopt Baryshnikov's (2000) subspecies descriptions, who only recognizes five subspecies in Africa, although the criteria to group subspecies, or otherwise, is not clear from the latter study.

From a phylogeographic perspective there are no obvious barriers to dispersal between proposed subspecies and the boundaries seemed to be strongly linked to political borders. In addition, the honey badger is highly vagile and the possibility of clinal variation has not been examined. It is therefore, reasonable to suggest that the subspecific designations in *Mellivora* are more than likely unreliable indicators of the true evolutionary relationships among geographic regions and should thus not be used for management decisions (Ball & Avise 1992).

General Biology

Overview

Despite the honey badgers fearsome reputation this species is elusive and is seldom seen in the wild. Perhaps as a consequence of this, not many scientific studies have focused on this species, leaving much of the natural history unclear, except for a single detailed study documenting the social organization of this species (Begg 2001a). To complicate matters, available data from field notes and anecdotal reports are often contradictory.

The honey badger has an extensive range throughout much of sub-Saharan Africa the Middle East and though to Asia (Smithers 1983; Skinner & Smithers 1990; Baryshnikov 2000; Fig. 1). Throughout this range there are inevitably areas where individuals do not occur. For example, their absence has been documented in the Free State Province, in South Africa and from the Malabar Coast, the lower Bengal and Ceylon (Finn 1929; Kingdon 1997). At present, the species is also characterized by a disjunct distribution between Africa and Asia. The reason for the absence of *Mellivora* from these regions are not certain, as the taxon can occupy a wide variety of habitats ranging from the fringes of deserts (the Sahara and pro-Namib) to the rain forests of the Democratic Republic of the Congo and from sea level to over 4000m in the afro-alpine steppes in the Bale Mountains in Ethiopia (Smithers 1983; Sillero-Zubiri 1996). Although unlikely the reported absence from these regions may also be a consequence of the elusive nature and the lack of studies that have focussed on this species



Fig. 1 Geographic distribution of the honey badger, *M. capensis* (Map redrawn from Finn 1929; Smithers 1983; Skinner & Smithers 1990; Kingdon 1997; Baryshnikov 2000).

Honey badgers are solitary foragers with primarily a carnivorous diet (Begg 2001a; Begg *et al.* 2003) but fruits, bulbs and tubers have also been reported in its diet (Fitzsimons 1919; Dragesco-Joffe 1993; Begg *et al.* 2003). Honey badgers are both opportunistic and generalist in feeding behavior with seasonal shifts in diet which seemingly reflects changes in the availability of primary prey items (Begg *et al.* 2003). The species is particularly well known for breaking into beehives to consume the honeycomb and bee larvae (Begg 2001b) but it has been suggested that there are strong geographic differences in diet (Stuart 1981; Kruuk & Mills 1983; Kingdon 1989; Skinner & Smithers 1990). For example, in stony territory or marshland where tortoises may be locally abundant these may be the preferred foods, whereas in cattle and elephant dominated regions insect larvae may be favored (Kingdon 1977). On the other hand there is a lack comparative data for other habitats and this may simply reflect its generalist feeding behavior. In areas undisturbed by man the honey badger is regularly active during the

day with increased activity during sunrise and sunset (Begg *et al.* 2003). It has been suggested that the foraging behavior of the honey badger, in areas occupied by humans, has shifted to nocturnal due to human activities (Skinner & Smithers 1990). This is likely to increase foraging costs and may place an additional stress on the survival of the species, especially in habitats with high human activity (Begg 2001a).

In a study in the southern Kalahari, large home ranges have been documented for *Mellivora* with an average size of $541 \pm 93 \text{ km}^2$ for males and $126 \pm 13 \text{ km}^2$ for females (Begg *et al.* 2005a). Females exhibit a loose territorial system, while males, contrary to the typical mustelid pattern of intrasexual territoriality, have a system of overlapping home ranges encompassing the home ranges of many females (up to 13 females; Begg *et al.* 2005a). The spatial organization probably reflects a polygamous mating system (Verwey *et al.* 2004). In solitary carnivores the female spatial patterns are generally determined by the abundance and dispersal of food, while that of males, at least during the mating season, is predominantly determined by the distribution of females (Lindstedt *et al.* 1986; Sandell 1989; Johnson *et al.* 2000). As activity patterns of species have been shown to change between habitats and home range sizes it is unknown whether the home ranges of honey badgers will be smaller in more mesic environments. The significantly larger home range of male honey badgers, compared to females, exceeds the values predicted for intersexual differences accounted for size dimorphism alone (Begg 2001a). This suggests factors other than energetics affect male home range size (Sandell 1989). It has been suggested that males are nomadic (Kruuk 1995; Begg 2001a). This difference in home range size coupled with the adoption of different tactics, i.e. roaming vs. staying, is expected to affect the levels of gene flow of males and females, necessitating the need to investigate the genetic structure employing both maternally (mitochondrial marker) and bi-parentally (nuclear markers) inherited markers in the present study.

Small litter sizes, usually one but up to four cubs (Neal & Cheeseman 1996; Johnson *et al.* 2000) have been reported for this species although in captivity litter sizes are commonly one with infrequent cases of two cubs, with no records of more than two cubs

(Johnstone-Scott 1975). In the most comprehensive study of the reproductive behavior of this species only a single cub was observed per litter out of 18 documented cases, with no evidence of more than one cub being born (Begg *et al.* 2005b). It has been proposed that individuals with low food availability have smaller litters (Boutin 1990) and it remains to be seen whether more than one cub will be produced in more productive environments. Cubs have an extended period of dependence (12-16 months) on their mothers, with the males playing no part in parental care (Begg *et al.* 2005b).

Conservation

The conservation status of the honey badger is ambiguous and it is apparent that the species is now absent in many areas where it previously occurred, e.g. Israel (Ben-David 1990) and parts of Morocco (Begg 2001a). Evidence also exists that population fragmentation may occur throughout its range (Smithers 1983; Comrie-Grieg 1985; Cuzin 1996) and in South Africa, where the distribution of the honey badger has been relatively well documented, reports exist of declining numbers/absence from certain areas (Coetzee 1977; Smithers 1986; Skinner & Smithers 1990; Rowe-Rowe 1992). At present the honey badger is not listed on the international red data book (IUCN 1999), but it appears on appendix III of CITES (Ghana & Botswana; Rowe-Rowe 1992). The South African Red Data Book (2004) lists the honey badger as near threatened, i.e. if the causal factors for its decline continue the honey badger may become threatened (Friedmann & Daly 2004).

The decline in numbers is due to several anthropogenic factors and includes factors such as agriculture and the associated destruction of suitable habitat. Direct conflict between man and the honey badger has also arisen because of the destruction of beehives (Kingdon 1977; Hepburn & Radloff 1998; Begg 2001b) leading to their rigorous extermination in some areas (Kingdon 1977). Small livestock farmers also kill the species (Kingdon 1989), and parts of their bodies are utilized for traditional medicine and human consumption (Cunningham & Zondi 1991; Monadjem 1998; Begg 2001a). In addition, honey badgers are often unintentionally exterminated by non-selective methods such as gin traps and poisons (Stuart 1990; Begg 2001c). Given that the species is already rare throughout its range and that they need fairly large home ranges, sharp declines in their numbers are to be expected. In particular, the destruction of suitable habitat coupled to factors such as the small litter size, extended period of cub dependence, increased maternal investment and a relatively short life span makes the honey badger vulnerable to local extinctions.

With local extinction's already being documented (Coetzee 1977; Smithers 1986; Rowe-Rowe 1992), translocations are being investigated as a management tool in order to re-establish populations across Africa (Begg, pers. comm.) Due to their large home range requirements and territorial behaviour, few nature reserves can maintain viable populations on their reserves. It is quite likely that their re-establishment on suitable habitat, not exclusive to nature reserves, will be required for long-term conservation of this species. At present no data is available and if this taxa goes extinct in certain areas there is no other alternative. In addition game farms are looking to reintroduce this charismatic species, and when source material for reintroductions cannot be found nearby there are no management plans to assist decision-making. The conflict between badgers and honey-farmers is intense and already caused a dramatic decline in their numbers in the Western Cape Province. Certainly conservation efforts should be directed towards conserving the habitat/ecosystem but in some instances where they have already experienced local extinctions this is not possible. Phylogeographic information is lacking for this taxon and subspecies designations are clearly unreliable (see above). Based on previous phylogeographic studies on highly mobile carnivores in Africa it is possible that the honey badger populations might show some level of genetic differentiation (Kingdon 1997; Girman *et al.* 2001).

Dr. Colleen Begg, Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, South Africa

Potential factors influencing gene flow among *Mellivora capensis* populations

Previous studies of mobile vertebrates with large home ranges have identified potential geographic barriers to gene flow in Africa. At the broader scale the distribution of the Miombo (*Brachystegia-Julbernardia*) woodland, that exhibited cycles of expansion and contraction throughout the Pleistocene (Hamilton 1976) has been suggested to influence the distribution of mammals (Coe & Skinner 1993). Carnivore examples can be found in taxa such as the black-backed jackal (*Canis mesomelas*) and bat-eared foxes (*Otocyon megalotis*; Kingdon 1997). On the contrary, an example where the expansion of the miombo woodlands did not act as a long-term effective barrier to gene flow can be found in the wild dog (*Lycaon pictus*; Girman *et al.* 2001). This latter species, like the honey badger, is highly mobile and is able to live in a variety of habitat types. Honey badgers also occur throughout the Miombo woodland and it is thus not expected that the expansion of the miombo forests acted as an effective barrier to gene flow in this species.

Ecological speciation has also recently been documented for the African elephant (Grubb *et al.* 2000; Roca *et al.* 2001; Comstock *et al.* 2002; Roca *et al.* 2005). This species was commonly thought to consist of a single taxon with two distinct morphological forms (savanna and forest elephants). There is now strong genetic evidence for species-level distinction between these two morphs (*Loxodonta africana* and *L. cyclotis* respectively). It has been argued that this speciation was as a result of a bottleneck or founder event in savanna elephants and this has been maintained through genetic isolation due to habitat preference. Although not expected it is possible that some honey badger populations might be more habitat specific. For example, the forest dwelling honey badgers might be confined to this habitat. Support for isolation of this taxon can be found in the morphological differences (completely black individuals have been documented, particularly in the Ituri forest region of the Democratic Republic of the Congo; Nowak 1999). If this holds, these populations might experience low levels of gene flow due to habitat preference.

The Great Rift Valley has also been proposed as a barrier to gene flow for a variety of vertebrates. Examples can be found in reptiles such as chameleons (*Rhampholeon*; Matthee *et al.* 2004), bird species such as the ostrich (*Struthio camelus*; Freitag & Robinson 1993), and mammals, such as the wild dog (Girman *et al.* 2001), springhare (*Pedetes capensis*; Matthee & Robinson 1997), lion (*Panthera leo*; Dubach *et al.* 2005), sable (*Hippotragus niger*; Pitra *et al.* 2002) and in the wildebeest (*Connochaetes taurinus*; Arctander *et al.* 1999). On the other hand, the Rift Valley did not appear to act as a barrier to gene flow in buffalo (*Syncerus caffer*; Simonsen *et al.* 1998), which occupies a large range of habitat types and is capable of moving large distances. Whether the Great Rift Valley system, coupled with climatic fluctuations and associated habitat changes, has served as a geographic barrier to gene flow in the honey badger, is uncertain but probably unlikely. Its large dispersal ability might enable it to traverse large tracts of inhospitable land thereby maintaining gene flow between geographically distinct regions.

Apart from the barriers specific to the African continent, the honey badger also displays a disjunct distribution between Africa and Asia (see Fig. 1), implying limited or no gene flow between the African and Asian populations. Historic levels of gene flow could have occurred via a land bridge between the Sinai and the Egypto-somalian domain, which was closed by the Red Sea and the Gulf of Suez in the late Pleistocene. Gene flow could have also taken place via land bridge connections between Africa and the Arabian Peninsula, up to 11 000 years before present (Delany 1989; Robinson & Matthee 1999). Infrequent exchange between Arabia and Africa may have also taken place, for example in the Egyptian-Sinai-Israel domain, when the Gulf of Suez emerged above sea level during the Pliocene establishing a land bridge (Tchernov 1989; Robinson & Matthee 1999)

Although the taxonomy within this species on the whole is not conclusive, one investigation that systematically assesses regional morphological variation within this species identified two lineages (Baryshnikov 2000) distinguishing between the African and Arabian subspecies up to Iran and the rest of Asia (see below). If this holds, it would

indicate that there has been limited gene flow between these two areas. The Himalayan orogenic fold belt, forming already during the late Miocene (in particular the Pontian and the Zagros), became an arid barrier to free biotic exchange between Africa and Arabia and north of this mountainous chain (Tchernov 1989). This could possibly have acted as a barrier thereby generating the phylogeographic structuring observed.

Choice of genetic markers

Phylogeography

Within most species there exists geographic variation in morphology and genetic constitution (Slatkin 1987). This geographic variation results from a number of forces and the observed phylogeographic structure of a species reflects patterns of historical fragmentation under the processes of mutation, genetic drift and natural selection. The pattern is constrained by gene flow between populations, which in turn is influenced by the dispersal capabilities of a species (Slatkin 1987). In wide ranging species, geographic distance is likely to become an important obstacle to dispersal, because at some geographic scale distance becomes an inhibiting factor to gene flow between regions. These patterns can range from continuous to discontinuous (Avice *et al.* 1987) and fast evolving molecular markers are extremely powerful to make inferences about the phylogeographic structuring of a species.

The mitochondria

One of the greatest advances in the field of molecular population genetics in the last few decades has been the detection and use of mtDNA in population genetic analysis (Avice 1998). Mitochondrial DNA has many characteristics that make it particularly suited as a molecular marker in evolutionary investigations. With the introduction of PCR techniques (Saiki *et al.* 1988) the rapid extraction of information from the mtDNA molecule has greatly improved evolutionary analyses, particularly at the population level where large sample sizes are generally required (Zhang & Hewitt 1996a). Although the mitochondrial genome possesses a number of useful characteristics for evolutionary studies at the population genetic level (such as effective haploidy, maternal inheritance and lack of recombination) there could be several associated problems that may confound its usefulness as a marker for population genetic studies. Things to consider include biparental inheritance (Kondo *et al.* 1990; Gyllensten *et al.* 1991; Kvist *et al.* 2003)

heteroplasmy (Awise 1991; Rand 1993; Hoelzel *et al.* 1994) and the occurrence of nuclear pseudogenes also known as numts (Lopez *et al.* 1994; Sorenson & Quinn 1998).

Animal mitochondrial genomes are comprised of a small, closed circular (generally) DNA molecule of approximately 14 000 - 19 000 base pairs (bp; Cantatore & Saccone 1987; Fairbanks & Andersen 1999). It is enclosed in the mitochondrion and contains 13 protein-coding genes (necessary for aerobic respiration), 22 tRNA genes, two rRNA genes and a noncoding segment of approximately 300-1000 bp long. Typically a somatic cell will contain 500-1000 mitochondria, derived from a limited number of mitochondria in the primordial germ cells, and each mitochondrion should thus contain multiple identical copies of the same mitochondrial genome.

Complete sequencing of the mitochondrial genome has revealed a highly compact arrangement of coding DNA, with only a single major segment of noncoding DNA. This non-coding region, also known as the control region, contains the origin of replication for both the light and heavy strands and their associated promoters. In mammals this stretch of noncoding DNA, is of variable length between and even within species and is found between the tRNA_{Pro} and tRNA_{Phe} genes. Because mtDNA replicates and transcribes within the mitochondrion, it is in close proximity to ROS (reactive oxygen species), produced by oxidative phosphorylation. This, combined with a low efficiency of mtDNA repair systems (mammalian mtDNA is replicated by γ -polymerase which lacks the ability to edit newly synthesized products and is more likely to incorrectly incorporate nucleotides than its nuclear counterpart, α -polymerase; Ciarrocchi *et al.* 1979) and the increased generation time of mtDNA has the effect that the overall mutation rate of the mitochondrial genome is greater than that found in nuclear DNA (Brown *et al.* 1979). Not surprisingly, the most rapidly evolving region of mammalian mitochondrial DNA appears to be the segment of the control region that contains the D-loop (Upholt & Dawid 1977; Walberg & Clayton 1981; Chang & Clayton 1985).

Phylogeographic inferences based on genealogies

Haplotypes detected in a species (for example, those based on the mtDNA control region sequences) represent a gene lineage that has survived through an organismal pedigree, and each gene phylogeny can differ from locus to locus (Ball *et al.* 1990). Because gene trees may differ from the species (or population) tree to which it belongs, reconstruction of the species (or population) tree from a single gene may lead to erroneous results (Pamilo & Nei 1988; Doyle 1992; Degnan 1993). This can be better understood if one disregards a simple tree with discrete thin branches (a bifurcating phylogeny) and rather see the genealogy as a cloud of gene histories which is more like a statistical distribution, with a central tendency but also a variance because of the multiplicity of gene trees (Maddison 1997). If population sizes have been small relative to the length of the phylogenetic branches, such as in the case of higher-level phylogenies, then a gene tree might be a fair representation of the species tree (Maddison 1997). However, the shallower the phylogenetic level being investigated, the more important it becomes to follow a multifurcating coalescence approach and to also combine a number of unlinked molecular markers (Pamilo & Nei 1988; Doyle 1992).

Recognizing the shortcomings of phylogeographic inferences based on a single genealogy, particularly one that exclusively tracks the maternal lineage, the use of multiple independent markers should be used to corroborate phylogenetic hypotheses (Slowinski & Page 1999; Zhang & Hewitt 2003). Furthermore the inclusion of biparentally inherited nuclear markers will be particularly powerful to address questions where sex specific gene flow is suspected/observed.

Nuclear markers

The nuclear genome is many times larger than that of the mtDNA genome. It contains multiple markers that can be effectively used to reconstruct independent genealogies. Nuclear markers such as allozymes and those based on restriction fragment length polymorphisms have been used to evaluate the genetic variation in closely related groups

of organisms (e.g. Gavin *et al.* 1991; Taylor *et al.* 2005). Recently, microsatellites have proven to be the nuclear marker of choice for intraspecific investigations (Zhang & Hewitt 2003).

Microsatellites are tandem repeated DNA sequences, with each motif generally consisting of 2 to 6 base pairs that are regularly dispersed throughout eukaryotic genomes (Tautz 1989; Weber & May 1989; Gyapay *et al.* 1994). Microsatellite loci have been shown to have high mutation rates (e.g. Weber & Wang 1993; FitzSimmons 1998), in the range of 1×10^{-4} and 5×10^{-6} in mammals (Dallas 1992; Edwards *et al.* 1992), and are thus among the most variable types of DNA sequence data (Weber 1990). It is therefore reasonable to suggest that they represent the most polymorphic loci to use within and between populations belonging to the same species (Jarne & Lagoda 1996). The mechanisms leading to the high mutation rates observed in microsatellite loci are poorly understood, although it has been proposed that DNA polymerase slippage during replication is the primary cause of mutation (Toth *et al.* 2000).

The high variability coupled to their co-dominant inheritance, presumable selective neutrality and seemingly simple modes of evolution have made them a targeted nuclear marker to infer the phylogeographic structure of a species (Zhang & Hewitt 2003). This coupled the challenges involved in obtaining nuclear DNA markers for use in population genetic studies (Zhang & Hewitt 2003) have lead to their wide use in fine scale evolutionary studies. Cross-species amplification has also contributed to their wide application because primers developed in a particular species often amplify polymorphic loci in a group of related taxa (e.g. Moore *et al.* 1991; Schlötterer *et al.* 1991). In addition, because this method is PCR based, small amounts of tissue are needed making it possible to use museum and non-destructive sampling methods. Despite these advantages there are drawbacks to the use of these markers. These are due principally to the lack of understanding of their molecular evolution (summarized in Zhang & Hewitt 2003 and references therein). A further drawback of this method is that, unlike DNA sequence data, the relationships between populations are based on the allele frequencies

of the loci in the populations and thus lacks the phylogenetic component contained in DNA sequence data. Genealogical patterns of evolutionary relationships are thus difficult to infer. Advances are being made in overcoming the technical and conceptual problems related to the use of nuclear DNA sequence data (reviewed in Zhang & Hewitt 2003), however, the methods of data generation and analysis are not as established as conventional microsatellite techniques and challenges still remain.

Aim

With the increasing concern about the conservation status of *M. capensis* relevant management practices are required. In formulating these management practices it is necessary to determine the phylogeographic structure of this species.

The main aim of the current study was to use molecular genetic markers to assess the genetic population structure of the honey badger. This was done with particular emphasis on honey badgers in Africa. A variable stretch of the mitochondrial control region and frequency data of five nuclear microsatellites were used to elucidate the phylogeographic structure of this species. The data were used to investigate whether there is any genetic evidence of geographic differentiation present in the species.

From this the following objectives were addressed:

- the information was used to determine preliminary evolutionary units for conservation.

- the data also enabled a preliminary assessment of the degree of concordance between phylogeographic patterns of the honey badger and some of the current subspecies designations.

CHAPTER 2: MATERIALS AND METHODS

Genetic samples

Due to the difficulties in obtaining samples from these elusive carnivores all attempts were made to increase the sample size. Skin biopsies or blood samples were taken from live caught individuals and was supplemented with the opportunistic sampling of carcasses such as road fatalities. These fresh tissues, consisting of 54 individuals, present samples taken over an eight-year period. The majority of these fresh samples were collected by two researchers (Keith and Colleen Begg) over an eight-year period (see Appendix). Tissue samples were stored in either a 20% DMSO- saturated salt solution or 90% ethanol, whereas whole blood samples were frozen at -20°C . The fresh material was mainly drawn from southern, eastern and central Africa (Fig. 2), as well as two individuals from Israel. In order to augment sample sizes skin biopsies of museum pelts or scrapings (from the brain or palate) from skulls were also obtained from an additional 30 animals. The latter specimens were predominantly of east African origin, but also included samples from southern Africa and a single sample from India.

Samples were regionally allocated into populations/groups on the basis of their sampling localities. Due to a lack of studies focusing on the mobility of this carnivore it has not been established if honey badgers will be less mobile in more mesic environments, individuals were grouped into populations based on a conservative approach. For the mitochondrial data, based on this conservative approach, individuals that were sampled more than 400km from the center of a geographic area were not included into any population. The population demarcation is also vaguely supported by ecological data suggesting that home ranges are large (up to 844km^2 ; see Chapter 1). This approach could not be followed employing the microsatellite data because of lower sample numbers, and in this instance samples were simply allocated to the most geographically proximate population with the largest sample size. Following this reasoning, the sampling in this study represent a total of six putative populations, namely Cape (South Africa); Kalahari (South Africa and Botswana); southern Mozambique; northern

Mozambique; northwestern Zambia and Kenya (Fig. 2). For the purposes of this study the boundary between eastern and southern Africa was taken as the western branch of the East African rift formation, namely the Western Rift Valley (Albertine Rift) and the Malawi Rift and associated lakes. This places the populations of northern Mozambique and Kenya in eastern Africa and the remainder of the populations in southern Africa.

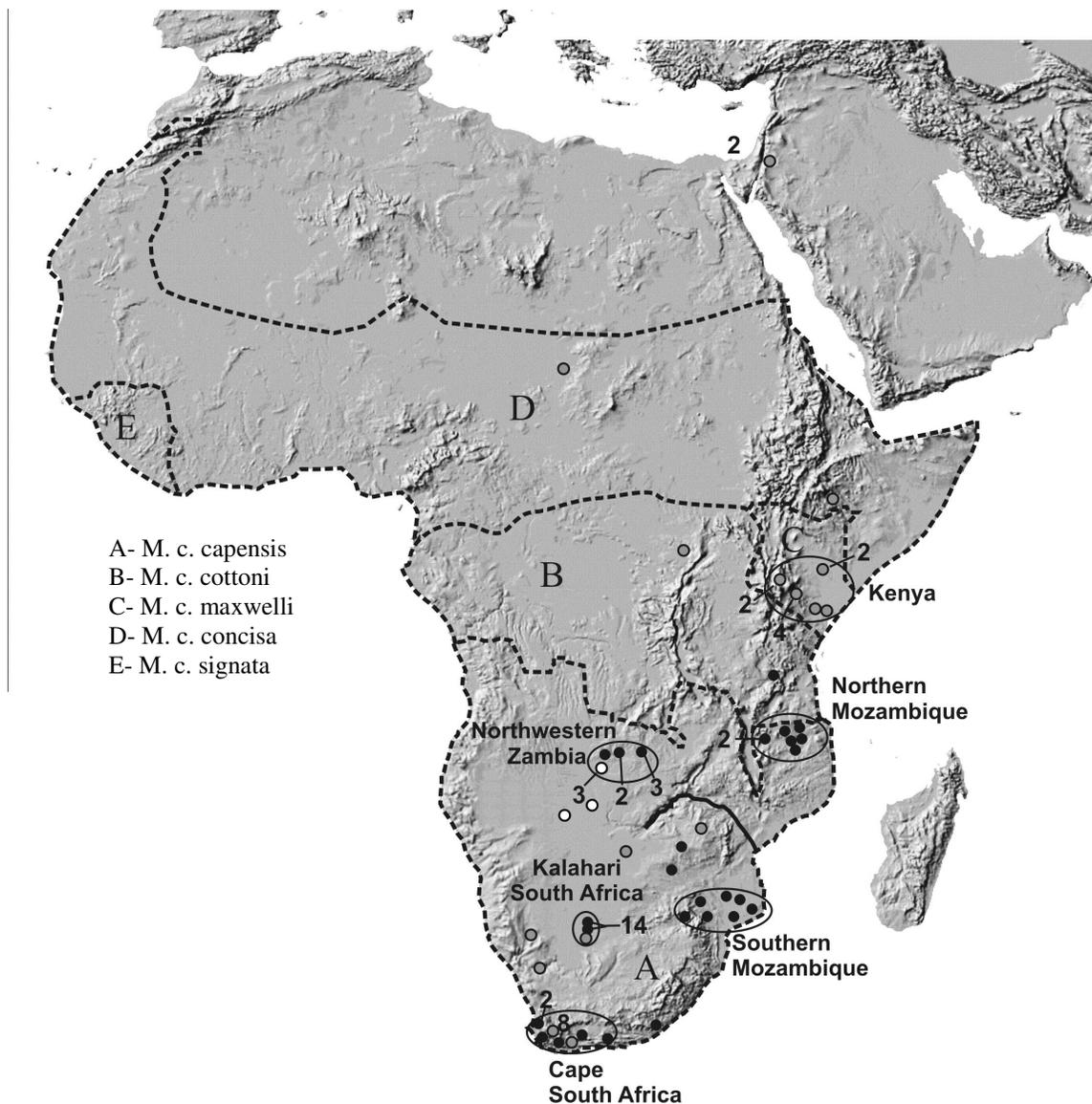


Fig. 2 Map showing sampling localities for the honey badger including the six putative populations. Black circles indicate samples used in both mitochondrial and microsatellite analyses, samples shaded gray indicate those used for mitochondrial analysis only and the samples shaded in white represent individuals used only in the microsatellite analysis. The Indian specimen is not indicated and only included in the mtDNA analyses. Dashed lines indicate subspecies designations in Africa (modified from Vanderlaar & Ten Hwang 2003) and the solid line shows the location of the lower Zambezi River. Subspecies designations are indicated by capital letters (A-E) and population abbreviations used in text are as follows; North-western Zambia: NWZA, Kenya: KENY, Northern Mozambique: NMOZ, Southern Mozambique: SMOZ, Kalahari, South Africa: KALA and Cape, South Africa: CAPE.

DNA extraction

Standard techniques based on proteinase K digestion followed by phenol/chloroform/iso-amyl alcohol procedures were used to extract total genomic DNA from fresh tissue samples (Sambrook *et al.* 1989; Hillis *et al.* 1996). Genomic DNA was extracted from nucleated blood using the Dneasy Tissue Kit (Quiagen- cat# 69504) according to the manufacturers protocol. All museum specimens were extracted in an area of the laboratory reserved for museum samples and were performed under sterile conditions in an extraction hood. Scrapings or skin pieces were washed in 100% EtOH, 70% EtOH and sterile H₂O respectively in an attempt to reduce surface contamination. Total genomic DNA was then extracted using the Dneasy Tissue Kit (Quiagen- cat# 69504) as above but with an extended digestion period allowing tissue to completely digest (up to a period of four days, with the daily addition of 20 µl [1 mg/ml] proteinase K). The elution of DNA from membranes was achieved with pre-warmed (70°C) elution buffer.

Control region sequencing

PCR (Saiki *et al.* 1988) was used to amplify the entire control region of 23 individuals using primers N777 and 12Srev-dloop (Table 1) originally designed for cetaceans and rabbits respectively, but with universal mammalian application. Amplification was performed using standard conditions in reactions containing 2-5 mM MgCl₂. The PCR profile consisted of 35 cycles of 94 °C for 30 s, 53-58 °C for 30 s and 72 °C for one min, preceded by a three minute denaturing step at 94 °C and followed by a final extension step at 72 °C for seven minutes. In instances where amplification was lacking or weak, bovine serum albumin, at a final concentration of 500µg/ml, was included (Kreder 1996). The PCR products were separated on 1% (w/v) agarose gels and purified using the QIAquick gel extraction kit (Quiagen- cat# 28706). Purified PCR products were cycle sequenced using the ABI Big Dye Terminator kits (Applied Biosystems) and the resulting fragments were analyzed on a 3100 ABI automated sequencer. Sequences were manually aligned.

From the sequences obtained for the 23 individuals there was evidence that more than a single fragment was being generated during amplification in the majority of the samples and, accordingly, the source of the sequences needed to be verified. Following identification of the authentic mtDNA copy (see Chapter 5) regions in the sequences were identified to design mtDNA specific primers (where the numt displays sequence differences from the authentic mt sequence). Difference between the numt and authentic mt sequence was most pronounced over a short variable stretch of approximately 250 bp. To determine whether this region of the mitochondrion contained sufficient variable sites to elucidate the phylogeographic structure in this species, an initial investigation was performed to examine the variability of this stretch of the control region. Control region specific primers were then designed and used for amplifying the 250 bp variable region of the control region using the conditions described above.

Table 1 Primers employed in this study with their respective nucleotide sequence and their origin.

Primer name	Primer sequence (5'-3')	Reference
N777	TAC ACT GGT CTT GTA AAC C	Hoelzel <i>et al.</i> 1991
12Srev-dloop	AAT WAW AAG GCC AGG ACC AAA C	C. A. Matthee (unpublished data)
L15162	GCA AGC TTC TAC CAT GAG GAC AAA TAT	Kocher <i>et al.</i> 1989
H15915	GTC ATC TCC GGT TTA CAA GAC	Kocher <i>et al.</i> 1989
Badger dll	TAC CCA TAT TCA TAT ACT RG	This study
Badger dlh2	GAT TGG ATC CAA CAT TAT CAT	This study
Badger dll2	ATC ACG AGC TTA ATC ACC AA	This study
Badger dlh-int	CGC AAG GAT TGA TGG TTT C	This study

Control region data analysis

Phylogenetic analysis

Phylogenetic relationships among haplotypes were determined using neighbour-joining methods (employing maximum likelihood distances) and parsimony in PAUP* version 4.0b10 (Swofford 2002). With model-based phylogenetics it is possible that the choice of

model could effects phylogenetic inferences and phylogenetic methods are less accurate or become inconsistent when the assumed model of evolution is incorrect (Huelsenbeck & Hillis 1993). MODELTEST version 3.0 was used to select the most appropriate model of DNA substitution for the data when calculating the ML distances (Posada & Crandall 1998). MODELTEST results include both the Akaike information criterion (AIC) and the hierarchical likelihood ratio test (hLRT) to calculate the best-fit model of nucleotide substitution. It has been suggested that AIC performs better than hLRTs in selecting the optimal model of substitution (Posada & Buckley 2004) and the model selected by the AIC method was thus used in the present study. The AIC is an estimator of the Kullback-Leibler (K-L) distance (Kullback & Leibler 1951). This distance is representative of the amount of information lost when model b is used to approximate model a (Posada & Buckley 2004). Ranking of candidate models is performed because the larger the AIC difference for a model the less probable that it is the best K-L model.

An estimate of the reliability of the nodes of the tree was obtained by using nonparametric bootstraps, which was first applied to phylogenetic analysis by Felsenstein (1985). This is a statistical technique that entails assembling data sets from the original data matrix by way of randomly sampling characters from the same data set until the replicate data set is the same size as the original. The bootstrap sample is then used to infer a phylogenetic tree and the process is repeated and summarized as the number of times that a particular phylogenetic relationship is observed across the number of replicates. Accordingly 1000 pseudo replicates were performed. To expedite the bootstrap analysis, bootstrap values were calculated on the haplotype data set only.

Chimerism

Chimerism and contamination has been shown to interfere with phylogenetic analyses (Olson & Hassanin 2003). Taking into account the presence of the closely related numt and the possibility that degraded DNA from the majority of the museum specimens had to be amplified in two stretches (one of the overlapping sequences could possibly be

derived from the numt) the following analysis was performed. The sequence data for the two overlapping fragments were analyzed separately (for all individuals irrespective of whether the region was sequenced in two overlapping fragments or not), with the numt sequence included in all analyses. This also serves as a control for possible cross taxa contamination by other samples because if the two separate fragments track different evolutionary trajectories within the same individual, a contaminated source for one of the fragments can be suspected.

Phylogenetic relationships relative to the numt

It has been argued that sequences originated from numts can be used as outgroups because of their presumably slower rate of evolution compared to their mitochondrial complement (Zischler *et al.* 1995). This method can be of particular value in situations where the choice of outgroup is problematic. For example, when outgroup taxa are distantly related to ingroup taxa the likelihood that nucleotide character states shared by a taxon and an outgroup will be based on random similarity rather than common descent, this increases with divergence between outgroup and ingroup taxa (Wheeler 1990). In other words, sequences accumulate so much change that the phylogenetic signal is lost (Maddison *et al.* 1992). The outgroup selected can then have implications on resultant inferences (Milinkovitch & Lyons-Weilert 1998), and in these instances the numt may offer the most suitable outgroup. It is important to recognize that it is only appropriate to employ a numt as an outgroup when certain criteria are met. The numt must have inserted itself into the nuclear genome prior to the divergence of the lineages being examined, must have been part of the same mitochondrial lineage being examined and must be free of recombination with other numts in the nuclear genome. Since there are no assurances that all these conditions have been met, results employing the numt to root the tree must be viewed in light of these constraints. Given there are no closely related extant taxa to the honey badger and because of the unresolved placement of the honey badger within the Mustelidae (Chapter 1), phylogenetic methods, as outlined above were repeated using the numt sequence to root the trees.

Network reconstruction

Phylogenetic trees assume bifurcating patterns that may be invalid for intraspecific comparisons because lines of descent are reticulate (Goldstein *et al.* 2000). Therefore, to depict phylogenetic, geographical and possible ancestor-descendent relationships among haplotypes, a median joining (MJ) network was constructed using the program Network (Bandelt *et al.* 1999). This method begins with minimum spanning trees, combined within a reticulate network. Employing parsimony, median vectors are added, which are the consensus sequences of three mutually close sequences. These median vectors can be interpreted as the extinct ancestral haplotypes or extant unsampled haplotypes. After each stage of median generation the process repeats itself with the enlarged set of sequences until all haplotypes have been connected. This method was performed on the haplotype data including and excluding the numt sequence from analyses (see above).

Population structure

In order to assess the genetic differentiation between the putative populations occurring in southern Africa, a hierarchical analysis of molecular variance was performed (AMOVA; Excoffier *et al.* 1992). To define optimal geographical groupings a spatial AMOVA (SAMOVA) was performed which implements an approach to define groups of populations that are geographically homogeneous and maximally differentiated from each other (Dupanloup *et al.* 2002). This could lead to the possible identification of genetic barriers between groups. The method is based on a simulated annealing procedure that maximizes the proportion of total genetic variance due to differences between groups of populations. These calculations were performed using the program SAMOVA 1.0. (Dupanloup *et al.* 2002) and were restricted to the six putative populations with sufficient sampling (Fig. 2) and two measures of genetic variation was used namely F_{ST} and Φ_{ST} . The Tamura-Nei distance (Tamura & Nei 1993) with a gamma correction estimated in PAUP* version 4.0b10 were used when estimating Φ_{ST} in Arlequin version 2.000 (Schneider *et al.* 2000). As mentioned above, individuals that could not be allocated to any population grouping (further than 400 km from the center of the population group)

were excluded from these analyses. However, analyses based on all samples of African origin (including all unallocated individuals to their closest population) resulted in the same overall patterns (data not shown).

Genetic diversity and isolation by distance

Nucleotide (π) and haplotypic diversity (δ) was estimated for the six populations (Nei 1987). F_{ST} values and Nei's standard genetic distance between populations were calculated. The geographic and genetic distances were used in a 2-way Mantel test (Mantel 1967) using the program MANTEL (Cavalcanti 1988-2001) to determine the relationship between genetic and geographic distance. Geographic distances were taken as the shortest terrestrial geographic distance between sampling sites and F_{ST} and Nei's standard genetic distance measures were used as the genetic distances.

Population Demographic Parameters and Selective Neutrality

The occurrence of a number of different mtDNA haplotypes within a region can result from a number of processes, namely from female gene flow with adjacent regions, or from a founder effect consisting a number of lineages or from mutations that accumulate *in situ* over time. Episodes of expansion, decline and separations leave characteristic signatures in the distribution of nucleotide site differences between pairs of individuals (Rogers & Harpending 1992). An episode of growth generates a wave that moves to the right over time, whereas the histogram of a stationary population is distinguished by a decrease of relative frequency values from 0 to the right, and is either multimodal or free of waves (Rogers & Harpending 1992). The mismatch distributions of haplotypes is expected to generate a histogram that is multimodal or free of waves looking at the first two processes above, whereas a unimodal distribution is expected for the third. To infer historical demographic patterns in the honey badger mismatch distributions between haplotypes were calculated using DNASP (Rozas & Rozas 1999) and Arlequin version 2.000. These were compared with a Poisson distribution indicating population expansions and bottlenecks, selective sweeps on the mtDNA and mutational rate

heterogeneity. The validity of the model was tested based on 1000 replicates, and obtaining the distribution of the test statistic SSD (sum of squared differences) between the observed and estimated mismatch distribution (Excoffier & Schneider 1999). The p-value is calculated as the proportion of simulated cases where the SSD value is larger than the original. A significant SSD value can be interpreted as evidence for departure from the estimated demographic model.

Mismatch distributions were initially performed on the entire mitochondrial data set, and then on the predominantly eastern African haplotypes consisting of the combined clades 3 and 4 and the predominantly southern African clades consisting of the combined clades 1 and 2. Mismatch distribution separate clades identified in this study, containing a minimum of five haplotypes, was also performed separately.

To estimate the time of expansion events two approximations were employed. The first was a relative approach based on the moment estimator of the time to expansion, τ , estimated in Arlequin. The other was based on the mean sequence divergence between clades as a measure of the divergence of lineages. Estimates of the rate of evolution of the control region of various mammal species has been suggested to be in the region of 10% sequence divergence per My (e.g. Aquadro & Greenberg 1982; Vigilant *et al.* 1991; Nachman *et al.* 1994; Stewart & Baker 1994) and this was subsequently used. It is however, important to realize that the estimated times are proportional to the mutation rates used and these are clearly not unequivocal. Furthermore, by using the mismatch distributions parameters to date divergences, it needs to be recognized that there are also large biases and standard errors associated with time estimates with these methods (Harpending *et al.* 1993).

To test for selective neutrality and to investigate mutation-drift equilibrium (under an infinite sites model) a number of tests were used, namely Fu and Li's D^* , Fu and Li's F^* (Fu & Li 1993), Tajima's D test (Tajima 1989) and Fu's F_s test (Fu 1997). All these tests were performed in Arlequin version 2.000 and DNASP and can be used to indicate

whether the observed haplotype frequencies have been produced by the forces of selection or genetic drift. These measures can also be used as evidence of recent demographic changes in populations, such as population expansions or bottlenecks. A range of tests of neutrality were used because the different tests are slightly different in principle and as such can be useful in clarifying different aspects of the evolutionary history of a species (see Fu and Li 1993; Simonsen et al 1995 and Przeworski 2002).

Microsatellite analysis

Microsatellite genotyping

In total five microsatellite loci: Gg10, Gg454, Ma1, Tt1 and Tt4 were used in this study. These were previously isolated and established to be polymorphic in other mustelid species (Davis & Strobeck 1998; Walker *et al.* 2001), and were optimized for amplification in a paternity study of the honey badger (Verwey *et al.* 2004). Data from the Kalahari population, excluding cubs, was obtained from the latter investigation. Amplification and scoring of fragment sizes was performed as described in Verwey *et al.* (2004). Attempts were made to amplify further loci that have been isolated in the Eurasian badger, *Meles meles* (Domingo-Roura *et al.* 2003). In loci Mel 8, Mel 9, Mel 11, Mel 12, Mel 13, Mel 15 and Mel 16 amplification was successful but they appeared monomorphic in a subsample of individuals from different populations and these loci were not included in the study.

Microsatellite data analysis

To detect null alleles, loci were tested for departures from Hardy-Weinberg equilibrium (HWE) using exact tests (Guo & Thompson 1992). Linkage disequilibrium between loci was tested using Fisher's exact tests and the number of alleles per locus was calculated using the program GENEPOP version 3.4 (Raymond & Rousset 1995). Estimates of the average observed and expected heterozygosity were performed by making use of CERVUS (Marshall *et al.* 1998).

Population structure

Similarly to the mitochondrial data SAMOVA was also performed on the microsatellite data. The output of the SAMOVA was used in hierarchical AMOVA's, using F_{ST} , and R_{ST} , to quantify genetic differentiation among populations under a stepwise mutation

model. The null distribution of pairwise values under panmixia was calculated by 1000 permutations.

A number of genetic distances were estimated using measures based on variance in allele frequencies and measures based on variance of repeat number. These included F_{ST} , R_{st} , DAS , DS and $\delta\mu^2$. These distances were computed using the software POPULATIONS 1.1.28 (Langella 2000). The trees were then visualized with TREEVIEW version 1.6.6 (Page 1996). To determine if there is a relationship between geographic and genetic distance, and in order to examine the possibility of isolation by distance, F_{ST} values between populations were included in a Mantel test.

Chapter 3: Results

Mitochondrial analysis

Control region diversity

Using original universal primers a total of 472 bp of homologous control region sequence was obtained for 23 randomly chosen individuals (for which fresh material was available for DNA extraction). This region contained 24 variable sites defining a total of 14 haplotypes. Within this segment was the region earmarked to amplify with primers specific to the mitochondrial copy only (Chapter 5). Pruning this segment to a 230 bp (length of sequence expected to be obtained from the 250 bp fragment) variable stretch earmarked for analyses by utilizing mtDNA specific primers, the number of variable positions slightly decreased to 19 and the number of haplotypes identified decreased to 10. All the additional haplotypes identified by the longer stretch differed by a maximum of an additional four site changes from the haplotypes identified employing the shorter stretch. Due to problematic amplification with museum specimens and to ensure orthologous mtDNA sequences were obtained, the remainder of the study focused on the shorter stretch to elucidate the mtDNA phylogeographic structure among individuals.

Mitochondrial diversity

In total 230bp of the control region of 78 individuals were obtained. The 230 bp stretch contained 33 variable sites, totaling 36 mutations, defining 41 haplotypes. Of these variable sites 29 were parsimony informative. When the numt sequence was included as an outgroup the number of variable sites increased to 46, 30 of which were parsimony informative. Overall haplotype diversity was high (0.970 ± 0.008 - including all specimens; 0.967 ± 0.009 if only samples of African origin were included), and overall nucleotide diversity was (0.0378 ± 0.0030).

Chimerism

Analysis of the mitochondrial data, partitioned according to the two overlapping fragments, revealed that none of the putative mitochondrial sequences were closely associated with the numt sequence. Furthermore the mt fragments, sequenced as two overlapping fragments, tracked broadly the same evolutionary trajectories in the same individuals when analyzed separately (data not shown). This suggests that numt and/or intraspecific contamination had not occurred.

MtDNA Phylogeography

The phylogenetic methods (NJ and parsimony) depict the same basal relationships among sampled haplotypes, and supported the presence of at least four distinct mtDNA clades (Figs. 3 & 4). In the NJ analyses using midpoint rooting, these four main groupings/clades were obtained with bootstrap support higher than 70% with the exception of clade 2 (Fig. 3). These same clades were depicted by the median joining networks indicating that haplotypes belonging to these clades are separated by a minimum of six mutational site changes from each other (Fig. 4). A single haplotype (NM1) is placed in an intermediate position between the southern and eastern African clades. This haplotype is grouped with the southern African haplotypes (clades 1 and 2) with a high bootstrap value and most closely related to clade 2. Haplotype NM1 is thus grouped within the latter clade.

Clades 1 and 2 represent haplotypes found almost exclusively in southern Africa (as defined above), clade 3 comprise haplotypes found exclusively in east Africa (as defined above) and clade 4 consist of haplotypes sampled predominantly in east Africa, Chad, and Israel. In broad terms, these four clades can be subdivided into southern (clade 1 and 2) and eastern/north Africa/Middle East clades (clades 3 and 4). However, two individuals sampled in southern Africa (ZIM3 and RSA2) cluster within clade 4, and three individuals sampled in east Africa (NM7, KEN4 and KEN5) cluster within clade 1,

while a single individual from northern Mozambique (NM1) clusters with haplotypes belonging to clade 2. These samples within these two broad geographic areas whose haplotypes do not cluster within their expected geographic clades are possibly introgressed. Clade 2 contains a sub network consisting of three haplotypes originated from four individuals (three from the NWZA population and an individual from the Ituri forest in central Africa). The haplotypes of the two southern African clades 1 and 2 are separated from one another by at least six mutational steps. The two southern African clades (1 and 2) are separated from the eastern African clade 3 by at least 17 mutational steps and from the eastern African clade 4 by a minimum of nine mutational steps. The haplotypes of the two eastern African clades (clades 3 and 4) are separated from one another by at least 10 mutational steps. Rooting the parsimony tree with the numt sequence reveals some support for the groupings identified above, and indicates that the lineages belonging to clade 4 (including the Indian sample) are the most basal (Fig. 5).

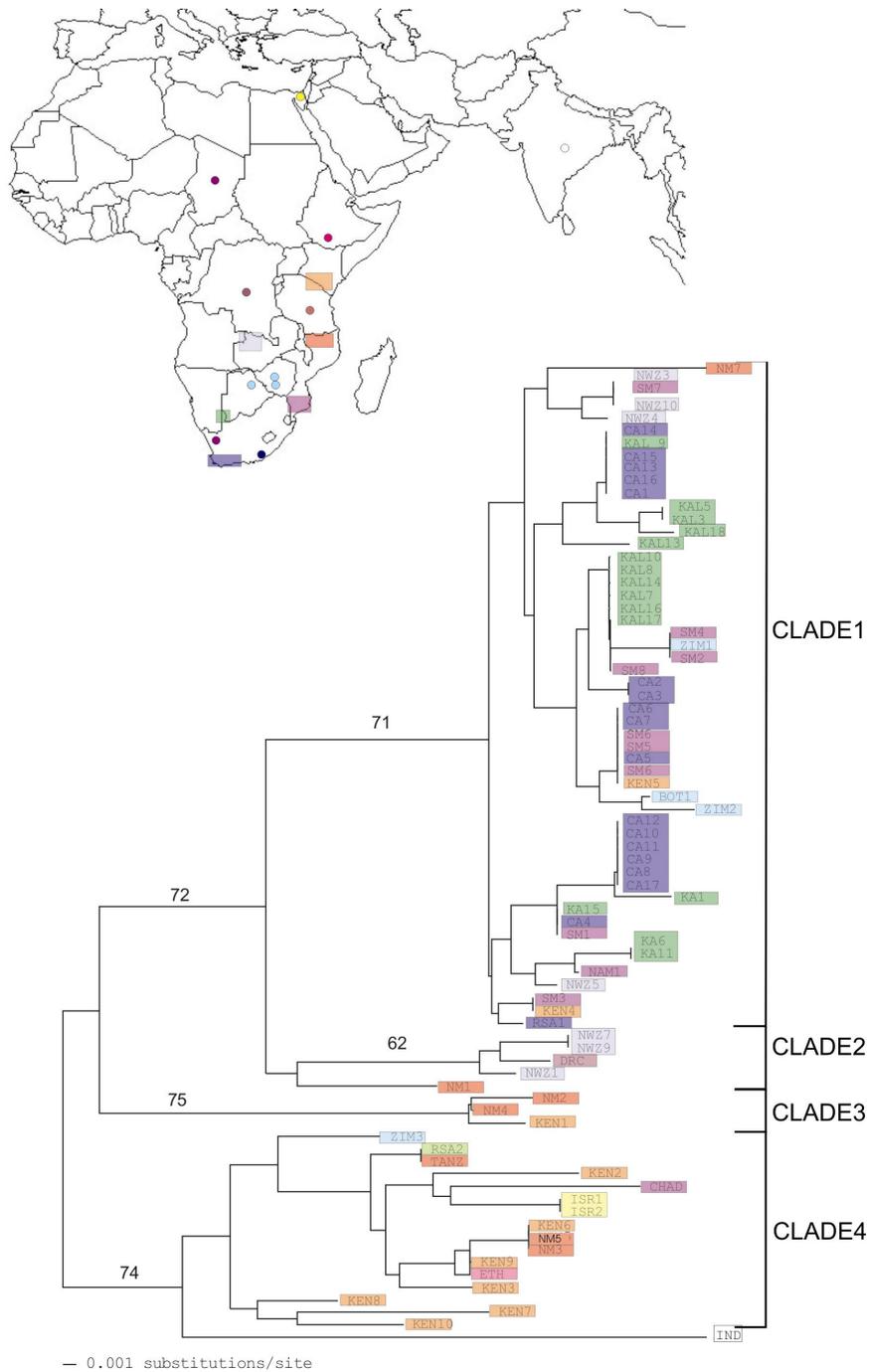


Fig. 3 Neighbor joining phylogenetic relationships between 41 haplotypes identified in the present study. The tree is based on HKY85 distances, with an assumed proportion of invariant sites (0.647) and a gamma rate for variant sites (0.223). Bootstrap values above 60 are shown. Colours are representative of the sampling locality and correspond to the inset map

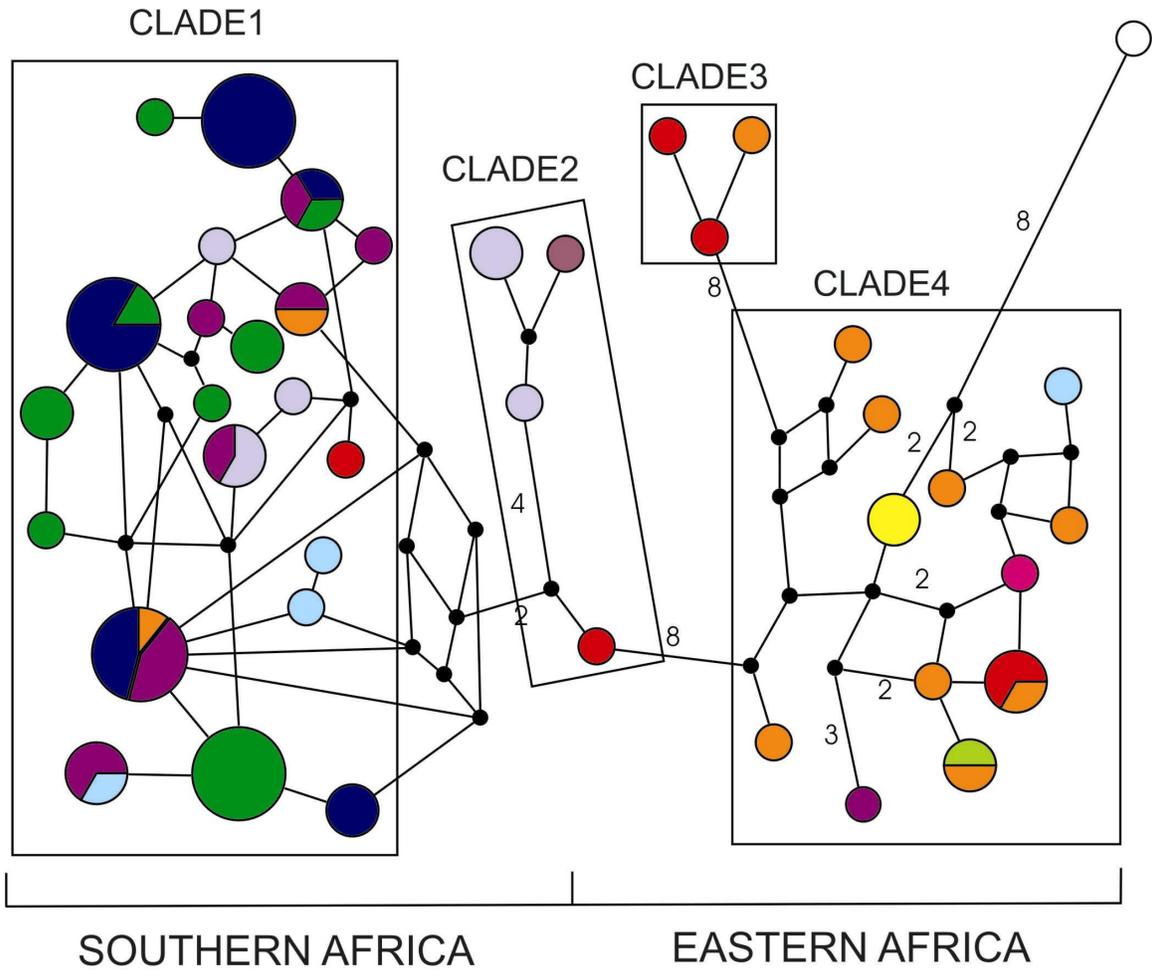


Fig. 4 Haplotype network of the 41 control region haplotypes identified in this study. Small black circles are unsampled or inferred haplotypes. Unless otherwise indicated the line connecting haplotypes represents one mutational change. Colours are representative of the sampling locality and correspond to Fig. 3.

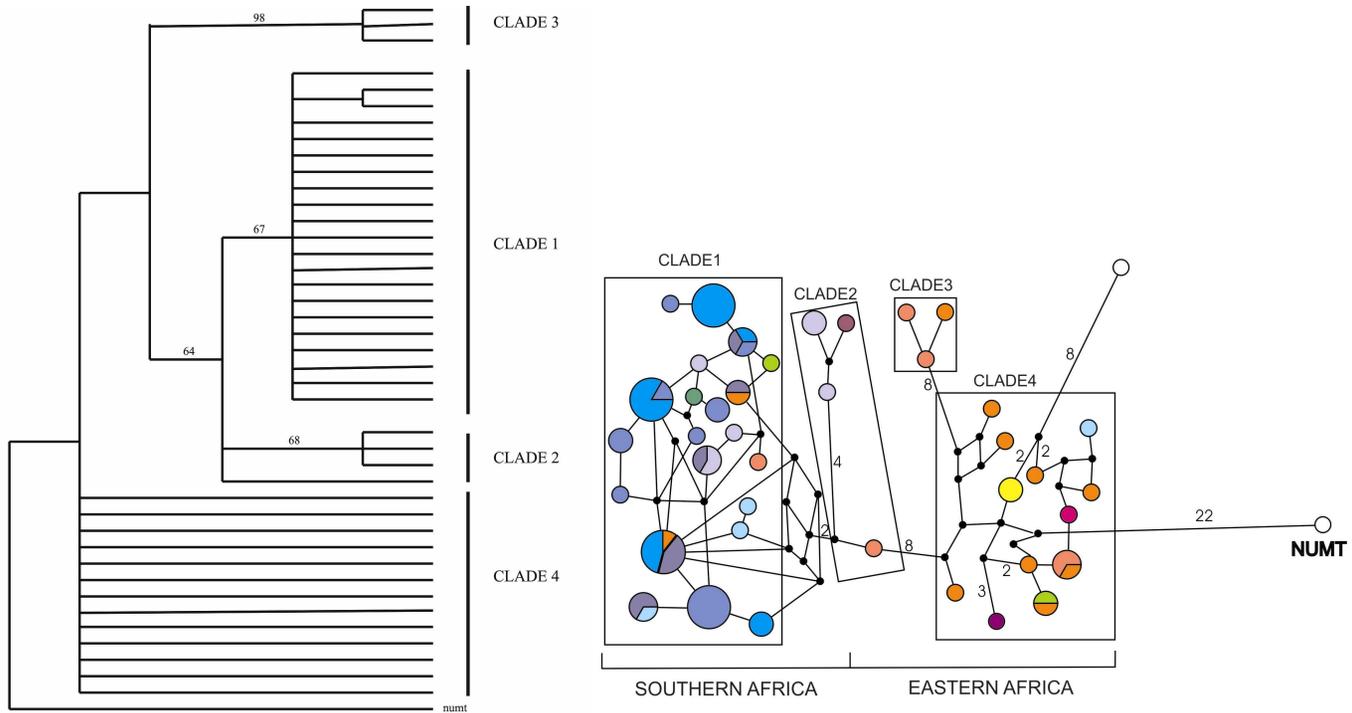


Fig. 5 a.) Phylogenetic relationships of haplotypes using parsimony methods with the numt sequence to root the tree. Clades identified in previous analyses are indicated. Bootstrap values above 60 are shown. b.) Haplotype network, as above, but including the numt in the network reconstruction to indicate the position of the numt relative to the other haplotypes.

Population structure

SAMOVA/ AMOVA analyses indicated that the highest genetic differentiation was identified when two groups were specified containing the two eastern African populations in the first group and the four southern African populations, in the second (Table 2). Although the F_{CT} and Φ_{CT} were not statistically significant, the significance level of these values are approaching being significant and if introgressed haplotypes are removed from analysis these values became statistically significant (data not shown). The three-group scenario, that separates the NWZA population from the remaining southern African populations, and four-group scenario, separating the two eastern African populations, only accounts for slightly less variation between groups, with the Φ_{CT} values for both scenarios and the three group F_{CT} values being significant. On the other hand, in the five-group scenario, in which the SMOZ population is separated from the remaining two southern African populations of CAPE and KALA, considerably less variation was described with both the F_{CT} or Φ_{CT} values being non-significant.

Table 2 Statistics generated employing groupings identified by SAMOVA. F_{ST} and Φ_{ST} values generated from mitochondrial data. Statistically significant results (generated from 1000 random permutation tests) are shown in bold.

	F_{ST} (p-value)	F_{CT} (p-value)	Φ_{ST} (p-value)	Φ_{CT} (p-value)
2 Group				
KENY & NMOZ	0.52	0.46	0.55	0.50
SMOZ, NWZA, KALA & CAPE	(0.000 ± 0.000)	(0.062 ± 0.000)	(0.000 ± 0.000)	(0.068 ± 0.002)
3 Group				
KENY & NMOZ	0.48	0.45	0.50	0.48
NWZA	(0.000 ± 0.000)	(0.015 ± 0.002)	(0.000 ± 0.000)	(0.015 ± 0.001)
SMOZ, KALA & CAPE				
4 Group				
KENY	0.46	0.46	0.49	0.49
NMOZ	(0.000 ± 0.000)	(0.051 ± 0.003)	(0.000 ± 0.000)	(0.047 ± 0.002)
NWZA				
SMOZ, KALA & CAPE				
5 Group				
KENY				
NMOZ	0.41	0.41	0.44	0.44
NWZA	(0.000 ± 0.000)	(0.077 ± 0.000)	(0.000 ± 0.000)	(0.066 ± 0.002)
SMOZ				
KALA & CAPE				

Mantel tests indicate significant isolation by distance among these sampled populations employing F_{ST} and Nei's standard genetic distance measures ($R = 0.83$; $P < 0.01$ for correlations involving both distance measures; Fig. 6). To test whether this relationship was heavily influenced by the two more distantly sampled eastern African populations, regression analyses were performed again using Nei's standard genetic distance but only on a subsample of the data (excluding the eastern African populations). The relationship remained significant for the subsample of populations consisting of only the southern African populations ($R = 0.61$; $P < 0.05$). Analysis was not performed within the eastern African populations due to the small number of sampling localities. When single samples from Chad, Israel and India are included the relationship remains significant ($R = 0.82$; $P < 0.01$).

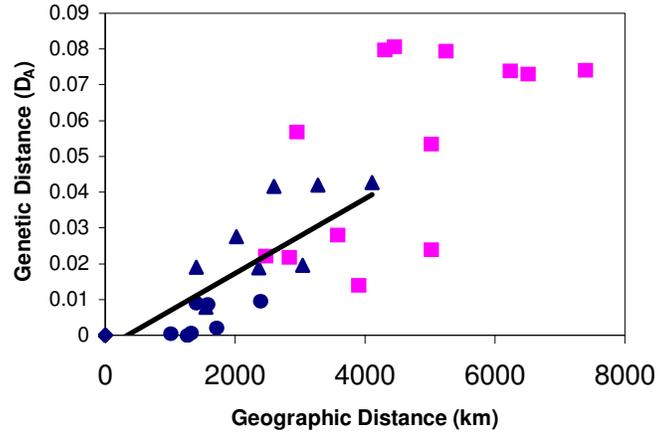


Fig. 6 Relationship between genetic distance and geographic distance. D_A based on sampled populations incorporating mitochondrial data. Triangles represent comparisons between main geographical regions, namely east and southern Africa, while circles represent comparisons within main geographical regions. Squares represent comparisons between the populations and the single haplotypes found in Chad, Israel and India. Solid line represents the fitted regression line calculated for populations only.

Pairwise comparisons between populations, using F_{ST} and Φ_{ST} , display concordant patterns (Table 3) with the majority of comparisons (11 out of 15 for both distance measures) being significant. Non-significant values are found between the southern African population of KALA and the two southern African populations of CAPE and SMOZ and between the NMOZ population and the populations of KENY and NWZA. The largest pairwise differences were found between the three southern African populations of KALA, SMOZ and CAPE, and the remaining three populations (NWZA, NMOZ and KENY) and between the KENY population and the NWZA population.

Table 3 F_{ST} and Φ_{ST} values between the honey badger populations used in this study. Figures above the diagonal represent F_{ST} values between populations and those below the diagonal represent Φ_{ST} values between populations based on mitochondrial data. Abbreviations as in Fig.2.

	CAPE	KALA	SMOZ	NWZA	NMOZ	KENY
CAPE		0.044	0.146	0.370	0.414	0.589
KALA	0.041		0.041	0.323	0.375	0.562
SMOZ	0.147	0.046		0.309	0.333	0.533
NWZA	0.383	0.338	0.322		0.130	0.372
NMOZ	0.430	0.393	0.341	0.116		0.021
KENY	0.613	0.587	0.553	0.382	0.014	

Population history

Mismatch distribution on the entire mitochondrial data set indicate a bimodal distribution (Fig. 7), although the observed data does not differ significantly from the expected distribution under a sudden expansion model as indicated by the raggedness statistic and the non-significant SSD statistics indicated no significant departure from the model of population expansion (Table 4). The peak on the left hand side of the graph is due largely to the pairwise differences within clades while the peak on the right is due predominantly to pairwise comparisons between clades. The geographical structuring of these clades, where clades 1 and 2 occur predominantly in southern Africa and clades 3 and 4 predominantly in eastern Africa and the mismatch distributions and neutrality tests suggest independent demographic histories for each clade.

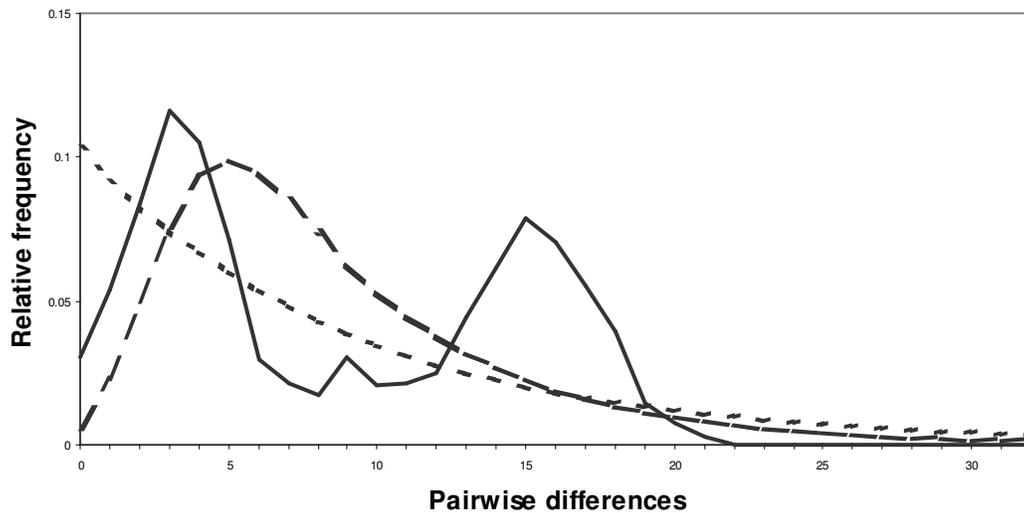


Fig. 7 Mismatch nucleotide distributions for the 41 mitochondrial DNA haplotypes. The solid line represents the observed distribution; the broken line the expected distribution under a model of growth/decline and the dotted line represents the expected distribution under a model of stasis.

Table 4 Results of tests for selective neutrality (Fu's F_s , Fu and Li's D^* , Fu and Li's F^* and Tajima's D) and the estimated value for the population parameter τ . (τ = the time passed since the expansion). SSD: statistic testing for significant departure of observed data from that expected from a sudden expansion model.

	τ	Fu's F_s (p-value)	Fu and Li's D^* (p-value)	Fu and Li's F^* (p-value)	SSD (p-value)	Raggedness Index (p-value)	Tajima's D (p-value)
Southern African haplotypes (clades 1 and 2)	2.187	-10.44 (<0.01)	0.88 (>0.10)	0.63 (>0.10)	0.009 (0.192)	0.017 (0.522)	-0.602 (>0.10)
Eastern African haplotypes (clades 3 and 4)	3.840	-4.77 (<0.01)	0.30 (>0.10)	0.36 (>0.10)	0.016 (0.606)	0.032 (0.786)	-0.238 (>0.10)
Southern and eastern African haplotypes	--	-14.57 (<0.01)	0.64 (>0.10)	0.75 (>0.10)	0.015 (0.359)	0.0136 (0.403)	0.635 (>0.10)
Clade 1	3.125	-10.76 (<0.01)	0.16 (>0.10)	0.06 (>0.10)	0.006 (0.110)	0.030 (0.350)	-0.192 (>0.10)
Clade 4	4.600	-4.74 (<0.01)	-0.38 (>0.10)	-0.36 (>0.10)	0.008 (0.560)	0.024 (0.610)	-0.132 (>0.10)

Population expansions for the clades representing each of the major geographic areas, as suggested by the mismatch distributions (Fig. 8 a & b), were supported by the raggedness statistic and the non-significant SSD statistics (Table 4). A negative and significant Fu's F_s provided more evidence for a demographic expansion for each of the combined southern and eastern populations respectively (Fu 1997). Fu and Li's D^* and F^* and Tajima's D were not significant, excluding the prospect that background selection is accountable for the digression from neutrality. This is confirmed by mismatch distributions on individual clades with sufficient samples, namely clades 1 and 4, that show similar patterns, indicating demographic expansions (Table 4; Fig. 8 c & d).

The age of an expansion affects the pattern of diversity considerably more for lower rather than for large Nm values. For low Nm values Tajima's D statistic is less powerful whereas Fu's F_s statistic is more powerful at rejecting the hypothesis of selective neutrality and population equilibrium for relatively recent expansions (Ray et al. 2003). The large negative and significant values of Fu's F_s statistic together with the small

negative and positive non-significant Tajima's D statistic indicate that the population expansion is relatively recent with low Nm values.

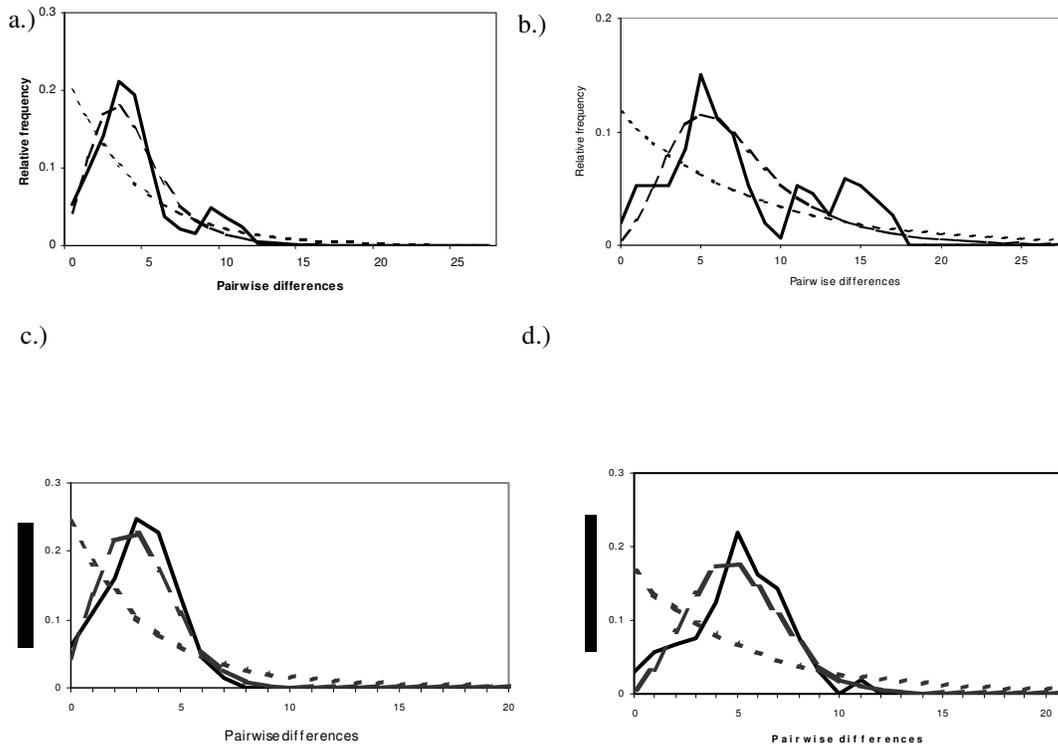


Fig. 8 Mismatch nucleotide distributions for: a.) clades 1 & 2; b.) clades 3 & 4; c.) clade 1 and d.) clade 4. The solid line represents the observed distribution, the broken line the expected distribution under a model of growth/decline and the dotted line represents the expected distribution under a model of stasis.

The τ values calculated indicate that the eastern African population is the older population/clade when compared to the southern African population. This notion is also supported by the higher level of mtDNA heterozygosity and haplotypic frequency in this region (KENY: $\delta = 1.000$, $SD = 0.034$; $\pi = 0.0487$, $SD = 0.0067$; NMOZ: $\delta = 0.952$, $SE = 0.096$; $\pi = 0.0749$, $SE = 0.0070$; Table 5). Further evidence for this was also provided by their basal position when the numt mtDNA sequence was used to root the tree. The rough estimated time of divergence between southern and eastern clades is approximately 521 700 - 782 800 years ago (assuming 10% sequence divergence per My ;Aquadro & Greenberg 1982).

Microsatellite analysis**Microsatellite variation**

The Kenyan population was constituted almost exclusively of museum samples and the microsatellites could not be reliably amplified for most loci. Microsatellite loci TT1 and Gg454 were successfully amplified in a number of the museum samples consisting the Kenyan population (data not shown). These two loci have the shortest absolute length of the microsatellite loci used (with some overlap in range). This consistent lack of amplification of the longer loci was likely due to the degraded nature of the DNA often found in museum samples. To gain preliminary insights into the nuclear DNA differentiation between populations in southern and eastern Africa all populations containing a minimum of seven individuals were included in the analyses (Fig. 2, Table 5). A number of departures from HWE were found: Gg10 in the CAPE, SMOZ and NMOZ populations and Gg454 in the CAPE and KALA populations. All departures were due to an excess of homozygotes, implying the potential presence of null alleles in these samples. As deviations from HWE were not found in other populations for these loci, they were retained for further analyses. Potential genotypic disequilibria were suggested in the two pairs of loci in two populations, namely Tt4 and Ma1 for the CAPE population and TT1 and TT4 for the NWZA population. As these genotypic disequilibria were only observed in single populations, they are unlikely to be linked. Furthermore comparisons for each locus-pair across all populations did not detect any linkage disequilibria. All loci were polymorphic with the number of alleles ranging from 7 - 17 per loci, totaling 51 alleles scored (Table 6). Most of the populations contained population specific alleles with KALA displaying two unique alleles, SMOZ and NMOZ displaying one unique allele each, and the NWZA population displaying six unique alleles. Only the CAPE population contained no unique alleles. It is noted here that although H_E and H_O is reported here only the H_E is discussed because it is a less biased measure of genetic variability (Nei & Roychoudhary 1974). This is particularly true with low sample sizes.

Table 5 Genetic diversity indices and numbers of *Mellivora* specimens used in the population analyses.

	Mitochondrial control region DNA						Microsatellite DNA (5 loci)		
	# sampled	# of haplotypes	haplotype diversity		nucleotide diversity		# sampled	H_o	H_E
			δ	SD	π	SD			
KENY	10	10	1.000	0.034	0.0487	0.0067	-	-	-
NMOZ	7	6	0.952	0.096	0.0749	0.0070	8	0.62	0.76
SMOZ	9	6	0.917	0.073	0.0119	0.0017	7	0.68	0.80
NWZA	6	5	0.933	0.122	0.0275	0.0046	11	0.63	0.79
KALA	15	8	0.838	0.085	0.0136	0.0016	14	0.75	0.78
CAPE	17	5	0.787	0.059	0.0114	0.0012	15	0.52	0.66
Total	78	41	0.967	0.009	0.0378	0.0030	55	0.68	0.79

Table 6 Number of alleles detected in *Mellivora* for the respective loci. For comparison the number of alleles detected in the taxa from which these loci were originally isolated are shown. The number of alleles detected in *Gulo gulo* loci (Gg-x) was determined from 16 individuals, those from *Martes americana* locus (Ma1) from 30 individuals and those at *Taxidea taxus* loci (Tt-x) from 19 individuals.

Locus	Primer sequence (5' - 3')	Repeat motif in original reference	# of alleles in isolated taxa	# of alleles in <i>M. capensis</i>	Reference
Gg10	F: TTGCAAAACTACAGGACATT R: CCTGGTACAGGTTTATATCAG	(CA) ₂₂	5	17	Walker <i>et al.</i> 2001
Gg454	F: CTTCTTACATAGTCAATGTTTTG R: TGCCATTTTCTCCAGAA	(CA) ₂₀	5	11	Walker <i>et al.</i> 2001
Ma1	F: ATTTTATGTGCGCTGGGTCTA R: TTATGCGTCTCTGTTTGTC	(TG) ₄ TA(TG) ₁₉	10	8	Davis & Strobeck 1998
Tt1	F: AACGGCTTCTAACCCTCCA R: CCCCCTTTTCATTCTTTA	(TG) ₂₀	8	8	Davis & Strobeck 1998
Tt4	F: GGTGAGACCCTGGAATAGAAA R: GCTAACCAACCTAGCAATGAT	(TG) ₁₈	9	7	Davis & Strobeck 1998

Genetic variation

The NJ phylogenetic trees constructed, based on a number of different genetic distances, using composite genotypes from the 55 individuals and on the five populations produced similar or identical topologies. Trees based on distances generated from population groupings suggest the clustering of southern African populations of the CAPE, SMOZ

and KALA separate from the NWZA and NMOZ populations (Fig. 9a). Trees based on individuals display little geographic structure, with minimal clustering of individuals according to geographic origins (Fig. 9b).

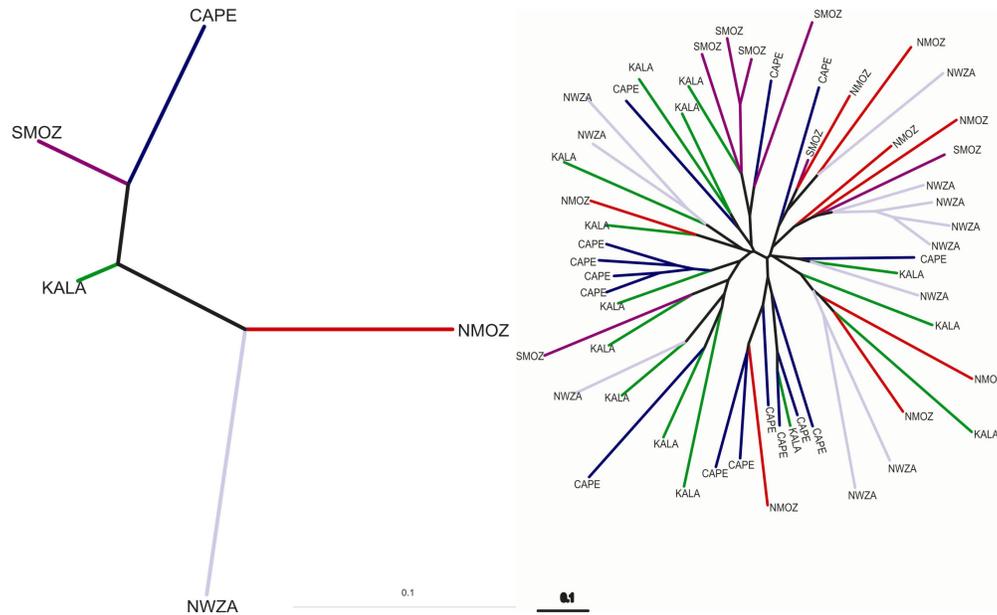


Fig. 9 Phylogenetic trees generated using a.) DAS distances between populations and b.) DAS distances between individuals. Both trees are based on the five microsatellite loci used in this study. Colours correspond to those used in Fig. 3.

The majority of pairwise comparisons between populations showed significant differentiation ($P < 0.05$), using F_{ST} (9 out of 10) and R_{ST} (7 out of 10), with pairwise comparisons ranging from 0.021 to 0.120 and 0.038 to 0.295 respectively (Table 7). Similar patterns of differentiation with regards to the mitochondrial data were found, with the largest F_{ST} and R_{ST} pairwise values found between the southern African populations, excluding the NWZA population, and the east African NMOZ population and the NWZA population. Likewise to the mitochondrial data, all non-significant comparisons were between adjacent populations.

Table 7 F_{ST} and R_{ST} values between the honey badger populations used in this study based on the microsatellite data. Figures above the diagonal represent F_{ST} values between populations and those below the diagonal represent R_{ST} values between populations. Note that the Kenyan population was not included in the microsatellite analysis. Abbreviations as in Fig. 2.

	CAPE	KALA	SMOZ	NWZA	NMOZ
CAPE		0.051	0.061	0.120	0.108
KALA	0.038		0.021	0.052	0.071
SMOZ	0.133	0.057		0.069	0.054
NWZA	0.108	0.091	0.115		0.055
NMOZ	0.270	0.295	0.194	0.064	

Statistically significant comparisons are shown in bold

SAMOVA based on R_{ST} values indicated, similar to the mtDNA, the greatest differentiation when two groups were specified. The east African population of NMOZ was placed separate to the other southern African populations (Table 8). Interestingly, the three-group scenario places KALA and CAPE populations together, NMOZ and NWZA together and SMOZ in a separate group. This contrasts with results obtained from the mtDNA data in that the SMOZ population was placed separately from the KALA and CAPE populations and the populations of NWZA and NMOZ are placed together. Also noteworthy was the differentiation between groups based on a stepwise mutation model decreased slightly, while those based on allele frequency differences increased slightly. In the four-group scenario, the KALA and CAPE populations are grouped while all other populations were placed separately. This is concordant with results obtained with SAMOVA using mtDNA data where these two populations are the last remaining grouped populations, indicating a close evolutionary association between these two populations. Lack of statistically significant R_{CT} and F_{CT} values are probably as a result of low sample sizes. The greater degree of differentiation detected by R_{ST} when compared to F_{ST} indicates that the respective groupings differ in regard to allele size distributions as well as allele frequencies. The mantel test, based on F_{ST} , detects a significant correlation between geographic and genetic distances ($R = 0.83$; $P < 0.01$; Fig. 10).

Table 8 Statistics generated employing groupings identified by SAMOVA. F_{ST} and R_{ST} values generated from microsatellite data. Statistically significant results (generated from 1000 random permutation tests) are shown in bold.

	F_{ST} (p-value)	F_{CT} (p-value)	R_{ST} (p-value)	R_{CT} (p-value)
2 Group				
NMOZ	0.076	0.011	0.214	0.138
SMOZ, NWZA, KALA & CAPE	(0.000 ± 0.000)	(0.379 ± 0.015)	(0.000 ± 0.000)	(0.179 ± 0.000)
3 Group				
NMOZ & NWZA	0.074	0.021	0.160	0.126
SMOZ	(0.000 ± 0.000)	(0.117 ± 0.011)	(0.000 ± 0.000)	(0.065 ± 0.000)
KALA & CAPE				
4 Group				
NMOZ	0.073	0.024	0.151	0.123
NWZA	(0.000 ± 0.000)	(0.101 ± 0.008)	(0.000 ± 0.000)	(0.091 ± 0.000)
SMOZ				
KALA & CAPE				

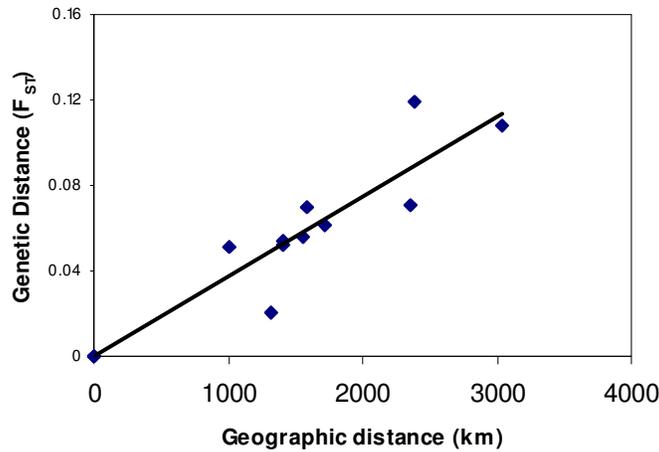


Fig. 10 Relationship between genetic distance and geographic distance. F_{ST} based on sampled populations incorporating microsatellite data. Diamonds represent comparisons between populations. Note that the Kenyan population was not included in the microsatellite analysis. Solid line represents the fitted regression line.

Chapter 4: Discussion

Intraspecific variation

In the present study the microsatellite loci exhibited higher variation when compared to overall nuclear heterozygosity found in other mustelid taxa that have similar life history traits such as population densities and dispersal capabilities. For example, the average expected heterozygosity for all microsatellite loci across populations (0.79) is higher than that recorded in the American mink (*Mustela vison*: $H_E = 0.61$; O'Connell *et al.* 1996; Brusgaard *et al.* 1998 a, b), wolverine ($H_E = 0.39 - 0.63$; Duffy *et al.* 1998; Davis & Strobeck 1998; Walker *et al.* 2001; Kyle & Strobeck 2001, 2002), fisher (*M. pennanti*: $H_E = 0.62$; Kyle *et al.* 2001) and American marten (*M. americana*: $H_E = 0.58 - 0.68$; Davis & Strobeck 1998; Kyle *et al.* 2000; Kyle & Strobeck 2003; Small *et al.* 2003) but similar to those found in the American badger ($H_E = 0.77$; Kyle *et al.* 2004). Moreover, given that the loci used in the present study were not designed for this species and the fact that small sample sizes were used is indicative of the high level of heterozygosity found in the honey badger. Although it must be kept in mind that small sample sizes could have resulted in underestimating the true diversity found in this species. This small sample size is however, expected to affect the allelic diversity to a greater degree than the expected heterozygosity (Nei *et al.* 1975).

The high level of nuclear DNA diversity is supported by the mitochondrial haplotypic diversity in the honey badger ($\delta = 0.97$), which was once again higher than that reported for other mustelids. For example, among European otter (*Lutra lutra*), only five haplotypes were detected in 129 individuals ($\delta = 0.48$; Cassens *et al.* 2000) and in the pine marten (*M. martes*) just 25 distinct control region sequences (six of these being introgressed from a closely related species) were found in 139 individuals (Davison *et al.* 2001). This is also consistent with the low haplotypic diversity found in analyses of polecats (*M. putorius* and *M. eversmannii*) and the European mink (*M. lutreola*; Davison *et al.* 2000). The most striking example of low haplotypic diversity typically observed in highly mobile mustelid species is probably found in the wolverine (Walker *et al.* 2001)

where only a single mitochondrial haplotype has been detected. Mitochondrial DNA nucleotide diversity of the honey badger ($\pi = 0.038$) compared to that found in the Eurasian otter ($\pi = 0.017$; Cassens *et al.* 2000), pine martens ($\pi = 0.013$; Davison *et al.* 2001) and polecats ($\pi = 0.009$; Davison *et al.* 2001) was also higher even though the published data were based on more extensive sampling.

From the above it is clear that in most instances the genetic diversity of the African honey badger is noticeably higher than the values found in European taxa. It has been proposed that the evolution of the principally European species were shaped by repeated periods of glaciation during the recent past (Cassens *et al.* 2000; Davison *et al.* 2001). In the cold adapted species it has been proposed that, during inter-glacials, the south Mediterranean probably represented the extent of their range and the refugial populations at range limits could have gone extinct (Davison *et al.* 2001). An alternative explanation is that these European species have lower effective population sizes leading to a loss of diversity. Haplotypic and nucleotide diversity of the honey badger, however, is also high when compared to non-European species such as the least weasel *M. nivalis* and ermine *M. erminea* although these species are considered relics of the last glacial period and the same processes may have shaped the genetic structuring of these species as well (Kurose *et al.* 1999). With respect to the European otter it has been proposed that glaciations resulted in a Pleistocene refugium resulting in the low genetic diversity in this species (Cassens *et al.* 2000).

The honey badger populations from east and southern Africa have not recently experienced any severe reductions in population size (also see below) and from a conservation perspective this species seems genetically vigorous. The lack of sufficient comparable genetic data from African mustelids, however, prevents any comparisons. Nevertheless, the level of mitochondrial diversity is high in the case of *M. capensis* when compared to that of the nearly extinct African wild dog, which has similar life history characteristics in terms of mobility. Although haplotype diversity is not indicated in the wild dog study only nine control region genotypes were found in 280 individuals ($\pi =$

0.001; Girman *et al.* 2001). The low diversity present in the African wild dog has been attributed to the recent population bottleneck but the pack structure of the wild dog where only a single female breeds will reduce the haplotype diversity expected in this species due to the smaller effective population size. Mitochondrial variance of the honey badger was comparable to the leopard *P. pardus*, a carnivore with a similar distribution to the honey badger similarly inhabiting many habitat types and feeding on a broad range of prey. In 69 leopards sequenced 33 haplotypes were defined. ($\pi = 0.012$; Uphyrkina *et al.* 2001). In concert these data tend to suggest that the recent evolutionary environment in Africa was possibly more stable than those in Europe during the same evolutionary period.

Phylogeography

Inspection of the genetic structure using SAMOVA/AMOVA, incorporating both microsatellite and mitochondrial data, indicate that the greatest amount of differentiation of the honey badger can be accounted for by a two group structure (one group including the southern African populations and the other including the east African population/s). Although microsatellite F_{CT} data indicate greater differentiation with an increase in the number of groups specified, the increase in differentiation is small, especially relative to the differentiation detected employing R_{CT} . The increase in differentiation with an increase in the number of groups is possibly an artifact of the small sample size resulting in the lack of detection of alleles in populations. The differentiation between eastern vs. southern Africa is also supported by the phylogenetic analyses where the majority of haplotypes sampled in southern Africa cluster within clades 1 and 2 and most individuals sampled in eastern Africa cluster within clades 3 and 4. Consistent with the geography, Clade 4 also includes individuals sampled from Israel and Chad (Figs. 3 & 4). The 3-group structure based on the mitochondrial data, supported by significant F_{CT} and Φ_{CT} values, also indicate a break between southern African and eastern African populations, with a possible further genetic break in central Africa (see below). This pattern is clearly supportive of a possible genetic break between eastern and southern Africa.

Alternatively, it is possible that the differentiation between eastern and southern Africa is merely the result of isolation by distance (a significant correlation is detected between geographic and genetic distance measures calculated for both mitochondrial and microsatellite data). Furthermore the observed groupings identified by SAMOVA analysis could merely be a consequence of the geographic arrangement of the sampling localities. For example, the opportunistic sampling strategy employed in this study could have resulted in the observed grouping of populations when an overall pattern of isolation by distance is present and particularly when one considers the differences in distances between the respective populations (the closer geographic proximity of the southern African populations of CAPE, KALA and SMOZ relative to the populations of NMOZ, KENY and NWZA). It is important to also realize that a significant association between geographic and genetic distance holds for the mitochondrial data even when introgressed haplotypes that do not group in their expected major geographic area, namely southern and eastern Africa, were excluded from the analyses.

The patterns observed could best be ascribed to being caused by a combination of demographic processes. It appears that a phylogeographic pattern of isolation by distance is present among the sampled populations for both the mtDNA and nuclear analyses. Apart from isolation by distance the phylogenetic analyses, network, mismatch distributions and SAMOVA/AMOVA strongly suggest the existence of at least two geographic groupings probably attributed to a geographic barrier to gene flow. Among the clades most of the intermediate haplotypes were lost but some movement among populations are evident from the introgressed haplotypes. It is interesting to note that a closer look at the isolation by distance analyses (employing mtDNA data) lend some support for a genetic break between eastern and southern Africa in that the majority of comparisons within the main geographic areas (i.e. southern and eastern Africa) fall below the fitted regression line while the majority of those between geographic areas lie above the regression line.

Further support for a genetic break between the two honey badger clades can be found in the congruence with other African fauna. A genetic break between eastern and southern Africa has been detected in numerous vertebrate species (Freitag & Robinson 1993; Matthee & Robinson 1997; Arctander *et al.* 1999; Girman *et al.* 2001; Matthee *et al.* 2004). A number of processes have been proposed to account for these patterns identified in other African taxa. As mentioned previously, the distribution of the Miombo (Brachystegia-Julbernardia) woodland is probably not an effective barrier to gene flow in the honey badger (also see Girman *et al.* 2001). The Great Rift Valleys has also been proposed as a potential barrier to gene flow (Freitag & Robinson 1993; Matthee & Robinson 1997; Arctander *et al.* 1999; Girman *et al.* 2001; Matthee *et al.* 2004). Given the distribution of the honey badger clades, the western branch of the east African rift valley system cannot be excluded as a potential barrier to gene flow in the honey badger. A third possible barrier could be the lower Zambezi River (Fig. 2) which is one of the largest rivers in Africa, and could potentially acted as a partial barrier to gene flow in this species. The evolution of the Zambezi and the surrounding drainage systems has been dynamic with the mid and lower Zambezi River being relatively stable for at least the last 2.5 My (Moore & Larkin 2001). However, the effect of a river as a barrier limiting gene flow in the honey badger may be questionable, as it has been reported that honey badgers are good swimmers and even pursue turtles in aquatic environments (Kingdon 1977). In addition the evidence that rivers act as a barrier to gene flow in terrestrial vertebrates is not strong (Gascon *et al.* 2000) although the Amazon River has been proposed as a partial barrier to gene flow in the jaguar *P. onca* (Eizirik *et al.* 2001).

Rifting in East Africa began as far back as 70 My ago, although most of the uplift and magmatic activity took place between 25 and 10 My ago (Partridge *et al.* 1995), with tectonic and magmatic events continuing to influence vertical movements of these geological structures up to the present (Ebinger 1989). It has been suggested that the western and eastern branches of the East African rift system had already been formed by approximately 17 My ago (Partridge *et al.* 1995). The estimated time of divergence between *Mellivora* southern and eastern clades is probably, 521 700 - 782 800 years ago.

Even if this estimate is an order of magnitude older or younger the actual divergence of clades, it is reasonable to suggest that the time of divergence between the two clades is considerably younger than the formation of the rift system. Because the potential physical barriers to gene flow are older than the divergence of the eastern and southern clades it is more likely that the rift acts as a partial barrier to gene flow if at all. Rifting in east Africa, in conjunction with climate fluctuations has been implicated in the fragmentation of habitat and the contraction of ranges, and the observed pattern is probably more likely a consequence of the climatic shifts and associated vegetation transformations during the Pleistocene, which could have led to severe habitat changes. These habitat changes might have affected the distribution of prey items causing temporal isolation among honey badger populations. In conjunction with weak physical barriers such as the lower Zambezi and the eastern rift, we can only suggest that a combination of these factors might have led to limited gene flow between the southern and eastern African populations. In fact, the presence of introgressed haplotypes tends to support the notion that no firm geographic barriers exist. In addition, certain life history characteristics of the honey badger, such as its ability to survive in harsh environments including the fringes of the Sahara and pro-Namib deserts, would suggest that the climatic shifts during the Pleistocene must have been particularly severe to have had any impact on the phylogeographic structuring of this species. Interestingly, the phylogeographic structure of the common warthog, *Phacochoerus africanus*, also able to survive in various, particularly arid, ecotypes has also been shown to be structured between eastern and southern Africa (Muwanika *et al.* 2003). Thus it appears that regional climatic changes lead to increased aridity, with resulting changes in habitat. Paleoclimatic evidence does indicate that for the late Pliocene and Pleistocene the African continent was experiencing a period of aridity (Hamilton & Taylor 1992).

The mismatch distribution performed on the entire mitochondrial data set displays a bimodal distribution. This combined with the geographical distribution of the clades, (clades 1 and 2 occurring almost exclusively in southern Africa and clades 3 and 4 occurring predominantly in eastern Africa) supports the interpretation of vicariance

followed by secondary contact (see Hoffmann *et al.* 2003). The presence of diverged mtDNA clades within a species that show some degree of sympatry has been interpreted as an indication of a vicariance event followed by secondary contact between previously isolated lineages. Another possible interpretation of the absence of intermediate haplotypes could be that it is a sampling artifact where geographically intermediate sampling localities (where intermediate lineages could be found) were not included in the present study (Templeton 1998). The sampling strategy, although highly opportunistic, does seem adequate to detect intermediate haplotypes between clades. The haplotype represented by individual NM1, which is in an intermediate position between the majority of the eastern and southern African lineages may represent a relic during the diversification of southern from eastern lineages.

SAMOVA/ AMOVA using mitochondrial data indicates that some phylogeographic structure is present between the two east African populations. This is suggested because the amount of variation that can be explained between groups increases to levels similar to that explained by the 2-group scenario, when going from three groups to four groups. This is achieved when separating the two east African populations, indicating that these populations contain some geographic structuring of haplotypes. It is unlikely that isolation by distance is the only force shaping the genetic structure in this region because we do not see the same level of differentiation in the southern African populations (which are not geographically further from each other than the two east African populations) which in general, share the same haplotypes or closely related haplotypes. Although speculative the Kenya (Gregory) rift valley of the east African rift system may act to impede gene flow within east Africa, namely between the populations of KENY and NMOZ, although further sampling in east Africa is needed before conclusions of this nature can be drawn (also see Matthee and Robinson 1997). Unfortunately because the KENY population was omitted from the microsatellite analysis no similar nuclear comparisons can be made. Interestingly in the 4-group scenario the three southern populations of SMOZ, CAPE and KALA form a group and the other three populations form their own groups. When this group is split in the 5-group scenario the

differentiation between groups decreases noticeably. This suggests close evolutionary associations for these three southern African populations.

Populations in east Africa show higher nucleotide and haplotype diversity, despite the inclusion of fewer samples from this area (Table 5). This coupled to the higher population parameter, τ , for these clades indicate that the eastern African assemblages are older than the southern African clades. It is also interesting to note that the Indian and Israel samples are most closely related to the east African assemblage, suggesting contact via land bridges from east Africa (see Robinson & Matthee 1999). Based on the genetic data it is suggested that *M. capensis* has spread into southern Africa from northern or east Africa (see below). Microsatellite H_E values show some support for this as the single population from east Africa, despite the lowest sampling, has, a moderate to high H_E level, whereas that of the southernmost population that has the most extensive sampling, has the lowest H_E values. Although tentative, rooting the NJ and parsimony trees with the numt sequence implies that the east African population is more basal than the southern African population. Furthermore the earliest fossil record for the genus *Mellivora* is approximately 10 My old and originates from Kenya (Bishop & Pickford 1975). East Africa may thus have acted as a refugia for this genus during older extreme climatic cycles with subsequent expansion into southern Africa. This is somewhat contrary to fossil data for the species suggesting that *M. capensis*, evolved in southern Africa following speciation with, *M. benfieldi*, found 3.5- 7 My ago in southern Africa (Hendey 1978).

The lower genetic differentiation at nuclear loci in comparison to the mitochondrial marker may reflect a pattern of male biased dispersal within this species. This is supported by observed sexual differences in the spatial organization of honey badgers, with larger average home ranges documented in males (541km²) than in females (126km²; Begg 2005a). Although if one considers that F_{ST} value has an upper bound of $1 - H_E$ (Hedrick 1999), the honey badger has an upper bound of approximately 0.3

indicating a higher level of differentiation than originally apparent, although still lower than that detected employing the mitochondrial loci.

Given the wide distribution of the honey badger it is possible that within the species, local populations might exist that have certain habitat preferences. For example, the forest dwelling honey badgers (*M. c. cottoni*; Fig. 2) are allegedly habitat specific and display morphological differences from the rest of this species (Nowak 1999). This possible habitat preference can lead to low levels of gene flow between forest dwellers and non-forest dwellers. Although rather speculative, the distinctiveness of clade 2, in particular the subnetwork consisting of three closely related haplotypes, from the other southern African clade could be evidence of a lineage becoming isolated due to habitat preference. The haplotypes of a number of individuals sampled in northwestern Zambia, bordering the proposed subspecies *M. c. cottoni*, cluster within this clade. Included in this clade is a museum sample from the Ituri forest situated in the Democratic Republic of the Congo where these morphological differences have been reported. This population is also the first southern African population to be placed in its own group separate from the other southern African populations in SAMOVA analysis based on mitochondrial data. This population also contains the highest number of unique alleles and highest H_E . Interestingly genetic evidence for character differentiation, between dark (found more frequently in woods) and 'typical' forms has been evidenced in the polecat *M. putorius*, a mustelid occurring in Europe (Lodé 2001). It is possible that genetic admixture between two previously isolated forms is taking place in this region. This could explain the observed pattern of haplotypes sampled in this region clustering in two distinct clades and the high microsatellite diversity. Dating the split of the two southern African maternal lineages places the divergence at 39 000- 56 500 for these two clades although, as previously mentioned, these dates must be treated with caution. Further sampling is needed in this area to investigate this possible habitat preference or whether the observed pattern is simply an artifact of sampling (this southern African population is the most geographically distant of all the southern African populations) in conjunction with a pattern of isolation by distance. Such a pattern of habitat association has also been

suggested in the highly vagile grey wolf (*Canis lupus*; Geffen *et al.* 2004). In this study it was suggested that phenotypic differences between grey wolf populations may be maintained by environmentally induced restrictions in gene flow and that this allows genetic drift and natural selection to act, resulting in genetic and morphological differentiations amongst populations of this species.

Currently the honey badger displays a disjunct distribution between Africa and Asia (Fig. 1), there is thus limited or no gene flow between present day African and Asian populations. The single haplotype found in two individuals from Israel cluster within the east African clade 4. Similarities have been found between honey badgers found in Africa and the Arabian Peninsula (Cheesman 1920; Pocock 1946). More recently morphological evidence based on skull morphometrics shows similarities between Arabian and African forms (Baryshnikov 2000). Although tentative, given the limited sampling in this region, it is possible that historical levels of gene flow could have occurred between Arabia and east Africa. This could have been facilitated by a landbridge between the Sinai and the Egypto-somalian domain, which was closed by the Red Sea and the Gulf of Suez during the late Pleistocene. Gene flow could have also taken place via landbridge connections, between Africa and the Arabian Peninsula, up to 11 000 years before present (Delany 1989).

Implications for conservation

Although an overall pattern of isolation by distance can be detected between eastern and southern Africa it is recommended that these geographic areas should be regarded as separate management units and translocations between these broad geographic regions should be discouraged. This is based on the independent demographic histories of these two geographic regions. Although there has been more recent gene flow between areas it has not been sufficient to homogenize the distribution of the mitochondrial clades and the greatest degree of differentiation using mitochondrial and microsatellite data is also found between these two regions. These regions may then have acquired adaptive changes (e.g. Hoekstra *et al.* 2005) warranting their separate management. The close association

between the three southern African populations of the CAPE, KALA and SMOZ indicate that translocations can be performed within this region, although taking into consideration the overall pattern of isolation by distance it would be good management practice to source genetic material for translocation from geographically proximate locations within this region. Within east Africa there is some evidence of geographic structuring of maternal haplotypes. Microsatellite data on the Kenyan population is lacking, and accordingly translocations between regions in east Africa must be discouraged until further sampling can be performed in this area. Taking a precautionary approach, until further sampling is done in central Africa, translocations should be avoided between southern, east and central. It should be noted that as a preliminary investigation only a portion of the range of the honey badger has been sampled. Further sampling and increasing the number of microsatellites employed (which will allow the use of analyses such as assignment tests) will allow for more rigorous testing of the phylogeographic structure of this species.

Subspecies descriptions

Our limited sampling makes the subspecies comparisons moot. Nevertheless, previous subspecies descriptions broadly correspond to the phylogeographic patterns detected. Most notably is that most of the southern African taxa share a close evolutionary history and are all considered part of the same subspecies, *M. c. capensis*. There are however some important differences, such as the exclusion of the NMOZ specimens. The genetic analyses rather suggest that this latter population shares a closer evolutionary association with individuals from Kenya. Nuclear data also indicate that the NMOZ population is most divergent relative to the other populations. There appears to be some structuring between these two latter populations, warranting its placement in a separate subspecies, *M. c. maxwelli*. Further sampling between these two regions and the inclusion of nuclear data for the Kenyan population are needed before firm conclusions can be drawn.

Conservation Implications at a Wider scale

The protection of threatened species, habitats and regions while employing limited economic resources has recently been on the forefront in the field of conservation biology (Master 1991; Amori & Gippoliti 2000). Although various criteria have been suggested to identify priority areas for conservation (e.g. Myers 1988; Freitag & van Jaarsveld 1997), molecular data may assist with interpretations of historical geography and determine how these events shaped the phylogeographic structuring of species. If many taxa share similar patterns of genetic structuring, the evolutionary forces responsible must have had an overall dominant effect. This has implications for conservation biology and the management of taxa. For example, such concordant genetic information may provide guidelines regarding translocations and taxonomic reassessments, particularly on inadequately investigated taxa,. Importantly too, in the absence of genetic, or other, data for species of conservation concern, concordant phylogeographic structuring detected by genetic markers, may provide provision for the recognition of regional reserves to preserve centers of biodiversity. Thus on a broader scale, looking at an ecosystem level to identify processes that effect habitat fragmentation for a number of taxa, broader conservation plans may be drawn up. These may help conserve many of the less charismatic taxa that may not receive conservation efforts. Considering that a number of genetic investigations have focused on east and southern Africa, using taxa representative of a wide variety of dispersal and life history characteristics (as suggested by Avise 1992), the time will soon be at hand to do a comprehensive regional case study of this area (see Avise 1992; Hewitt 2004). It is hoped that this can then be used in the active management of the area for the preservation of the biodiversity of Africa.

CHAPTER 5: Co-amplification of Numt Sequences and the Confirmation of MtDNA Control Region Sequences

Introduction

MtDNA has frequently been employed to determine genetic structure, particularly among mammal species, and although these studies have proved very informative they have also highlighted certain pitfalls with the use of this marker. For most animal species investigated it appears that the mitochondrial DNA within an individual is homoplasic at the nucleotide sequence level, although length heteroplasmy has been observed in some taxa (see Rand 1993 for review). The use of mtDNA as a population genetic marker, however, can be plagued by the existence of multiple mitochondrial or mitochondrial-like sequences (Zhang & Hewitt 1996a). These pseudo sequences can either be due to mtDNA heteroplasmy (Awise 1991; Rand 1993; Hoelzel *et al.* 1994) or the existence of nuclear copies of mitochondrial sequences (Lopez *et al.* 1994; Mourier *et al.* 2001; Zhang & Hewitt 1996b). Heteroplasmy can be the result of 'paternal leakage' (Kondo *et al.* 1990; Gyllensten *et al.* 1991; Kvist *et al.* 2003), mutations of the mitochondrial genome within a cell (Lunt *et al.* 1998; Steel *et al.* 2002) or heteroplasmy in the oocyte of the organism (Steel *et al.* 2002).

The endosymbiotic theory seems to provide the evolutionary basis for the wide occurrence numts in the nuclear genome. The serial endosymbiosis theory proposes that the mitochondria arose from symbiotic free-living prokaryotic ancestors (Margulis 1970). With the sequencing of the complete mt genomes of many organisms as well as the sequencing of complete genomes of α -proteobacteria, particularly *Rickettsia*, the evidence is strong that the mitochondrion, and accompanying genome, are of endosymbiotic origin derived from a once free living bacterium, probably an α -proteobacterium (Gray 1999; Gray *et al.* 1999, 2001). If this (almost universally accepted) theory holds, some function has transferred from mitochondrial to nuclear genomes as mtDNA does not encode all the proteins required for organellar function

(examined mitochondrial genomes contain 1 - 3 orders of magnitude fewer genes than their free living ancestors) and nuclear gene products are transported into the mitochondrion where they play a role in biochemical pathways (Hartl & Neupert 1990). The movement of mtDNA into the nuclear genome must then be an important process in the evolution of these two genomes (with many mitochondrial genes being transferred to the nucleus, while others have been replaced by nuclear genes of similar function). Transfer of DNA from the mitochondrion to the nucleus is an ongoing process and can result in the transferred gene becoming functional. Typically, though, sequences recently transposed from mitochondrial DNA into nuclear DNA are non-functional because of differences in the mechanisms regulating the control and expression of mtDNA and nuclear DNA and the presence of frameshifting indels (Lopez *et al.* 1994). These nonfunctional copies of mtDNA in the nuclear genome are thus expected to evolve as nuclear pseudogenes and are typically referred to as numts.

Numts have been documented in plants (e.g. Blanchard & Schmidt 1995) and animals (reviewed recently by Bensasson *et al.* 2001). In some taxa they have been shown to be numerous (e.g. Zhang & Hewitt 1996b; Mourier *et al.* 2001). These pseudogenes range in size from small fragments of transferred DNA to the insertion of almost the entire mt genome in *Arabidopsis thaliana* (Arabidopsis Genome Initiative 2000) and humans (Mourier *et al.* 2001). Numts have also recently received more attention because they are often preferentially amplified during PCR and can significantly obscure phylogenetic studies.

Data suggest that numts arise as a result of non-homologous recombination of nuclear DNA with fragments of mtDNA that have leaked from damaged mitochondria and that this is a continuous evolutionary process (Mourier *et al.* 2001). The duplication of pre-existing numts could also be involved in their growth in number in the nuclear genome (Lopez *et al.* 1994). The large numbers of studies that are being published reporting the occurrence of numts (more than 82 species - Bensasson *et al.* 2001), suggest that this process may be frequent (also see Richly & Leister 2004). Many of the studies reporting

numts propose that the origin of mt sequences generated by PCR for phylogeographic investigations must be evaluated particularly if these sequences were generated employing universal primers (e.g. Zhang & Hewitt 1996a; Mirol *et al.* 2000).

Various ways have been suggested to recognize or identify the amplification of multiple mitochondrial or mitochondrial-like sequences. The first evidence of a mtDNA numt is often the existence of non-specific PCR amplification, which results in extra products in restriction profiles or PCR. Numts also typically result in sequence ambiguities, frameshift mutations, unexpected phylogenetic placements or nucleotide sequences different from those expected and misplaced stop codons in coding regions (Arctander 1995; Collura & Stewart 1995; Zhang & Hewitt 1996a). In the honey badger heterozygous reads in DNA were evident in our attempts to sequence the mitochondrial control region by making use of the universal primers listed previously (Chapter 2).

The main aim of this part of the study was to ensure the generation of mtDNA sequence for phylogeographic comparisons. A technique had to be designed to ensure that only orthologous mtDNA sequences are used.

Materials and Methods

Total genomic extraction and subsequent amplifications and sequencing were performed as described previously (Chapter 2). In many of the individuals examined either the chromatopherograms could not be interpreted because of the prevalence of double peaks throughout the sequence and/or contained significant background noise (heterozygous substitutions). PCR products were run out for approximately five hours at 80V on 2% (w/v) low melting point agarose gels and two PCR products were noted. Both products were sequenced.

In order to determine the genomic origin of the sequences a mitochondrial-enriched extraction was performed on a single individual, displaying evidence of multiple PCR products. It was expected that if the additional PCR products were caused by a numt, only the authentic mitochondrial copy would be amplified, whereas if the duplicate copy was of mtDNA origin, sequence ambiguities would still be present. This procedure was performed from frozen heart muscle tissue, employing a modified protocol for plasmid organelles (e.g. Birnboim & Doly 1979; Taylor & Brose 1988; Welter *et al.* 1989) as proposed by Palva & Palva (1985). This method was chosen as it is a rapid and simple technique, does not require monitoring with a pH meter or the use of an ultracentrifuge and, importantly for population genetic studies, does not require large amounts of fresh tissue. It is important to realize, however, that this procedure does not guarantee that numts will not be amplified (Birnboim & Doly 1979; Beckman *et al.* 1993) but, as a first step, we regarded this protocol as sufficient for the purpose of this study.

To further optimize the mitochondrial amplification of the PCR product, decreasing amounts of template DNA was used in subsequent PCR's. This can potentially lower the concentration of nuclear DNA below a threshold necessary for co-amplification. The dilution that contained the lowest concentration of template DNA but still resulted in

amplification products of sufficient quality to sequence was used for sequencing. It is important to also realize that during amplification “universal” primers may be especially prone to amplifying numts (Sorenson & Quinn 1998). This is because numts are more likely to evolve slower than the mt copy following the transposition event and may thus be more similar to the ancestral sequence (Arctander 1995, Collura & Stewart 1995). Thus these methods of employing a mitochondrial enriched extraction and decreasing DNA concentrations can be compromised if the primers used preferentially amplify the nuclear insertions. To determine whether either of the PCR products was preferentially amplified the stringency of the original PCR conditions were made more stringent, by increasing the primer annealing temperature and/ or reducing $MgCl_2$ concentrations.

The identification of a nuclear copy is more difficult in non-coding areas such as the control region as they lack coding structure. Consequently, in addition to identifying the mitochondrial sequence obtained from the mt enriched extraction as being of mitochondrial origin the authenticity was further examined by linking this to adjacent coding regions to examine their coding structure. Amplification of approximately 2 600 bp of the mitochondrion, including cytochrome b, control region and 12S rRNA sequences was obtained using the primers L15162 and 12Srev-dloop (Table 1). The PCR and sequencing conditions were similar to those described in Chapter 2. From this single PCR product two regions were sequenced. The first segment, a fragment approximately 800 bp on the 3' side of the cytochrome b gene was sequenced using the L15162 primer and the second fragment, approximately 500 bp on the 5' side of the control region, was sequenced using the primer N777. In an attempt to further confirm the authenticity of the mtDNA the linked cytochrome b stretch was translated to amino acid codons. This was done based on the assumption that the non-functional nuclear homolog should be free of selective forces and could have experienced mutation events that inactivate the mitochondrial cytochrome b insert (indicative of the pseudo copy).

Results and Discussion

The universal primers resulted in two PCR fragments differing slightly in length. The key difference between the two fragments was the existence of three indels at the 5' end of the control region (15, 1 and 12 bp in length respectively; Fig. 11). The sequences showed a fairly high amount of homology but at least 23 site changes separate them in the 230bp region used for the phylogeographic study (Fig. 11).

```

numt      CCCTGATTTCTCACCTCCCTAATTTCAATTGATATATTTAATAATATTT
mt        CCCTGGTACCCAT-----ATTCATATACTCAGCAGGCC-
Ambiguities -----*_*_*-----*_*_*-----*_*_*-----*_*_*-----

numt      ACTGTGCTTCCCCAGTATGTACCTTTTCTTTTCCCTCCCCTATGTATAT
mt        ACTGTGCTTCCCCGGTATGTCGG-----CCCCCTATGTACCT
Ambiguities -----*_*_*-----*_*_*-----*_*_*-----*_*_*-----

numt      CGTGCATTAATGGTTTGCCCCATGCATATAAGCATGTACATATCATGATT
mt        CGTGCATTAATGCTTTGCCCCATGCATATAAGCATGTACATACTATGATT
Ambiguities -----*_*_*-----*_*_*-----*_*_*-----

numt      GGCTTTACATGCATGTATTTCACTTAGACCACGAGCTTGATCACCAGGCC
mt        GATCTTGCATACGTGCATATCACCTAGATCAGGAGCTAATCACCAGGCC
Ambiguities -----*_*_*-----*_*_*-----*_*_*-----*_*_*-----

numt      TCRAGAAACCATCAATCCTTGKCGCATGTGTACCTCTTCTCTCTCGGGC
mt        TCGAGAAACCATCAATCCTTGCGCGATGTGTACCTCTTCTCGCTCCGGGC
Ambiguities -----*_*_*-----*_*_*-----*_*_*-----*_*_*-----

numt      CCATCAAACGTGGGGTTTCTWTACTGAAACTATACCTGGCATCTGGTTC
mt        CCATCAAACGTGGGGTTTCTACGTTGAAACTATACCTGGCATCTGGTTC
Ambiguities -----*_*_*-----*_*_*-----*_*_*-----

numt      TTACTTCAGGGCCATGACAGTCCTYAATCCAATCCTACTAACCTCTCAA
mt        TTACTTCAGGGCCATGATAGTGTGAATCCAATCCTACTAACCTCTCAA
Ambiguities -----*_*_*-----*_*_*-----*_*_*-----

numt      TGGGACATCTCGATGGACTAATGACTAATCAGCCCATGATCACACATAAC
mt        TGGGACATCTCGATGGACTAATGACTAATCAGCCCATGATCACACATAAC
Ambiguities -----*_*_*-----*_*_*-----*_*_*-----
    
```

Fig. 11 Alignment of 400bp of the presumed numt and authentic mitochondrial sequence, with ambiguities between sequences shown with a *. Both sequences were obtained from the same individual. Underlined regions on the mitochondrial sequence indicate where designed primers are situated.

By increasing the stringency of the original PCR (by adjusting annealing temperature and MgCl₂ concentration), neither of the PCR products was preferentially amplified, as determined by the brightness of their respective bands on 2% (w/v) agarose gels. In addition the amplification of both types of sequence using the same primer pair indicate a high degree of homology in regions where primer annealing occurs. This indicates that in

the honey badger, primer annealing will not preferentially amplify any nuclear copy that may be present in the mitochondrial enriched extraction and lend further support for using the mitochondrial enriched extraction procedure.

A single unambiguous sequence with no background artifacts was obtained when the mt enriched DNA extract was used in PCR. This putative control region sequence was linked to a cytochrome b sequence that did not experience any frame shift mutations or stop codons in approximately 800 bp sequenced. This stretch of the cytochrome b gene contains 38 residue positions conserved across a wide variety of taxa (Esposti 1993), and all of these were conserved in the putative honey badger cytochrome b sequence. This provides additional evidence that the sequence is a functional cytochrome b gene of mtDNA origin. The latter is however based on the assumption that enough time has lapsed since the nuclear insertion event.

Primers that selectively amplify the putative mt sequences were designed by comparing the authentic mt sequences to the identified numt (Sorenson & Quinn 1998). Due to the high degree of homology between the control region and its nuclear homologue the primer Badger DLL (Table 1) was designed on the 5' end of the control region where the numt displays an insert, so as to be specific to the mitochondrial copy only. Another internal primer Badger DLH (Table 1) was also designed to amplify a shorter stretch, of approximately 250 bp, in samples containing degraded DNA particularly the museum skin samples. Further internal primers were designed to amplify this stretch in two overlapping fragments in highly degraded samples (Table 1). The primers were specifically designed to ensure that annealing was more specific to the mitochondrial copy as opposed to the numt (Fig. 11). Furthermore at least one of the primers in a pair used in PCR was highly specific to the mitochondrial copy and in theory could not amplify the numt due to 3 prime incompatibility.

Additional evidence for the authentication of the mtDNA copy lies in that fact both fragments were sequenced for eight individuals. The reactions using primers designed to

preferentially amplify “mtDNA copies” produced variable sequences among individuals (Chapter 3). In these same individuals, the sequences of the other type (identified as a numt), are identical in sequence, displayed the same indels and only displayed heterozygous sites in some individuals (Fig. 11). The sampling of the eight individuals comprised a large geographic area (CAPE, KALA and KENY). What was particularly striking is that the individual from KENY, whose mitochondrial sequence exhibits significant sequence differentiation when compared to the other southern African haplotypes (Figs. 3 & 4) are identical when numts are compared. The identical sequences identified as a numt provides further evidence for the nuclear origin of these sequences as it is consistent with observations that the rate of evolution of non-coding regions in the nuclear genome is slower than the control region of the mitochondria (10 to 40 times; Ward *et al.* 1991; Lundstrom *et al.* 1992; Arctander 1995). In a direct comparison numt sequences display 23-29 bp differences relative to the authentic mitochondrial sequences over the 230bp stretch sequenced, whereas the extant *Mellivora* mitochondrial lineages identified in this study contain a maximum of 22 bp differences between lineages.

It is unlikely that the remaining sequence was generated as a result of mitochondrial heteroplasmy because the mt enriched extraction procedure generated only a single unambiguous copy. Additionally the differences in sequences do not involve the variation in the number of repeats in the control region which accounts for most of the mutations resulting in heteroplasmy in animals and is not consistent with a repetitive region in the 3' end of the control region that has been described in carnivores (Hoelzel *et al.* 1994). Furthermore it is unlikely that heteroplasmy, by any of the three mechanisms resulting in more than one type of mt DNA molecule in the cell, would have resulted in variable, geographically partitioned, mt haplotypes whereas the other copy is identical, excluding ambiguous sites, with no apparent geographic partitioning. The presence of consistent ambiguous sites in an individual (even after re-sequencing), when it appears that only one type of sequence is being amplified, is difficult to explain if caused by heteroplasmy, but is more readily explained if a numt is invoked, as the ambiguous sites can be elucidated as a heterozygous state in the nuclear copy.

The identification of a nuclear introgression from the mitochondrion is not unusual (e.g. Quinn 1992; Arctander 1995) particularly in mammals (Zullo *et al.* 1991; Collura & Stewart 1995; Lopez *et al.* 1994; Zischler *et al.* 1995). It is likely that their presences in evolutionary studies are often overlooked (Thalmann *et al.* 2004). Following this, it should be noted that in a comprehensive study of the population structure of the Scandinavian wolverine, nuclear DNA variation was detected in the absence of any mtDNA variation (a single haplotype was observed; Walker *et al.* 2001). In the light of the conserved nature of the numt detected in the honey badger, the latter study may warrant further investigation.

REFERENCES

- Amori G, Gippoliti S (2000) What do mammalogists want to save? Ten years of mammalian conservation biology. *Biodiversity and Conservation*, **9**, 785-793.
- Anderson E (1989) The phylogeny of the mustelids and the systematics of ferrets. In: *Conservation biology and the Black-footed ferret* (eds. Seal US, Thorne ET, Bogan MA, Anderson SH), pp 10-20. Yale University Press, New Haven, Connecticut.
- Aquadro CF, Greenberg BD (1982) Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics*, **103**, 287-312.
- Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796-815.
- Arctander P (1995) Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proceedings of the Royal Society of London, B Series*, **262**, 13-19.
- Arctander P, Johansen C, Coutellec-Vreto, M-A (1999) Phylogeography of three closely related African bovids (Tribe Alcelaphini). *Molecular Biology and Evolution*, **16**, 1724-1739.
- Avise JC (1991) Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Annual Review of Genetics*, **25**, 45-69.
- Avise JC (1992) Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos*, **63**, 62-76.
- Avise JC (1998) The history and purview of phylogeography: a personal reflection. *Molecular Ecology*, **7**, 371-379.
- Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders SC (1987) Intraspecific phylogeography: the mitochondrial bridge between

- population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.
- Avise JC, Ball RM (1990) Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys of Evolutionary Biology*, **7**, 45-67.
- Avise JC, Lansman RA (1983) Polymorphism of mitochondrial DNA in populations of higher animals. In: *Evolution of Genes and Proteins*, (eds. Nei M, Koehn RK), pp. 147- 164. Sinauer, Massachusetts.
- Avise JC, Lansman RA, Shade RO (1979) The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. Population structure and evolution in the genus *Peromyscus*. *Genetics*, **92**, 279-295.
- Avise JC, Wollenburg K (1997) Phylogenetics and the origin of species. *Proceedings of the National Academy of Sciences of the USA*, **94**, 7748– 7755.
- Ball RM JR, Avise JC (1992) Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. *The Auk*, **109**, 626-636.
- Ball RM JR, Neigel JE, Avise JC (1990) Gene genealogies within the organismal pedigrees of random mating populations. *Evolution*, **44**, 1109-1119.
- Bandelt HJ, Forster P, Roehl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37-48.
- Baryshnikov CF (1988) The taxonomic position of *Mellivora* species of the USSR fauna Byulleten Moskovskogo Obshchestra Ispytatteley prirody. *Otdelenie Biologii*, **93**, 50-58.
- Baryshnikov CF (2000) A new subspecies of the honey badger *Mellivora capensis* from central Asia. *Acta Theriologica*, **45**, 45-55.
- Beckman KB, Smith MF, Orrego C (1993) Purification of mitochondrial DNA with Wizard™ minipreps DNA purification system. *Promega Notes Magazine*, **43**, 10-15.
- Begg CM (2001a) Feeding ecology and social organisation of honey badgers (*Mellivora capensis*) in the southern Kalahari. PhD Thesis, University of Pretoria.

- Begg CM, Begg KS, Du Toit JT, Mills MGL (2003) Sexual and season variation in the diet and foraging behaviour of a sexually dimorphic carnivore, the honey badger (*Mellivora capensis*). *Journal of Zoology, London* , **260**, 301-306.
- Begg CM, Begg KS, du Toit JT, Mills MGL (2005a) Spatial organization of the honey badger *Mellivora capensis* in the southern Kalahari: home-range size and movement patterns. *Journal of Zoology, London* , **265**, 17-22.
- Begg CM, Begg KS, du Toit JT, Mills MGL (2005b) Life-history variables of an atypical mustelid, the honey badger *Mellivora capensis*. *Journal of Zoology, London* , **265**, 17-22.
- Begg KS (2001b) Report on the conflict between beekeepers and honey badgers *Mellivora capensis* with reference to their conservation status and distribution in South Africa. *Unpublished report for the Endangered Wildlife Trust, Johannesburg*.
- Begg KS (2001c) Viscous Circles: gin traps in South Africa. *Africa Geographic*, **6**, 40-46.
- Ben-David M (1990) Phototrapping honey badgers. *Mustelid and Viverrid Conservation*, **3**, 14-15.
- Bensasson D, Zhang D-X, Hartl DL, Hewitt GM (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution*, **16**, 314-321.
- Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, **7**, 1513-1523.
- Bishop WW, Pickford MHL (1975) Geology, fauna and palaeoenvironments in the Ngorora Formation, Kenya Rift Valley. *Nature*, **254**, 185-192.
- Blanchard JL, Schmidt GW (1995) Pervasive migration of organellar DNA to the nucleus in plants. *Journal of Molecular Evolution*, **41**, 397-406.
- Boutin S (1990) Food supplementation experiments with terrestrial vertebrates: pattern, problems and the future. *Canadian Journal of Zoology*, **68**, 203-220.
- Brown WM (1983) Evolution of animal mitochondrial DNA. In: *Evolution of Genes and Proteins*, (eds. Nei M, Koehn RK), pp. 62- 88. Sinauer, Massachusetts.

- Brown WM, George M Jr, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the USA*, **76**, 1967-1971.
- Brusgaard K, Holm L-E, Lohi O (1998b) Two polymorphic mink (*Mustela vison*) dinucleotide repeat loci. *Animal Genetics*, **29**, 468-469.
- Brusgaard K, Malchenko SN, Christensen K, Lohi O, Kruse T (1998a) A polymorphic mink (*Mustela vison*) dinucleotide repeat. *Animal Genetics*, **29**, 467-468.
- Bryant HN, Russell AP, Fitch WD (1993) Phylogenetic relationships within the extant Mustelidae (Carnivora): appraisal of the cladistic status of the Simpsonian subfamilies. *Zoological Journal of the Linnean Society*, **108**, 301-334.
- Cantatore P, Saccone C (1987) Organization, structure, and the evolution of mammalian mitochondrial genes. *International Review of Cytology*, **108**, 149-208.
- Carroll RL (1988) *Vertebrate Paleontology and Evolution*. pp. 478-483. W.H. Freeman and Company, San Francisco.
- Cassens I, Tiedemann R, Suchentrunk F, Hartl GB (2000) Mitochondrial DNA variation in the European otter (*Lutra lutra*) and the use of spatial autocorrelation in conservation. *Journal of Heredity*, **91**, 31-35.
- Cavalcanti MJ (1988–2001) MANTEL- test for association between 2 symmetric distance matrices with permutation iterations. Setor de Paleovertebrados, Departamento de Geologia e Paleontologia, Museu Nacional do Rio de Janeiro, Quinta da Boa Vista, São Cristóvão, 20940-040, Rio de Janeiro.
- Chang DD, Clayton DA (1985) Priming of human mitochondrial DNA replication occurs at the light strand promoter. *Proceedings of the National Academy of Sciences of the USA*, **82**: 351-355.
- Cheesman RE (1920) Report on the mammals of Mesopotamia: collected by members of the Mesopotamian Expeditionary Force, 1915 to 1919. *Journal of the Bombay Natural History Society*, **27**, 323–336.
- Ciarrocchi G, Jose JG, Linn S (1979) Further characterization of a cell-free system for measuring replicative and repair DNA synthesis with cultured human fibroblasts

- and evidence for the involvement of DNA polymerase α in DNA repair. *Nucleic Acids Research*, **7**, 1205-1219.
- Coe MJ, Skinner JD (1993) Connections, disjunctions and endemism in the eastern and southern African mammal faunas. *Transcripts of the Royal Society of South Africa*, **48**, 233-255.
- Coetzee CG (1971) Order Carnivora. In: *The Mammals of Africa: An Identification Manual* (eds. Meester J, Setzer HW). Smithsonian Institution Press, Washington.
- Coetzee PW (1977) Present distribution and status of some of the mammals of Albany. Albany Divisional Council, Grahamstown.
- Collura RV, Stewart C-B (1995) Insertions and duplications of mtDNA in the nuclear genomes of Old World monkeys and hominoids. *Nature*, **378**, 485-489.
- Comrie-Grieg J (1985) Death trap for two honey badgers. *African Wildlife*, **39**, 119.
- Comstock KE, Georgiadis N, Pecon-Slattery J, Roca AL, Ostrander EA, O'Brien SJ, Wasser SK (2002) Patterns of molecular genetic variation among African elephant populations. *Molecular Ecology*, **11**, 2489-2498.
- Cracraft J (1983) Species concepts and speciation analysis. *Current Ornithology*, **1**, 159-187.
- Cracraft J (1997) Species concepts in systematics and conservation biology: an ornithological viewpoint. Pages 325-339. In: *Species, the units of biodiversity* (eds. Claridge MF, Dawah HA, Wilson MR), pp 325 - 339. Clarendon Press, London.
- Crandall KA, Bininda-Emonds ORB, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution*, **15**, 290-295.
- Cunningham AB, Zondi AS (1991) Use of animal parts for the commercial trade in traditional medicines. Institute of Natural Resources, University of Natal *Working Paper*, **76**, 42.
- Cuzin, F (1996) Répartition actuelle et statut des grands Mammifères sauvages du Maroc (Primates, Carnivores, Artiodactyles). *Mammalia*, **60**, 101-124

- Dallas JF (1992) Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mammalian Genome*, **5**, 32 – 38.
- Davis CS, Strobeck C (1998) Isolation, variability, and cross-species amplification of polymorphic microsatellite loci in the family Mustelidae. *Molecular Ecology*, **7**, 1776-1778.
- Davison A, Birks JDS, Brooks RC, Messenger JE, Griffiths HI (2001) Mitochondrial phylogeography and population history of pine martens *Martes martes* compared with polecats *Mustela putorius*. *Molecular Ecology*, **10**, 2479-2488
- Dawid IB, Blackler AW (1972) Maternal and cytoplasmic inheritance of mtDNA in *Xenopus*. *Developmental Biology*, **29**, 152-161.
- Degnan SM (1993) The perils of single gene trees – mitochondrial versus single-copy nuclear DNA variation in white-eyes (Aves: Zosteropidae). *Molecular Ecology*, **2**, 219-225.
- Delany MJ (1989) The zoogeography of the mammal fauna of southern Arabia. *Mammal Review*, **19**, 133-152
- Domingo-Roura X, MacDonald DW, Roy MS, Marmi J, Terradas J, Woodroffe R, Burke T, Wayne RK (2003) Confirmation of low genetic diversity and multiple breeding females in a social group of Eurasian badgers from microsatellite and field data. *Molecular Ecology*, **12**, 533-539.
- Doyle JJ (1992) Gene trees and species trees: Molecular systematics as one character taxonomy. *Systematic Botany*, **17**, 144-163.
- Dragesco-Joffe A. (1993) *La vie Sauvage au Sahara*. Lausanne, Switzerland. Delachaux et Niestle, Lausanne and Paris.
- Dragoo JW, Honeycutt RL (1997) Systematics of mustelid-like carnivores. *Journal of Mammalogy*, **78**, 426-443.
- Dubach J, Patterson BD, Briggs MB, Venzke K, Flamand J, Stander P, Scheepers L , Kays RW (2005) Molecular genetic variation across the southern and eastern geographic ranges of the African lion, *Panthera leo*. *Conservation Genetics*, **6**, 15-24.

- Duffy AJ, Landa A, O'Connell M, Stratton C, Wright JM (1998) Four polymorphic microsatellites in wolverine, *Gulo gulo*. *Animal Genetics*, **29**, 63-72.
- Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571-2581.
- Ebinger CJ (1989) Tectonic development of the western branch of the East African rift system. *Bulletin of the Geological Society of America*, **101**, 885-903.
- Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric tandem repeat loci in four human population groups. *Genomics*, **12**, 241-253.
- Eisenberg JF (1989) An introduction to the Carnivora. In: *Carnivore behavior, Ecology and Evolution* (ed. Gittleman JL), pp 1-9. Comstock Publishing Associates, Ithaca, New York.
- Eizirik E, Kim J-H, Menotti-Raymond M, Crawshaw JR PG, O'Brien SJ, Johnson WE (2001) Phylogeography, population history and conservation genetics of jaguars (*Panthera onca*, Mammalia, Felidae). *Molecular Ecology*, **10**, 65-79.
- Esposti MD, De Vries S, Crimi M, Ghelli A, Patarnello T, Meyer A (1993) Mitochondrial cytochrome b: evolution and structure of the protein. *Biochimica et Biophysica Acta*, **1143**, 243-271.
- Excoffier L, Schneider S (1999) Why hunter-gatherer populations do not show sign of Pleistocene demographic expansions. *Proceedings of the National Academy of Sciences of the USA*, **96**, 10597-10602.
- Excoffier L, Smouse PE, Quattro, JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-494.
- FairbanksDJ, Andersen WR (1999) *Genetics: The Continuity of Life*. Belmont, CA. Wadsworth Publishing Co.
- Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, **39**, 783-791.
- Finn F (1929) *Sterndale's Mammalia of India*. Thacker, Spink & Co., Calcutta, India.

- FitzSimmons NN (1998) Single paternity of clutches and sperm storage in the promiscuous green turtle (*Chelonia mydas*). *Molecular Ecology*, **7**, 575-584.
- Fitzsimons FW (1919) *The natural history of South Africa. Mammals*. Volume Two. Longmans, Green & Co, London.
- Freitag S, Robinson TJ (1993) Phylogeographic patterns in mitochondrial DNA of the ostrich (*Struthio camelus*). *The Auk*, **110**, 614-622.
- Freitag S, van Jaarsveld AS (1997) Relative occupancy, endemism, taxonomic distinctiveness and vulnerability: prioritizing regional conservation actions. *Biodiversity and Conservation*, **6**, 211-232.
- Friedlander M (1980) Monospermic fertilization in *Chrysopa carnia* (Neuroptera; Chrysopidae): Behaviour of the fertilizing spermatozoa prior to syngamy. *International Journal of Insect Morphology and Embryology*, **9**, 53-57.
- Friedmann Y, Daly B (2004). *Red data book of the Mammals of South Africa: A Conservation Assessment: CBSG Southern Africa, Conservation Breeding Specialist Group (SSC/IUCN)*. Endangered Wildlife Trust. South Africa.
- Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915-925.
- Fu Y-X, Li W-H (1993) Statistical tests of neutral mutations. *Genetics*, **133**, 693-709.
- Gascon C, Malcolm JR, Patton JL, da Silva MNF, Bogard JP, Loughheed SC, Peres CA, Neckel S, Boag PT (2000) Riverine barriers and the geographic distribution of Amazonian species. *Proceedings of the National Academy of Sciences of the USA*, **97**, 13672-13677.
- Gavin TA, Howard RA, May B (1991) Allozyme variation among breeding populations of red-winged blackbirds: the california conundrum. *The Auk*, **108**, 602-611.
- Geffen E, Anderson MJ, Wayne RK (2004) Climate and habitat barriers to dispersal in the highly mobile grey wolf. *Molecular Ecology*, **13**, 2481-2490.
- Geist V (1992) Endangered species and the law. *Nature*, **357**, 274-276
- Giles RE, Blanc H, Cann HM, Wallace DC (1980) Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences of the USA*, **77**, 6715-6719.

- Girman DJ, Vila C, Geffen E, Creel S, Mills MGL, McNutt JW, Ginsberg J, Kat PW, Mamiya KH, Wayne RK (2001) Patterns of population subdivision, gene flow and genetic variability in the African wild dog (*Lycaon pictus*) *Molecular Ecology*, **10**, 1703-1723.
- Goldstein PZ, DeSalle R, Amato G, Vogler AP (2000) Conservation genetics at the species boundary. *Conservation Biology*, **14**, 120-131.
- Gray MW (1999) Evolution of organellar genomes. *Current Opinion in Genetics and Development*, **9**, 678-687.
- Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. *Science*, **283**, 1476-1481.
- Gray MW, Burger G, Lang BF (2001) The origin and early evolution of the mitochondria. *Genome Biology 2: reviews*. 1018.1-1018.5.
- Gresson RAR (1940) Presence of the sperm middle-piece in the fertilized egg of the mouse (*Mus musculus*). *Nature* **145**: 425.
- Groves C (2004) The what, why and how of primate taxonomy. *International Journal of Primatology*, **25**, 1105- 1126.
- Groves C (2004) The what, why and how of primate taxonomy. *International Journal of Primatology*, **25**, 1105- 1126.
- Grubb P, Groves CP, Dudley JP, Shoshani J (2000) Living African elephants belong to two species: *Loxodonta africana* (Blumenbach, 1797) and *Loxodonta cyclotis* (Matschie, 1900). *Elephant*, **2**, 1-4.
- Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361-371.
- Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millassean P, Marc S, Bernardi G, Lanthrop M, Weissenbach J (1994) The 1993-94 Genethon human linkage map. *Nature Genetics*, **7**, 246-339.
- Gyllensten UB, Wharton D, Josefsson A, Wilson AC (1991) Paternal inheritance of mitochondrial DNA in mice. *Nature*, **352**, 255-257.
- Haig SM (1998) Molecular contributions to conservation. *Ecology*, **79**, 413-425.

- Hallgrímsson B, Maiorana V (2000) Variability and size in mammals and birds. *Biological Journal of the Linnean Society*, **70**, 571-595.
- Hamilton AC (1976) The significance of patterns of distribution shown by forest plants and animals in tropical Africa for the reconstruction of palaeoenvironments: a review. *Palaeoecology of Africa*, **9**, 63-97.
- Hamilton AC, Taylor D (1992) History of climate and forests in tropical Africa during the last 8 million years. In: *The History of Tropical Forests* (ed. Myers N), pp 65-78. Kluwer: Dordrecht.
- Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. *Current Anthropology*, **34**, 483-496.
- Harrison DL, Bates JJ (1991) *The mammals of Arabia*, second edition. Harrison Zoological Museum, Kent, United Kingdom.
- Hartl FU, Neupert W (1990) Protein sorting to mitochondria: evolutionary conservations of folding and assembly. *Science*, **247**, 930-938.
- Hendey QB (1978) Late Tertiary Mustelidae (Mammalia, Carnivora) from Langebaanweg, South Africa. *Annals of the South African Museum*, **76**, 329-357.
- Hepburn HR, Radloff SE (1988) *Honeybees of Africa* Springer, New York.
- Hewitt GM (2004) The structure of biodiversity- insights from molecular phylogeography. *Frontiers in Zoology*, **1**, 4.
- Hillis DM, Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, **42**, 182-192.
- Hoekstra HE, Krenz JG, Nachman MW (2005) Local adaptation in the rock pocket mouse (*Chaetodipus intermedius*): natural selection and phylogenetic history of populations. *Heredity*, **94**, 217-228.
- Hoelzel AR, Hancock JM, Dover GA (1991) Evolution of the cetacean mitochondrial d-loop region. *Molecular Biology and Evolution*, **8**, 475-493.
- Hoelzel AR, Lopez JV, Dover GA, O' Brian SJ (1994) Rapid evolution of a heteroplasmic repetitive segment in the mitochondrial DNA control region of carnivores. *Journal of Molecular Evolution*, **39**, 191-199.

- Hoffmann FG, Owen JG, Baker RJ (2003) mtDNA perspective of chromosomal diversification and hybridization in Peters' tent-making bat (*Uroderma bilobatum*: Phyllostomidae). *Molecular Ecology*, **12**, 2981-2993.
- Hollister N (1918) *East African mammals in the United States National Museum*. Part 1. Insectivora, Chiroptera, and Carnivora. Smithsonian Institution Bulletin, Washington, New York.
- Huelsenbeck JP, Hillis D (1993) Success of phylogenetic methods in the four taxon case. *Systematic Biology*, **42**, 247-264.
- Jarne P, Lagoda PLJ (1996) Microsatellites: from molecules to populations and back. *Trends in Ecology and Evolution*, **11**, 424-429.
- Johnson DDP, MacDonald DW, Dickman AJ (2000) An analysis and review of models of the socio-biology of the Mustelidae. *Mammal Review*, **30**, 171-196.
- Johnson WE, Eizirik E, Roelke-Parker M, O'Brien, SJ (2001) Applications of genetic concepts and molecular methods to carnivore conservation In: *Carnivore Conservation* (eds. Gittleman JL, Funk SM, MacDonald DW, Wayne RK). Cambridge University Press, London.
- Johnstone-Scott R (1975) Hand rearing a honey badger *Mellivora capensis* at Howletts Zoo Park, Bekesbourne. *International Zoo Yearbook*, **15**, 241-244.
- Johnstone-Scott R (1981) Notes on the management and breeding of the African ratel or Honey badger (*Mellivora capensis*). *Symposium of the Association of British Wild Animal Keepers*, **5**, 6-15.
- Kingdon (1977) *East African mammals: an atlas of evolution in Africa*. Volume three, part A. Academic press, London, United Kingdom.
- Kingdon J (1989) *East African mammals: Carnivores*, pp 232-233. University of Chicago Press, Chicago.
- Kingdon J (1997) *The Kingdon Field Guide to African Mammals*. Academic Press, San Diego, California.
- Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification

- and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the USA*, **86**, 6196-6200.
- Koepfli K-P, Wayne RK (1998) Phylogenetic relationships of otters (Carnivora: Mustelidae) based on mitochondrial cytochrome *b* sequences. *Journal of Zoology*, **246**, 401-416.
- Kondo R, Satta Y, Matsuura ET, Ishiwa H, Takahata N, Chigusa SI (1990) Incomplete maternal transmission of mitochondrial DNA in *Drosophila*. *Genetics*, **126**, 657-663.
- Kreader CA (1996) Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Applied and Environmental Microbiology*, **62**, 1102-1106.
- Kruuk H (1995) *Wild otters: predation and populations*. Oxford University Press, New York.
- Kruuk H, Mills MGL (1983) Notes on the food and foraging of the honey badger (*Mellivora capensis*) in the Kalahari Gemsbok National Park. *Koedoe*, **26**, 153-157.
- Kullback S, Leibler RA (1951) On information and sufficiency. *The Annals of Mathematical Statistics*, **22**, 79-86.
- Kurose N, Masuda R, Yoshida MC (1999) Phylogeographic variation in two mustelines, the least weasel *Mustela nivalis* and the ermine *M. erminea* of Japan, based on mitochondrial DNA control region sequences. *Zoological Science*, **16**, 971-977.
- Kvist L, Martens J, Nazarenko AA, Orell M (2003) Paternal leakage of mitochondrial DNA in the great tit (*Parus major*). *Molecular Biology and Evolution*, **20**, 243-247.
- Kyle CJ, Davis CS, Strobeck C (2000) Microsatellite analysis of North American pine marten (*Martes americana*) populations from the Yukon and Northwest territories. *Canadian Journal of Zoology*, **78**, 1150-1157.
- Kyle CJ, Robitaille JF, Strobeck C (2001) Genetic variation of fisher (*Martes pennanti*) populations across North America. *Molecular Ecology*, **10**, 2341-2347.

- Kyle CJ, Strobeck C (2001) Genetic structure of North American wolverine (*Gulo gulo*) populations. *Molecular Ecology*, **10**, 337-347.
- Kyle CJ, Strobeck C (2002) Connectivity of peripheral and core populations of North American wolverines. *Journal of Mammalogy*, **83**, 1141-1150.
- Kyle CJ, Strobeck C (2003) Genetic homogeneity of Canadian mainland marten populations underscores the distinctiveness of Newfoundland pine martens (*Martes americana atrata*). *Canadian Journal of Zoology*, **81**, 57-66.
- Kyle CJ, Weir RD, Newhouse NJ, Davis H, Strobeck C (2004) Genetic structure of sensitive and endangered northwestern badger populations (*Taxidea taxus taxus* and *T. t. jeffersonii*). *Journal of Mammalogy*, **85**, 633-639.
- Langella, O (2000) POPULATIONS 1.2.28. Accessed from <http://www.cnrs-gif.fr/pge>, December 2004.
- Ledje C, Arnason U (1996) Phylogenetic relationships within caniform carnivores based on analysis of the mitochondrial 16S rRNA gene. *Journal of Molecular Evolution*, **43**, 641-649.
- Lindstedt SL, Miller, BJ, Buskirk, SW (1986) Home range, time and body size of mammals. *Ecology*, **67**, 413-418.
- Lodé T (2001) Genetic divergence without special isolation in polecat *Mustela putorius* populations. *Journal of Evolutionary Biology*, **14**, 228-236.
- Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien S J (1994) *Numt*, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *Journal of Molecular Evolution*, **39**, 174-190
- Lundstrom R, Tavaré S, Ward RH (1992) Modelling the evolution of the human mitochondrial genome. *Mathematical Biosciences*, **112**, 319-335.
- Lunt DH, Whipple LE, Hyman BC (1998) Mitochondrial DNA variable number tandem repeats (VNTRs): utility and problems in molecular ecology. *Molecular Ecology*, **7**, 1440-1455.
- Maddison DR, Ruvolo M, Swofford DL (1992) Geographic origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. *Systematic Biology*, **41**, 111-124.

- Maddison WP (1997) Gene trees in species trees. *Systematic Biology*, **46**, 523-536.
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209-220.
- Margulis, L (1970) *Origin of Eucaryotic Cells*. Yale University Press, New Haven.
- Marmi J, López-Giráldez JF, Domingo-Roura X (2004) Phylogeny, evolutionary history and taxonomy of the Mustelidae based on sequences of the cytochrome b gene and a complex repetitive flanking region. *Zoologica Scripta*, **33**, 481-499.
- Marshall TC, Slate J, Kruuk L, Pemberton JM (1998) Statistical confidence for likelihood-based paternity in natural populations. *Molecular Ecology*, **7**, 639-655.
- Martin LD (1989) Fossil history of the terrestrial Carnivora. In: *Carnivore Behaviour, Ecology and Evolution* (ed. Gittleman JL), pp. 536-568. Cornell University Press, Ithaca, New York.
- Master LL (1991) Assessing threats and setting priorities for conservation. *Conservation Biology*, **5**, 559-563.
- Matthee CA, Robinson TJ (1997) Mitochondrial DNA phylogeography and comparative cytogenetics of the springhare, *Pedetes capensis* (Mammalia: Rodentia). *Journal of Mammalian Evolution*, **4**, 53-73.
- Matthee CA, Tilbury CR, Townsend T (2004) A phylogenetic review of the African leaf chameleons: genus *Rhampholeon* (Chamaeleonidae): the role of vicariance and climate change in speciation. *Proceedings of the Royal Society of London, B Series*, **271**, 1967-1975.
- Mayr E (1940) Speciation phenomena in birds. *American Naturalist*, **74**, 249-278.
- Mayr E (1963) *Animal species and evolution*. McGraw-Hill, New York.
- Mayr E, Ashlock PB (1991) *Principles of systematic zoology*, second edition. McGraw-Hill, New York.
- Milinkovitch MC, Lyons-Weilert J (1998) Finding optimal ingroup topologies and convexities when the choice of outgroups is not obvious. *Molecular Phylogenetics and Evolution*, **9**, 348-357.

- Mirol PM, Mascheretti S, Searle JB (2000) Multiple nuclear pseudogenes of mitochondrial cytochrome *b* in *Ctenomys* (Caviomorpha, Rodentia) with either great similarity to or high divergence from the true mitochondrial sequence. *Heredity*, **84**, 538-547.
- Monadjem A (1998) *The mammals of Swaziland*. Conservation Trust of Swaziland and Big Game Parks, Swaziland.
- Moore AE, Larkin PA (2001) Drainage evolution in south-central Africa since the break up of Gondwana. *South African Journal of Geology*, **103**, 47-68.
- Moore SS, Sargeant LL, King TJ, Mattick JS, Georges M, Hetzel DJS (1991) The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics*, **10**, 654 – 660.
- Moritz C (1994) Defining ‘Evolutionary Significant Units’ for conservation. *Trends in Ecology and Evolution*, **9**, 373-375.
- Mourier T, Hansen AJ, Willerslev E, Arctander P (2001) The human genome project reveals a continuous transfer of large mitochondrial fragments to the nucleus. *Molecular Biology and Evolution*, **18**, 1833-1837.
- Muwanika VB, Nyakaana S, Siegismund HR, Arctander P (2003) Phylogeography and population structure of the common warthog (*Phacochoerus africanus*) inferred from variation in mitochondrial DNA sequences and microsatellite loci. *Heredity*, **91**, 361-372.
- Myers N (1988) Threatened biotas: ‘Hotspots’ in tropical forests. *Environmentalists*, **8**, 1-20
- Nachman MW, Boyer SN, Searle JB, Aquadro CF (1994) Mitochondrial DNA variation and the evolution of robertsonian chromosomal races of house mice, *Mus domesticus*. *Genetics*, **136**, 1105-1120.
- Neal E, Cheeseman C (1996) *Badgers*. T & D Poyster Ltd, London.
- Nei M (1972) Genetic distance between populations. *American Naturalist*, **106**, 283-291.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

- Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1-10.
- Nei M, Roychoudhary AK (1974) Sampling variances of heterozygosity and genetic distance. *Genetics*, **76**, 379-390.
- Nowak RM (1999) *Walker's Mammals of the World*, volume 2, sixth edition, pp1005-1006. The Johns Hopkins University Press, Baltimore and London.
- O' Connell M, Wright JM, Farid A (1996) Development of PCR primers for nine polymorphic american mink *Mustela vison* microsatellite loci. *Molecular Ecology*, **5**, 311-312.
- O' Brien SJ, Mayr E (1991) Bureaucratic mischief: Recognising endangered species and subspecies. *Science*, **227**, 1428- 1434.
- Olson LE, Hassanin A (2003) Contamination and chimerism are perpetuating the legend of the snake-eating cow with twisted horns (*Pseudonovibos spiralis*). A case study of the pitfalls of ancient DNA. *Molecular Phylogenetics and Evolution*, **27**, 545-548.
- Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences*, **12**, 357-358.
- Palva TK, Palva ET (1985) Rapid isolation of animal mitochondrial DNA by alkaline extraction. *FEBS Letters*, **192**, 267-270.
- Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution*, **5**, 568-583.
- Partridge TC, Wood BA, de Menocal PB (1995) The influence of global climate change and regional uplift on large-mammalian evolution in east and southern Africa. In: *Paleoclimate and evolution with emphasis on human origins* (eds. Vrba ES, Denton GH, Partridge TC, Burckle LH), pp 331-355. Yale University Press, London.
- Pitra C, Hansen A, Lieckfeldt D, Arctander P (2002) An exceptional case of historical outbreeding in African sable antelope populations. *Molecular Ecology*, **11**, 1197-1208.

- Pocock RI (1946) External and cranial characters of some rare Asiatic mammals recently exhibited by the society. *Proceedings of the Zoological Society of London*, **115**, 310-318.
- Pocock RL (1921) On the external characters and classification of the Mustelidae. *Proceedings of the Zoological Society of London*, **1921**, 803- 837.
- Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology*, **53**, 793-808.
- Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817-818.
- Przeworski M (2002) The signature of positive selection at randomly chosen loci. *Genetics*, **160**, 1179–1189.
- Quinn TW (1992) The genetic legacy of mother goose- phylogeographic patterns of lesser snow goose *Chen caerulescens caerulescens* maternal lineages. *Molecular Ecology*, **1**, 105-117.
- Rand DM (1993) Endotherms, ectotherms, and mitochondrial genome-size variation. *Journal of Molecular Evolution*, **37**, 281-295.
- Ray N, Currat M, Excoffier L (2003) Intra-Deme Molecular Diversity in Spatially Expanding Populations. *Molecular Biology and Evolution*, **20**, 76-86.
- Raymond M, Rousset F (1995) GENEPOP version 1.2: population genetic software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248-249.
- Richly E, Leister D (2004) NUMTs in Sequenced Eukaryotic Genomes. *Molecular Biology and Evolution*, **21**, 1081-1084.
- Robinson TJ, Matthee (1999) Molecular genetic relationships of the extinct ostrich, *Struthio camelus syriacus*: consequences for ostrich introductions into Saudi Arabia. *Animal Conservation*, **2**, 165-171.
- Roca AL, Georgiadis N, O'Brien SJ (2005) Cytonuclear genomic dissociation in African elephant species. *Nature Genetics*, **37**, 96-100.
- Roca AL, Georgiadis N, Pecon-Slattery J, O'Brien SJ (2001) Genetic evidence for two species of elephant in Africa. *Science*, **293**, 1473-1477.

- Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pair-wise genetic differences. *Molecular Biology and Evolution*, **9**, 552-569.
- Rosevear DR (1974) *The carnivores of West Africa*. The British Museum, London.
- Rowe-Rowe DT (1992) The carnivores of Natal. *The Lammergeyer*, **25**, 1-47.
- Rozas J, Rozas R (1999) DnaSP, Version3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174-175.
- Ryder OE (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution*, **1**, 9-10.
- Ryder OE, Shaw JH, Wemmer CM (1988) Species, subspecies and *ex situ* conservation. *International Zoo Yearbook*, **27**, 134-140.
- Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487-491.
- Saitou N, Nei M (1987) The neighbour-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406-425.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: A laboratory manual*, second edition. Cold Spring Harbour Laboratory Press, New York.
- Sandell M (1989) The mating tactics and spacing patterns of solitary carnivores pp 164-182. In: *Carnivore Behaviour, Ecology and Evolution* (ed. Gittleman JL), pp 164-182. Cornell University Press, Ithaca, New York.
- Schlötterer C, Amos B, Tautz D (1993) Conservation of polymorphic simple sequences in cetacean species. *Nature*, **354**, 63-65.
- Schmidt-Kittler N (1981) Zur Stammesgeschichte der marderverwandten Raubtiergruppen (Musteloidea, Carnivora). *Eclogae Geologicae Helveticae*, **74**, 753-801.
- Schneider S, Roessli D, Excoffier L (2000) Arlequin: A software for population genetics data. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

- Schreber JCD (1776) *Die säugethiere in Abbildungen nach der Natur, mit Beschreibungen*. Wolfgang Walther, Erlangen, Germany. (Not seen, cited in Vanderlaar & Ten Hwang 2003).
- Sillero-Zubiri C (1996) Records of honey badger *Mellivora capensis* (Carnivora, Mustelidae) in afroalpine habitat, above 4000m. *Mammalia*, **60**, 323-325
- Simonsen BT, Siegismund HR, Arctander P (1998) Population structure of African buffalo inferred from mtDNA sequences and microsatellite loci: high variation but low differentiation. *Molecular Ecology*, **7**, 225-237.
- Simonsen KL, Churchill GA, Aquadro CF (1995) Properties of Statistical Tests of Neutrality for DNA Polymorphism Data. *Genetics*, **141**, 413-429
- Simpson GC (1945) The principles of classification and a classification of the mammals. *Bulletin of the American Museum of Natural History*, **85**, 1-350.
- Skinner JD, Smithers RHN (1990) *The mammals of the southern African subregion*. University of Pretoria, Pretoria.
- Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787-792.
- Slowinski JB, Page RDM (1999). How should species phylogenies be inferred from sequence data. *Systematic Biology*, **48**, 814-825.
- Small MP, Stone KD, Cook JA (2003) American marten (*Martes americana*) in the Pacific Northwest: population differentiation across a landscape fragmented in time and space. *Molecular Ecology*, **12**, 89-103.
- Smithers, RHN (1983) *The mammals of the southern African subregion*, pp 435-438. University of Pretoria, South Africa.
- Smithers, RHN (1986) The South African Red Data Book- Terrestrial Mammals. *South African National Scientists Progress Report*, **125**, 75-76.
- Sorenson MD, Quinn TW (1998) Numts: A challenge for avian systematics and population biology. *The Auk*, **115**, 214-221.
- Steel DJ, Trewick SA, Wallis GP (2000) Heteroplasmy of mitochondrial DNA in the ophiuroid *Astrobrachion constrictum*. *Journal of Heredity*, **91**, 146-149.

- Stewart DT, Baker AJ (1994) Patterns of sequence variation in the mitochondrial D-loop region of shrews. *Molecular Biology and Evolution*, **11**, 9-21.
- Storr GCC (1780) *Prodromus methodi mammalium*. Litteris Reissianis, Tubinae, Germany. (Not seen, cited in Vanderlaar & Ten Hwang 2003).
- Stuart CT (1981) Notes on the mammalian carnivores of the Cape Province, South Africa. *Bontebok*, **1**, 1-58.
- Stuart CT (1990) The conservation status of mustelids and viverrids in Southern Africa. *Mustelid and Viverrid Conservation*, **3**, 16.
- Swofford DL (2002) PAUP*: *Phylogenetic Analysis Using Parsimony (*and Other Methods*. Version 4.0 beta 10. Sinauer Associates, Sunderland, MA.
- Tajima F (1989) The effect of change in population size on DNA polymorphism. *Genetics*, **123**, 597-601.
- Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512-526.
- Tautz D (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, **17**, 6463-6471.
- Taylor DE, Brose EC (1988) Modified Birnboim-Doly method for rapid detection of plasmid copy number. *Nucleic Acids Research*, **16**, 9056.
- Taylor DJ, Sprenger HL, Ishida S (2005) Geographic and phylogenetic evidence for dispersed nuclear introgression in a daphniid with sexual propagules. *Molecular Ecology*, **14**, 525-537.
- Tchernov E (1989) The paleobiogeographical history of the southern Levant. In: *The zoogeography of Israel. The distribution and abundance at a zoogeographical crossroad* (eds. Yom-Tov Y, Tchernov E), pp 159-250. Dordrecht: Monographic Biologicae.
- Tedford RH (1976). Relationship of pinnipeds to other carnivores (Mammalia). *Systematic Zoology*, **25**, 363-374.
- Templeton AR (1998) Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381-397.

- Thalmann O, Hebler J, Poinar HN, Pääbo S, Vigilant, L (2004) Unreliable mtDNA data due to nuclear insertions: a cautionary tale from analysis of humans and other great apes. *Molecular Ecology*, **13**, 321-335.
- Toth G, Gaspari Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Research*, **10**, 967-981.
- Upholt WB, Dawid IB (1977) Mapping of mitochondrial DNA of individual sheep and goats: rapid evolution in the D-loop region. *Cell*, **11**, 571-583.
- Uphyrkina O, Johnson WE, Quigley H, Miquelle D, Marker L, Bush M, O' Brien SJ (2001) Phylogenetics, genome diversity and origin of modern leopard, *Panthera pardus*. *Molecular Ecology*, **10**, 2617-2633.
- Ursprung H, Schabtach E (1965) Fertilization in tunicates: loss of the paternal mitochondrion prior to sperm entry. *Journal of Experimental Zoology*, **159**, 379-383.
- Vanderhaar JM, Ten Hwang Y (2003) *Mellivora capensis*. *Mammalian Species*, **721**, 1-8.
- Verwey R, Begg C, Begg K, Matthee CA (2004) A microsatellite perspective on the reproductive success of subordinate male honey badgers, *Mellivora capensis*. *African Zoology*, **39**, 305-308.
- Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. *Science*, **253**, 1503-1507.
- Vogler AP, DeSalle R (1994) Diagnosing units of conservation management. *Conservation Biology*, **8**, 354-363.
- Walberg MW, Clayton DA (1981) Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. *Nucleic Acids Research*, **9**, 5411-5421.
- Walker CW, Vilà C, Landa A, Lindén M, Ellegren H. (2001) Genetic variation and population genetic structure in Scandinavian wolverine (*Gulo gulo*) populations. *Molecular Ecology*, **10**, 53-63.

- Waples RS (1991) Pacific salmon, *Oncorhynchus* spp., and the definition of "species" under the Endangered Species Act. *Marine Fisheries Review*, 53, 11-22.
- Ward RH, Frazier BL, Dew-Jager K, Pääbo S (1991) Extensive mitochondrial diversity within a single Amerindian Tribe. *Proceedings of the National Academy of Sciences USA*, **88**, 8720-8724.
- Weber JL (1990) Informativeness of human (dC-dA)_n. (dG-dT)_n polymorphisms. *Genomics*, **7**, 524-530.
- Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44**, 388-396.
- Weber JL, Wang C (1993) Mutation of human short tandem repeats. *Human Molecular Genetics*, **2**, 1123-1128.
- Welter C, Dooley S, Blin N (1989) A rapid protocol for the purification of mitochondrial DNA suitable for studying restriction fragment length polymorphisms. *Gene*, **83**, 169-172.
- Wheeler WC (1990) Nucleic acid sequence phylogeny and random outgroups. *Cladistics*, **6**, 363-367.
- Williams ST, Knowlton N (2001) Mitochondrial pseudogenes are pervasive and often insidious in the snapping shrimp genus *Alpheus*. *Molecular Biology and Evolution*, **18**, 1484-1493.
- Wozencraft WC (1989) The phylogeny of the Recent Carnivora. In: *Carnivore behaviour, ecology and evolution* (ed. Gittleman JL), pp 495-535. Cornell University Press, Ithaca, New York.
- Wozencraft WC (1993) Order Carnivora. In: *Mammal Species of the World: a Taxonomic and Geographic Reference* (eds. Wilson DE, Reeder DA), pp. 279-348. Smithsonian Institution Press, Washington.
- Wyss AR, Flynn JJ (1993) A phylogenetic analysis and definition of the Carnivora. In: *Mammal Phylogeny: placentals* (eds. Szalay FS, Novacek MJ, McKenna MC), pp 32-52. Springer-Verlag, New York.

- Zhang D-X, Hewitt GM (1996a) Nuclear integrations: challenges for mitochondrial DNA markers. *Trends in Ecology and Evolution*, **11**, 247-251.
- Zhang D-X, Hewitt GM (1996b) Highly conserved nuclear copies of the mitochondrial control region in the desert locust *Schistocerca gregaria*: some implications for population studies. *Molecular Ecology*, **5**, 295-300.
- Zhang D-X, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **12**, 563-584.
- Zischler H, Geisert H, von Haeseler A, Pääbo S (1995) A nuclear 'fossil' of the mitochondrial D-loop and the origin of modern humans. *Nature*, **378**, 489-492.
- Zullo S, Sieu LC, Slighton JL, Hadler HI, Eisenstadt JM (1991) Mitochondrial D- loop sequences are integrated into the rat nuclear genome. *Journal of Molecular Biology*, **221**, 1223-1235.

Appendix

Complete list of samples and their sampling locality used in this study. Also indicated is the population to which the sample was allocated during analysis

		Assigned Population	Assigned Population	Data obtained:	Data obtained:	
Sample Name	Collection Locality	Mitochondrial Data	Microsatellite Data	Mitochondrial	Microsatellite	Collected by/ Reference
CA1	Hermanus, Western Cape Province, RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA2	Velddrift, Western Cape Province, RSA	CAPE	CAPE	X	X	J. Isham
CA3	R27 near Ganzekraal , Western Cape Province, RSA	CAPE	CAPE	X	X	J. Isham
CA4	Pearly Beach, Western Cape Province, RSA	CAPE	NOT APPLICABLE	X	NOT TYPED	Les Friebus
CA5	Albertina, Western Cape Province, RSA	CAPE	NOT APPLICABLE	X		De Wildt Cheetah and Wild Life Centre
CA6	Viljoenshof area, Western Cape Province , RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA7	Wolvengat, Western Cape Province, RSA	CAPE	NOT APPLICABLE	X	NOT TYPED	Keith and Colleen Begg
CA8	Wolvengat, Western Cape Province, RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA9	Wolvengat, Western Cape Province, RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA10	Wolvengat, Western Cape Province, RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA11	Wolvengat, Western Cape Province, RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA12	Wolvengat, Western Cape Province, RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA13	Wolvengat, Western Cape Province, RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA14	Wolvengat, Western Cape Province, RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA15	Caledon, Western Cape Province, RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA16	Goukamma Reserve, Western Cape Province, RSA	CAPE	CAPE	X	X	Keith and Colleen Begg
CA17	Paarl, Western Cape Province, RSA	CAPE	CAPE	X	X	Conrad Matthee
KAL1	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL3	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL5	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL6	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL7	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL8	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg

Appendix continued

KAL9	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	NOT APPLICABLE	X	NOT TYPED	Keith and Colleen Begg
KAL10	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL11	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL13	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL14	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL15	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL16	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL17	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
KAL18	Kgalagadi Transfrontier National Park, RSA/ Botswana	KALA	KALA	X	X	Keith and Colleen Begg
SM1	Kruger National Park, RSA	SMOZ	NOT APPLICABLE	X	NOT TYPED	De Wildt cheetah and wild life centre
SM2	Kruger National Park, RSA	SMOZ	SMOZ	X	X	Conrad Matthee
SM3	Limpopo Province, RSA	SMOZ	NOT APPLICABLE	X	NOT TYPED	De Wildt cheetah and wild life centre
SM4	Maputo province, Mozambique	SMOZ	SMOZ	X	X	Keith and Colleen Begg
SM5	Inhambane, central Mozambique	NOT APPLICABLE	SMOZ	NOT SEQUENCED	X	Keith and Colleen Begg
SM6	Lamahasha, Maputo province, Mozambique	SMOZ	SMOZ	X	X	Keith and Colleen Begg
SM7	Maputo, Mozambique	SMOZ	SMOZ	X	X	Keith and Colleen Begg
SM8	Maputo, Mozambique	SMOZ	SMOZ	X	X	Keith and Colleen Begg
SM9	Ruwa, Harare, Zimbabwe	SMOZ	SMOZ	X	X	Bulawayo Museum
SM10	Zinave, Zimbabwe	SMOZ	SMOZ	X	NOT TYPED	Bulawayo Museum
NM1	Niassa Reserve, Mozambique	NMOZ	NMOZ	X	X	Keith and Colleen Begg
NM2	Niassa Reserve, Mozambique	NMOZ	NMOZ	X	X	Keith and Colleen Begg
NM3	Niassa Reserve, Mozambique	NMOZ	NMOZ	X	X	Keith and Colleen Begg
NM4	Niassa Reserve, Mozambique	NMOZ	NMOZ	X	X	Keith and Colleen Begg
NM5	Lichinga, Mozambique	NMOZ	NMOZ	X	X	Keith and Colleen Begg
NM6	Niassa province, near Lichinga, Mozambique	NOT APPLICABLE	NMOZ	NOT SEQUENCED	X	Keith and Colleen Begg

Appendix continued

NM7	Niassa Reserve, Mozambique	NMOZ	NMOZ	X	X	Keith and Colleen Begg
NWZ1	Mfumbuwe, Zambia	NWZA	NWZA	X	X	Keith and Colleen Begg
NWZ2	Mfumbuwe, Zambia	NOT APPLICABLE	NWZA	NOT SEQUENCED	X	Keith and Colleen Begg
NWZ3	Near Manyinga, Zambia	NWZA	NWZA	X	X	Keith and Colleen Begg
NWZ4	Katala, Zambia	NWZA	NWZA	X	X	Keith and Colleen Begg
NWZ5	Mukundwiji River, N.W Province, Zambia	NWZA	NWZA	X	X	Keith and Colleen Begg
NWZ6	Close to Angola border, N.W Province, Zambia	NOT APPLICABLE	NWZA	NOT SEQUENCED	X	Keith and Colleen Begg
NWZ7	Close to Angola border, N.W Province, Zambia	NWZA	NWZA	X	X	Keith and Colleen Begg
NWZ8	Sesheka District, Zambia	NOT APPLICABLE	NWZA	NOT SEQUENCED	X	Amatole Museum
NWZ9	Zambezi (Balovale), Zambia	NWZA	NWZA	X	X	Amatole Museum
NWZ10	Western Province, Zambia	NWZA	NWZA	X	X	Bulawayo Museum
NWZ11	Kalichero, Zambia	NOT APPLICABLE	NWZA	NOT SEQUENCED	X	Bulawayo Museum
KEN1	Segeera ranch, Kenya	KENY	NOT APPLICABLE	X	NOT TYPED	Keith and Colleen Begg
KEN2	Tsavo National Park, Kenya	KENY	NOT APPLICABLE	X	NOT TYPED	Nairobi Museum
KEN3	Karen, Nairobi, Kenya	KENY	NOT APPLICABLE	X	NOT TYPED	Nairobi Museum
KEN4	Karen, Ngong Hills, Kenya	KENY	NOT APPLICABLE	X	NOT TYPED	Nairobi Museum
KEN5	Molo, Rift Valley, Kenya	KENY	NOT APPLICABLE	X	NOT TYPED	Nairobi Museum
KEN6	Londiani, Rift Valley, Kenya	KENY	NOT APPLICABLE	X	NOT TYPED	Nairobi Museum
KEN7	Central Kenya	KENY	NOT APPLICABLE	X	NOT TYPED	Nairobi Museum
KEN8	Nairobi, Kenya	KENY	NOT APPLICABLE	X	NOT TYPED	Nairobi Museum
KEN9	Kenya (no other info)	KENY	NOT APPLICABLE	X	NOT TYPED	Nairobi Museum
KEN10	Nairobi, Kenya	KENY	NOT APPLICABLE	X	NOT TYPED	Nairobi Museum
TANZ	Lukwati, Tanzania	UNNASSIGNED	NMOZ	X	X	Sun Tissue Bank
ETH	Bale Mountains, Ethiopia	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	Karen Laurenson
ZIM1	Valindre Farm near Bulawayo, Zimbabwe	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	Bulawayo Museum
ZIM2	Ruwa Harare, Zimbabwe	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	Bulawayo Museum

Appendix continued

ZIM3	Kwe Kwe, Zimbabwe	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	Bulawayo Museum
BOT1	Odiakwe, Botswana	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	Bulawayo Museum
NAM1	Grootfontein (Nuitsas), Namibia	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	Amatole Museum
RSA1	East London, Western Cape Province, RSA	UNNASSIGNED	CAPE	X	X	Amatole Museum
RSA2	Northern Cape Province, Namaqualand, RSA	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	Amatole Museum
DRC	Ituri Forest, Democratic Republic of the Congo	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	Nairobi Museum
CHAD	Zakouma National Park, Chad	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	C. Denys
ISR1	Israel	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	Yoram Yom- Tov
ISR2	Israel	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	Yoram Yom- Tov
IND	Bori Province, India	UNNASSIGNED	NOT APPLICABLE	X	NOT TYPED	London Museum