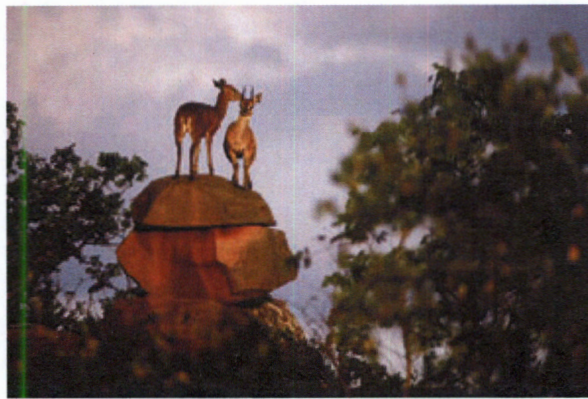

Phylogeographic Analysis Reveals Strong Geographical Structuring in the Klipspringer, *Oreotragus oreotragus*.



NIKKI LE ROEX

Thesis submitted in fulfillment of the requirements for the degree of Master of Science (Molecular and Cell Biology) at the University of Cape Town.

SUPERVISOR: Dr Colleen O’Ryan

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Contents

DECLARATION

ACKNOWLEDGEMENTS

ABSTRACT

Chapter 1: INTRODUCTION

1.1	<i>Oreotragus oreotragus</i>	1
1.2	Subspecies designations within <i>Oreotragus</i>	4
1.3	Conservation genetics.....	5
1.4	Mitochondrial DNA.....	9
1.5	Nuclear integrations.....	11
1.6	Sample types.....	13
1.7	Research objectives.....	15

Chapter 2: MATERIALS AND METHODS

2.1	Sample collection.....	16
2.2	DNA extraction.....	18
2.3	DNA amplification.....	19
2.3.1	Primerless PCR.....	19
2.3.2	The cytochrome <i>b</i> fragment.....	20
2.3.3	The control region fragment.....	20
2.4	PCR purification and sequencing.....	25
2.5	Data analysis.....	26
2.5.1	Sequence identity confirmation.....	26
2.5.2	Identification of mtDNA control region fragment.....	26
2.5.3	Genetic variation and phylogenetic analysis.....	27
2.5.4	Population structure.....	29

Chapter 3: RESULTS

3.1	Sequence identity confirmation.....	30
-----	-------------------------------------	----

3.2 Identification of authentic mtDNA control region sequences.....	30
3.2.1 Sequence types.....	30
3.2.2 Haplotype diversity and phylogenetic relationships.....	32
3.2.3 Identification of recombinants.....	35
3.3 Authentic mtDNA analyses.....	36
3.3.1 The cytochrome <i>b</i> fragment.....	36
3.3.2 The control region fragment.....	39
(i) Genetic variation and phylogenetic analysis.....	39
(ii) Population structure.....	45
3.3.3 Subset analysis: <i>Oreotragus oreotragus</i>	47
(i) Genetic variation and phylogenetic analysis.....	47
(ii) Population structure.....	48

Chapter 4: DISCUSSION

4.1 Identification of authentic mtDNA control region sequences.....	50
4.1.1 Identification of the Numt sequences.....	50
4.1.2 Identification of PCR recombinants.....	51
4.1.3 Numts, recombinants and sequence databases.....	52
4.2 Mitochondrial DNA analysis.....	53
4.2.1 Genetic variation.....	53
4.2.2 Phylogeography.....	55
4.2.3 Population differentiation and subspecies designations.....	58
4.2.4 Implications for conservation.....	60
4.2.5 Conclusion.....	62

REFERENCES.....	63
------------------------	-----------

APPENDICES

Appendix 1.....	I
Appendix 2.....	III
Appendix 3.....	V

Acknowledgements

I would like to thank a number of people who contributed to this study. In particular, to my supervisor, Dr Colleen O’Ryan, for her guidance and advice, to Dr Jacqueline Bishop and Dr Woody Cotterill for their input and assistance, and to everyone else in lab 405 for their unfailing support.

I would like to extend a special thank you to Professor Paul van Helden for taking the time and interest to assist in my quest for samples, and to Dr Terry Robinson for so generously providing me with klipspringer samples of his own. Without their help, this study would not have been possible. I would also like to thank Denise Hammond (Iziko Museum, Cape Town) for the provision of additional samples.

Finally, I would like to thank my family and friends for their support throughout this time. Particulary Nic, who motivated me to work so much harder.

Abstract

The klipspringer, *Oreotragus oreotragus*, occurs from the Western Cape of South Africa through to north-east Africa. Anthropological factors such as hunting have eradicated the klipspringer in parts of its former range and efforts have begun to reintroduce klipspringer back into these areas. This has highlighted the need for the proper understanding of the genetic variation, phylogeographic and population structure of the species, as well as the genetic validity of the 11 subspecies designations. Mitochondrial control region and cytochrome *b* sequence data were used to investigate these topics using blood, faecal and museum specimens. Cytochrome *b* and control region fragments were amplified in 83 and 60 samples respectively, out of a total sample set of 89 individuals. The generation of authentic mitochondrial DNA control region fragments proved difficult, with both Numt and PCR recombinant sequences identified after amplification. High levels of genetic variation were observed in the klipspringer, with cytochrome *b* and control region haplotype diversities of 0.78 and 0.98 respectively. Phylogenetic and network analyses showed the distinct geographical clustering of individuals into two major groups, south/south-western (S/SW) and east/north-eastern (E/NE), with the S/SW further split into two divergent groups. This suggests that the two S/SW groups were separated and isolated as a result of shifting climatic conditions in the late Pliocene/early Pleistocene, after which secondary contact occurred and klipspringer moved upwards from southern Africa to colonise the east. The results from this study have conservation implications with respect to translocation policies, suggesting that translocations should only occur within the geographic groups identified in order to avoid outbreeding depression.

Chapter 1

Introduction

1.1 *Oreotragus oreotragus*

The klipspringer, *Oreotragus oreotragus* (Zimmerman 1783), has a wide distribution range that extends from the Western Cape region of South Africa to north-eastern Africa, although within this range their distribution is fragmented and irregular (Figure 1). This fragmented distribution is a result of specific habitat requirements, with suitable habitat types including mountainous areas, rocky hills or outcrops, and gorges with rocky sides (Wilson & Child 1965). Taxonomically, the klipspringer belongs to the family Bovidae (Mammalia, Artiodactyla, Ruminantia) which is comprised of 137 extant species, that make up 45 genera. Within Bovidae, the klipspringer is suggested to belong to the tribe Neotragini (dwarf antelope) and to be most closely related to the duiker (*Cephalophus spp.*), although the tribes within Bovidae, as well as their phylogenetic affinities, have been the subject of much debate (Hassanin & Douzery 1999, Matthee & Robinson 1999). The klipspringer's diet consists predominantly of leaves, fruit, berries, seedpods and flowers, and although they are water-independent browsers, they will drink water if it is available in the surrounding area (Skinner & Chimimba 2005).

Klipspringers are normally seen in pairs, but can also be found as solitary individuals, family groups, or even as larger congregations of up to six individuals. These larger congregations

are thought to be transient, as if they are disturbed the group dissolves rapidly (Skinner & Chimimba 2005). Klipspringer are also highly territorial, with both males and females defending their territory through visual displays of dominance, scent marking with pre-orbital glands, and by the creation of dung heaps (Dunbar & Dunbar 1974).



Figure 1: Distribution map of *Oreotragus oreotragus* (Institute of Applied Ecology 1998, <http://gorilla.bio.uniroma1.it/amd/amd184b.html>). Green shaded areas represent the areas inhabited by klipspringer.

Klipspringers are preyed on by a variety of predators such as leopards, hyenas, baboons, eagles and caracals and when threatened they rely on loud, repetitive alarm calls and their extraordinary ability to negotiate the rocky terrain to evade their predators (Tilson & Norton 1981). Breeding seasons appear to vary according to local conditions as lambs can be seen at

most times of the year. although much debate surrounds the length of their gestation period (Norton 1980). Generally a single lamb, of approximately 1 kg, is born in the safety of rocks or heavy vegetation (Wilson & Child 1965). The lambs rely almost entirely on camouflage for protection during the first few months of their lives, and when approached, merely flatten themselves against the ground in an attempt to blend into the landscape (Norton 1980).

The coat colour of the klipspringer varies slightly according to their area of origin, with those from the Western Cape having a yellow coat flecked with brown. The hairs are hollow, flattened and bristly and are used to assist in the regulation of body temperature by facilitating the radiation of heat during the hottest parts of the day, and acting as a thick, insulating layer during periods of extreme cold (Norton 1980). In the recent past, these spongy hairs were the material of choice for stuffing horse saddles, which led to the extensive hunting of these antelope. This caused their numbers to decrease dramatically, and they have consequently become extinct in various parts of their range (Skinner & Chimimba 2005). One of the regions that the klipspringer has been eliminated from is Table Mountain National Park (TMNP) in the Western Cape of South Africa. In addition to extensive hunting, another contributing factor to their extinction in this region is thought to be the inadvertent release of a number of Himalayan tahrs (*Hemitragus jemlahicus*) in the area, which utilise the same habitat as the klipspringer (Skead 1985). The reintroduction of klipspringer into the area began in 1999, with the eradication of the tahrs following shortly thereafter. To date, park management have translocated a total of 56 individuals from various locations in the Western Cape, although the origin and/or success of these translocations is unclear (pers. comm., Mr Cheney, TMNP). In order to stock the park with additional klipspringer, it is important to identify the most genetically appropriate source population or populations, so that inbreeding, founder effects and outbreeding depression are minimised and thereby genetic diversity is maximised.

1.2 Subspecies designations within *Oreotragus*

The genus *Oreotragus* was first described by Zimmerman (1783) from specimens collected in the Cape of Good Hope, South Africa. Across the broad geographic distribution of *Oreotragus*, Ansell (1972) described 11 different subspecies of klipspringer based on variations in physical appearance. Within South Africa, *O. o oreotragus* is described from the Western and Northern Cape, and *O. o transvaalensis* is described inhabiting Gauteng, Kwa-ZuluNatal and possibly southern Mozambique. Namibia and Angola are home to *O. o tyleri*, while *O. o stevensoni* is a resident of Zimbabwe and, to a lesser extent, Botswana. Tanzania boasts three of the klipspringer subspecies, with *O. o centralis* in the south-west (also in northern Malawi and Zambia), *O. o aceratos* in the south-east (also in northern Mozambique) and *O. o schillingsi* occupying the rest of the country. *O. o schillingsi* is interesting in that the females of this subspecies are the only *Oreotragus* females to possess well-developed horns. *O. o aureus*, *O. o saltatrixoides* and *O. o somalicus* are described from northern Kenya, Ethiopia and Sudan, and northern Somalia respectively. *O. o porteousi* is thought to be found in only a single province in Nigeria and is consequently the only subspecies considered to be seriously threatened. However, the precise limits within which these 11 subspecies occur are unclear (Skinner & Chimimba 2005).

The subspecies concept itself has been the subject of much debate and controversy over the years (Manier 2004; Monroe 1982). Originally, the word subspecies was used to distinguish between geographical varieties, with all distinct natural populations not significantly different enough to be separate species being classified in this manner (Mayr 1982). The naming of subspecies continued to be popular until the 1950's, at which time a number of concerns were raised as to the relevance and accuracy of this classification (Mayr 1982; Inger 1961; Wilson

& Brown 1953). Such criticisms included the fact that many subspecies were described based on a single specimen, and that instances of poor geographical sampling often obscured gradual clinal differentiation (Manier 2004; Barrowclough 1982). Two divergent philosophies about the underlying purpose of the subspecies classification further exacerbated this controversy (Manier 2004). Some argued that subspecies represented incipient species, while others maintained that subspecies represented patterns of geographical variation (Mayr 1982). Inconsistent use of the subspecies concept, based on the discrepancies between these two schools of thought, led to the suggestion to abandon it completely (Manier 2004). However, its practical value in naming and identifying differing populations below the species level (Storer 1982), and in contributing to the understanding of the geographic variation of species, sees this form of classification still widely practised today. It is generally agreed, however, that the validity of previously designated subspecies should be tested using other more robust methods (Barrowclough 1982; Zusi 1982).

The identification of conservation units within a species can be immensely useful in conservation biology. Moritz (1994) proposed the use of ESUs (evolutionary significant units) and MUs (management units) according to the level of genetic and/or ecological distinctiveness found in populations of a particular species/subspecies. ESUs are used to describe populations that have differentiated over time to become reciprocally monophyletic, while MUs are used to describe populations that differ significantly with respect to allele frequency distributions (Moritz 1994). Phylogenetically distinct lineages with restricted geographical distribution would qualify as ESUs, and may warrant special protection in order to preserve the evolutionary history of the species in question, while MUs require structured conservation strategies to avoid problems such as inbreeding and outbreeding depression (Crandall *et al.* 2000)

1.3 Conservation genetics

Genetic diversity refers to the amount of variation present within a population, and is vital for a population to adapt, and therefore evolve with environmental change (Amos & Harwood 1998; Frankel & Soule 1981). The genetic diversity of a population is influenced by a number of factors, such as effective population size, genetic drift, mutation rate, patterns of migration, inbreeding and mating systems, population structure, founding effects and other demographic events (Amos & Balmford 2001; Amos & Harwood 1998). Inbreeding, or the breeding of individuals that are more related than would be found by chance, can lead to increases in homozygosity and the frequency of recessive deleterious alleles in the population, as well as cause a reduction in the fitness of progeny (Gemmell *et al.* 2004; Keller & Waller 2002; Frankham 1995). Outbreeding depression can be equally harmful and occurs when genetically differentiated individuals/populations interbreed and produce offspring with reduced fitness. This is of particular concern for conservation management strategies that involve the translocation and reintroduction of individuals (Rebholz & Harley 1999; Storfer 1999).

In the past two decades, significant advancement has been seen in the acquisition of DNA sequence data. This progress has been driven most notably by the development of automated sequencing technology and is continually improved with the constant addition of data to public sequence databases (Emerson *et al.* 2001). The use of sequence data has advanced the progress of a variety of fields, such as the study of human evolution, virology, conservation biology and phylogeography (Emerson *et al.* 2001). Phylogeography can be defined as the geographical distribution of genealogical lineages, and this knowledge of distribution patterns along with that of mutation rates can be used to provide insight into the demographic history and biogeographical range of populations (Avice 1987). Phylogeographic studies provide a

highly informative approach to investigating levels of diversity within and among populations, as well as population structure and demographic history, and are often employed to answer questions in conservation biology (Arbogast & Kenagy 2001; Ricklefs & Schluter 1993).

There are numerous factors that influence the phylogeographic patterns of taxa, and these include dispersal capabilities and life history, lineage sorting, vicariance and isolation by distance (Arbogast & Kenagy 2001). The dispersal abilities and life history traits of a species can have a profound effect on the rate and direction of gene flow and thus on the degree of population structuring within a species. For example, species that utilise broadcast spawning as a reproductive strategy are predicted to show high levels of genetic variation but low levels of population structuring (Burton 1998). Lineage sorting refers to the loss of haplotypes over time due to random genetic drift, which is particularly prevalent in small populations. Vicariance occurs when populations, and therefore gene flow, are separated by environmental events or barriers, such as the formation of mountain ranges or changes in the course of a river (Beebe & Rowe 2004). Isolation by distance is generally observed in the absence of any physical barrier to gene flow, and is the result of individuals mating with their closest neighbours. In this way, populations at opposite ends of a species' biogeographic range can be reasonably divergent (Slatkin 1993).

Within the field of conservation genetics, phylogeographic data can be used for a variety of applications, for example the determination of the source populations of introduced species, the determination of suitable source populations for translocations, or to establish whether populations are indigenous or introduced by estimating the time of divergence between populations from different geographical areas (Beebe & Rowe 2004). Phylogeographic data can also be used to make complex taxonomic classifications, such as in cases where there is a

continuous cline of variation, making it imperative to sample the entire biogeographical range of a species (Beebee & Rowe 2004). A number of recent studies, e.g. Nersting and Arctander (2001) and Lecis and Norris (2004), have examined the phylogeographic structure of taxa, with a particular emphasis on the conservation implications of the resultant data.

Nersting and Arctander (2001) investigated the phylogeographic structure of the greater kudu (*Tragelaphus strepsiceros*) and impala (*Aepyceros melampus*) in Africa using mitochondrial DNA markers. Results for the greater kudu suggested that the colonisation of new areas occurred from southern Africa towards eastern Africa. The impala results showed a divergence of lineages between south western Africa and other areas with a recognised subspecies, the black-faced impala (*A.m. petersi*), appearing genetically distinct. In addition, the authors found evidence of hybridisation between the vulnerable black-faced impala and the common impala, possibly due to cohabitation in areas previously occupied by only black-faced impala. These findings have profound conservation implications with respect to government legislation, in order to prevent the movement of common impala into areas where only black-faced impala are indigenous.

Lecis and Norris (2004) studied the genetic diversity and population structure of the critically endangered Sardinian newt (*Euproctus platycephalus*). They found that the four populations sampled on the island of Sardinia showed significant population differentiation and strong geographical structuring. The two clades discovered were attributed to range expansions and contractions during periods of adverse and favourable conditions. Coupled with years of drought, the authors found it unlikely that any gene flow could currently exist among the populations and suggested that management plans should be implemented such that the populations are managed as separate units and migration is promoted as much as possible.

1.4 Mitochondrial DNA

Animal mitochondrial genomes are usually closed-circular molecules of approximately 15-20 kb in size and play a vital role in processes such as metabolism, disease, apoptosis and aging (Ballard & Whitlock 2004). Almost all animal mitochondrial genomes encode the same 37 genes, of which two are rRNAs, 13 are protein subunits of the oxidative phosphorylation enzymes, and 22 are tRNAs required for the translation of mitochondrial proteins (Ballard & Whitlock 2004; Boore 1999). There is also a single, large non-coding region thought to be involved in replication control and gene transcription, which is commonly referred to as the control region or D-loop (Boore 1999).

Mitochondrial DNA (mtDNA) is commonly the genetic marker of choice in phylogeographic and phylogenetic research, and is used in more than 80% of these studies (Ballard & Dean 2001; Avise 1998). It has a number of properties which contribute to its effectiveness in elucidating patterns of genetic variation within and among populations. MtDNA has a high copy number in most tissues, is therefore easy to amplify using the polymerase chain reaction (PCR), and has been successfully amplified from museum and archaeological specimens (Su *et al.* 1999). The mitochondrial genome has a high mutation rate compared to the nuclear genome as a result of a deficiency of repair enzymes and increased exposure to oxidative damage, and thus variation is found within almost all populations (Ballard & Whitlock 2004; Gemmell *et al.* 2004; Ballard & Dean 2001). The mitochondrial genome is also small in comparison to the nuclear genome, and is therefore more susceptible to the effects of genetic drift and bottleneck events (Gemmell *et al.* 2004). Lastly, mtDNA is maternally inherited and does not generally undergo recombination, which means that each haplotype has just one ancestor in the previous generation (Ballard & Whitlock 2004; Birungi & Arctander 2000).

All of the above properties facilitate the identification of unique haplotypes within populations and the inference of relationships between them.

The cytochrome *b* and control region (D-loop) of the mitochondrial genome are both used extensively in phylogeographic studies (Eggert *et al.* 2002). Cytochrome *b* is a protein-encoding gene which evolves at a moderate rate, allowing resolution at the species and genera level. The control region is regulatory in function and evolves much more rapidly, thus making it more suitable for studies at the population level (Tang *et al.* 2005). Because it is non-coding, the control region generally shows substantially more variation than cytochrome *b*, including both substitutions and insertions/deletions (indels) (Tang *et al.* 2005). Consequently, the control region is the most variable region in the entire mammalian mitochondrial genome and is characterised by vast changes in sequence composition and length (Birungi & Arctander 2000). Although it is generally accepted that the control region mutates at a rate that is two to five times higher than that of the mitochondrial protein-encoding genes, there are exceptions to this principle (Tang *et al.* 2005). In a number of organisms, such as rats and mice (Brown *et al.* 1986) and salmonid fishes (Bernatchez & Danzmann 1993), the control region has been shown to have a slower substitution rate than the protein-encoding genes (Tang *et al.* 2005). However, as cytochrome *b* has been used to elucidate the phylogenetic relationships within the family Bovidae (Hassanin & Douzery 1999; Matthee & Robinson 1999) of which the klipspringer forms part, the faster-evolving control region is more likely to be the more suitable marker to resolve the relationships at the population level.

1.5 Nuclear integrations

Despite the benefits of using mtDNA, there are also a number of known difficulties associated with the amplification of mitochondrial genes, including phenomena such as biparental inheritance and heteroplasmy (Zhang & Hewitt 1996). Chief amongst these concerns is the co-amplification, or preferential amplification, of nuclear mitochondrial-like sequences, or Numts (Venkatesh *et al.* 2006; Bensasson *et al.* 2001; Zhang & Hewitt 1996). The integration of mitochondrial-like sequences into the nuclear genome was first described in 1967 (Bensasson *et al.* 2001), and is proposed to occur via non-homologous recombination between nuclear and mitochondrial DNA during chromosomal repair (Venkatesh *et al.* 2006; Bensasson *et al.* 2001). Once incorporated into the nuclear genome, Numts can undergo duplication events, thus resulting in the presence of multiple copies within a single species (Venkatesh *et al.* 2006; Antunes & Ramos 2005; Zhang & Hewitt 1996). The transfer of mitochondrial material into the nuclear genome is not limited to specific types of genes; Numts have been found for protein-coding, RNA-coding and non-coding regions. In addition, insertion may involve a large section of the mitochondrial genome as opposed to only a single gene (Zhang & Hewitt 1996); this is seen in the domestic cat, which has approximately half of its mitochondrial genome (7.9kb) inserted into the nuclear genome and tandemly repeated (Bensasson *et al.* 2001; Zhang & Hewitt 1996).

Numts have been reported in a wide variety of species, including gorillas (Anthony *et al.* 2007), the domestic cat (Lopez *et al.* 1994) and grasshoppers (Bensasson *et al.* 2000) and, if mistakenly included in analyses, can negate the benefits of mtDNA, inflate estimates of diversity and mislead phylogenetic and population genetic studies (Anthony *et al.* 2007; Zhang & Hewitt 1996). Since the genetic code differs between the nuclear and mitochondrial

genomes in animals, Numts become non-functional pseudogenes upon incorporation into the nuclear genome (Venkatesh *et al.* 2006; Antunes & Ramos 2005; Bensasson *et al.* 2001). Their position in the nucleus also means that Numts are under different mutational constraints compared to their mitochondrial counterparts, as the mutation rate of the nuclear genome is much slower than that of the mitochondria (Bensasson *et al.* 2001; Zhang & Hewitt 1996). Numts are therefore more similar to an ancestral form of the gene in question, and usually show little or no intra-specific sequence variation (Bensasson *et al.* 2001; Zhang & Hewitt 1996).

Due to their conserved nature and slow mutation rate, Numts can be co-amplified along with the mitochondrial target, or even preferentially amplified, during PCR reactions. This occurs particularly when conserved PCR primers are used to target the region of interest, because a better match is made between the conserved primer sequence and the slower-evolving Numt, or when gene rearrangement has taken place in the mitochondrial but not the nuclear copies (Bensasson *et al.* 2001; Zhang & Hewitt 1996). Numerous other factors influence the likelihood of Numt amplification, including the proportion of mitochondrial to pseudogene target, the DNA extraction procedure, as well as the type of tissue used for extraction (Bensasson *et al.* 2001). Signs of Numt amplification include: the consistent generation of more than one or different PCR products; repeated sequence ambiguities; the occurrence of inexplicable indels (insertions or deletions), frameshift mutations or stop codons within coding regions and the generation of unexpected or non-sensical phylogenetic trees (Zhang & Hewitt 1996). Circumventing the problem of Numt contamination can be difficult, but may be overcome by using mtDNA-specific primers, extracting DNA from tissue that is rich in mtDNA e.g. heart or muscle, or by purifying mtDNA from total extracted DNA before PCR amplification of the target region (Zhang & Hewitt 1996). If these methods prove

unsuccessful, it is sometimes possible to determine which sequences are nuclear versus mitochondrial, once they have been sequenced, by their divergent patterns of evolution (Bensasson *et al.* 2001)

1.6 Sample types

The advent of PCR has enabled both invasive and non-invasive sampling of biological materials for the extraction of DNA (Whittier *et al.* 1999). While blood and tissue samples are conventionally the most straightforward to work with, the expense and difficulty in obtaining these samples, as well as the necessary permits for animals in protected areas or reserves, has become increasingly problematic for researchers (Whittier *et al.* 1999; Kohn & Wayne 1997). The growing need to answer questions on the phylogeography, population structure and demographic history of populations and species has meant that other potential sources of DNA have had to be utilised more frequently, in particular faecal and museum material (Rohland *et al.* 2004; Kohn & Wayne 1997).

Material from specimens held in museum collections, usually bone or teeth, are an indispensable and relatively inexpensive way to obtain information on the historic distributions of species and extinct populations, ensure extensive species-wide sampling and allow the genetic comparison of historic and present-day populations (Rohland *et al.* 2004). Although the DNA extracted from museum samples is often degraded, PCR can successfully amplify sufficient amounts of DNA from very small or degraded samples (Kohn & Wayne 1997; Höss & Pääbo 1993). In the past, museums allowed very restricted use of their collection material, as samples had to be ground up or otherwise destroyed to obtain sufficient quantities of DNA,

but recent methodological advances have enabled the extraction of DNA from bones and teeth without damage to the sample material (Rohland *et al.* 2004). The main problem associated with museum material is contamination, and extreme care must therefore be taken during the extraction procedure (Höss & Pääbo 1993). Numerous studies using museum samples have been carried out successfully, including research on musk deer (Su *et al.* 1999), elephant seals (Weber *et al.* 2000) and extinct marsupials (Westerman *et al.* 2004) among others.

Faecal sampling is the least invasive of all sampling methods, and although it does require fairly extensive field work, samples of this type are relatively easy to collect and store for later processing (Whittier *et al.* 1999, Reed *et al.* 1997). Faecal material contains a variety of substances including cells which are shed from the intestinal lining of the host, micro-organisms, undigested food, bile salts and bilirubin (Kohn & Wayne 1997, Reed *et al.* 1997). The actual concentration of the host cells relative to the other components is low, which can hinder the amplification of target DNA, and degradation can also occur due to the high bacterial content (Reed *et al.* 1997). In addition, substances that are co-extracted along with the DNA from faecal material, such as bilirubin, are often found to inhibit PCR reactions, and can also contribute to inconsistent results (Whittier *et al.* 1999; Kreader 1996; Deuter *et al.* 1995). Nevertheless, the optimisation of PCR conditions, the use of species-specific primers targeting short segments of the target DNA region, and the addition of compounds such as bovine serum albumin (BSA) can help to reduce these problems (Kohn & Wayne 1997; Kreader 1996). Recent studies have successfully extracted and amplified both nuclear and mitochondrial gene regions from the faecal samples of primates (Whittier *et al.* 1999), seals (Reed *et al.* 1997), African elephants (Eggert *et al.* 2002) and Asian elephants (Fernando *et al.* 2000) for a variety of applications, including the investigation of population structure, phylogeography, and assigning species and sex to individuals.

1.7 Research objectives

When devising and implementing conservation management strategies, it is essential to understand all the components of a species' life history: the phenotypic and/or genotypic basis of subspecies classifications, aspects of population structure, phylogeographic history, demographic history, and the extent and distribution of variation (Avice 2000). With this in mind, the objectives of this study included the following:

- 1) To examine the level of cytochrome *b* and mitochondrial control region variation found in klipspringers from various parts of their biogeographical range
- 2) To use these data to elucidate the phylogeographic structure of klipspringer from different geographic regions across its broad distribution in Africa, and
- 3) To use the results of these analyses to:
 - i. assess the genetic support for the existing subspecies classifications of klipspringer (as per Ansell 1972), and
 - ii. infer past demographic events and thus provide information that could be useful for guiding conservation management plans for the re-introduction of klipspringers, by translocation, into areas that were historically occupied by this species.

Chapter 2

Materials and Methods

2.1 Sample collection

A total of 90 klipspringer samples were available for this study. Samples consisted of a combination of blood, faecal samples and museum material (Table 1). Although a total of 56 animals have been reintroduced into TMNP, only 14 samples were obtained from this group. These were obtained as blood samples from individuals originating from Limietberg Nature Reserve, Paarl ($n=7$) and Waterfall Nature Reserve, Tulbagh ($n=7$). Samples were stored at 4° C. Samples from Karoo National Park ($n=22$), Kruger National Park ($n=3$) and the Cape Point section of TMNP ($n=7$) were obtained as dung samples. Dung samples were kept on ice until they reached the laboratory, where they were stored at -20° C in 96% ethanol. Museum samples ($n=8$) in the form of teeth or bone were obtained from the mammal collection of the Iziko Museum, Cape Town. These samples were all from different locations in South Africa; unfortunately some of these locations were unknown. Additional museum samples were obtained from Namibia ($n=4$), Malawi ($n=2$), Tanzania ($n=15$), Transvaal/KwaZulu-Natal ($n=8$), the Western Cape ($n=3$), Kenya ($n=1$), Uganda ($n=1$), Ethiopia ($n=1$) and Somalia ($n=1$). Unfortunately no further information was available for these additional samples.

Although every attempt was made to obtain additional samples from other National Parks throughout South Africa, this was not possible as a result of the logistical constraints of SANParks. A summary of all the sample locations, and the sample numbers, is given below in Table 1. The samples used in this study represent nine of the 11 klipspringer subspecies, and are representative of a large part of the species' range (Figure 2).

Table 1: Klipspringer samples obtained and analysed in this study.

Country	Location	Number of individuals	Subspecies	Source material
South Africa	Unknown	2	Unknown	Museum
	Tulbagh	7	<i>O. o. oreotragus</i>	Blood
	Paarl	7	<i>O. o oreotragus</i>	Blood
	Kruger National Park	3	<i>O. o transvaalensis</i>	Dung
	Karoo National Park	22	<i>O. o oreotragus</i>	Dung
	Cape Point	8	<i>O. o oreotragus</i>	Dung
	Western Cape	6	<i>O. o oreotragus</i>	Museum
	Transvaal/KwaZulu-Natal	8	<i>O. o transvaalensis</i>	Museum
Malawi	Nchisi (2), Shire Highlands (1), Mlanje (1)	4	<i>O. o centralis</i>	Museum
Tanzania	Mashwa (6), Makau (2), Wembere (1)	9	<i>O. o schillingsi</i>	Museum
	Muhesi (3), Kisigo (2), Unknown (1)	6	<i>O. o aceratos</i>	Museum
	Windhoek (1), Otavi (1), Cunene Falls (1), Unknown (1)	4	<i>O. o tyleri</i>	Museum
Ethiopia	Unknown	1	<i>O. o saltatrixoides</i>	Museum
Uganda	Unknown	1	<i>O. o aureus</i>	Museum
Kenya	Unknown	1	<i>O. o aureus</i>	Museum
Somalia	Unknown	1	<i>O. o somalicus</i>	Museum



Figure 2: Countries from which klipspringer samples were obtained (main map) relative to their distribution (inset) (<http://geography.about.com/library/blank/blxafrica.htm>). Countries coloured yellow indicate those sampled in this study.

2.2 DNA extraction

DNA was extracted from the blood samples following proteinase K digestion, using a standard isoamyl alcohol/salt extraction method (Mullenbach *et al.* 1989). Dung samples were extracted using a Qiagen Stool Kit, as per the manufacturer's instructions, with modification of the first lysis step. This step was performed overnight because it was found to yield better results. The museum samples (teeth and bone) were soaked in 5ml of 5M guanidium thiocyanate (GuSCN) buffer for 1-2 weeks at 37°C with gentle agitation (Rohland *et al.* 2004), after which 1ml of the GuSCN buffer was processed through a Qiagen DNeasy Tissue

Extraction kit, as per the manufacturer's instructions. GuSCN was used because it is a powerful denaturant, and it denatures proteins, lyses cells and inactivates nucleases (Reed *et al.* 1997). After elution, samples were concentrated by evaporation at 65°C overnight, and resuspended in 50µl dH₂O. Control extraction blanks were prepared with all batches of museum samples. The remaining museum samples were obtained as extracted DNA from the Department of Botany and Zoology, University of Stellenbosch (Prof T.J. Robinson). DNA was stored at 4°C. All DNA extractions were checked for quality by electrophoresis on 1.5% agarose gels and visualised by ethidium bromide (EtBr) under UV light.

2.3 DNA amplification

2.3.1 Primerless PCR

All museum samples were subjected to a pre-amplification step by carrying out a 'primerless' PCR on the extracted DNA. Primerless PCR, proposed by Stemmer (1994), is particularly useful for degraded DNA such as that extracted from museum samples. It has been used successfully to reconstruct DNA templates in studies such as the northern elephant seal (Weber *et al.* 2000). In this method the short, degraded DNA fragments bind to each other in the absence of primers, and after elongation by the *Taq* DNA polymerase, result in longer, continuous stretches of DNA (Weber *et al.* 2000). The PCR reactions were performed on a Hybaid Thermal Reactor, and the conditions used for the amplification were as follows: denaturation at 94°C for 3min; 35 cycles of denaturation at 92°C for 1min, annealing at 48°C for 1min and elongation at 72°C for 1min, and then a final elongation step at 72°C for 10min. Reactions were performed in a final volume of 40µl, containing 4-6µl template DNA, 1X

GoTaq Reaction Buffer, 1.5mM MgCl₂, 0.2mM DNTP's and 1.25U of GoTaq DNA Polymerase (Promega). Negative controls were included in each PCR.

2.3.2 *The cytochrome b gene fragment*

Various conserved primers were initially used to attempt to amplify a 500 base pair (bp) fragment of the cytochrome *b* gene, but the target DNA seemed to be too degraded for the successful amplification of a fragment of that length. Internal cytochrome *b* primers cyt1 (5'- CCT AGC CAT AGT ACA CCT AC -3') and cyt2 (5'- GAC GCC TCC TAG TTT ATT GG -3') were then designed from an alignment of the klipspringer cytochrome *b* sequences found in GenBank (accession numbers AF036288, AF022052). These primers were used to amplify a 251bp fragment of the cytochrome *b* gene. PCR conditions were optimised with respect to MgCl₂ concentration, primer concentration and annealing temperature, and were performed on a Hybaid Thermal Reactor. Cycling conditions were as follows: denaturation at 94°C for 3min; 35 cycles of denaturation at 94°C for 30sec, annealing at 58°C for 30sec and elongation at 72°C for 45sec, and then a final elongation step at 72°C for 7min. Reactions were performed in 20µl, containing 1-2µl DNA template, 1X GoTaq Reaction Buffer, 1.5mM MgCl₂, 0.2mM DNTPs, 0.5µM of each primer, and 1.25U of GoTaq DNA Polymerase (Promega).

2.3.3 *The Control Region Fragment*

Numerous primer sets, both published conserved primers and primers designed in this study, were used to amplify various fragments of the control region, in order to identify the authentic

mtDNA control region sequence from nuclear copies and PCR recombinants. A summary of the primer sets used is given below.

(i) Primers C and E

Initially, a 700bp fragment of the control region was amplified with the conserved primers C and E (Wilkinson & Chapman 1991) (Table 2). PCR conditions were optimised with respect to MgCl₂ concentration, primer concentration and annealing temperature and were performed on an ABI GeneAmp® 2700 PCR System. Cycling conditions were as follows: denaturation at 94°C for 4min, 35 cycles of denaturation at 94°C for 45sec, annealing at 54°C for 45sec and elongation at 72°C for 1min, and a final elongation step at 72°C for 7min. Reactions were performed in 40µl, containing 4µl DNA template, 1X GoTaq Reaction Buffer, 1.5mM MgCl₂, 0.2mM dNTPs, 0.5µM of each primer, and 1.25U of GoTaq DNA Polymerase (Promega). While this amplification was successful in the blood samples, it did not yield results in the dung or museum samples despite extensive attempts at optimisation, which included the addition of bovine serum albumin (BSA) in an attempt to minimise possible inhibition. In order to circumvent this problem a pair of nested primers was designed.

(ii) Primers C2 and E2

Primers C2 and E2 (Table 2) were designed as nested primers, using the sequences obtained from the blood samples, in conjunction with the sequences of primers C and E (Wilkinson & Chapman 1991). A two-step, or nested, PCR successfully amplified a 600bp fragment from the dung samples. The initial amplification was performed with primers C and E according to the same protocol as for the blood samples, but with the addition of 1000ng/µl BSA to

minimise inhibition (Kreader 1996). The resulting PCR product was then used as the template in the nested PCR. Nested PCRs were performed on a PTC-100™ Programmable Thermal Controller (MJ Research) with the following conditions: denaturation at 94°C for 4min, 35 cycles of denaturation at 94°C for 45sec, annealing at 54°C and elongation at 72°C for 1min, and a final elongation step at 72°C for 7min. The annealing temperature of the initial PCR was decreased to 50°C for samples that failed to amplify after the nested PCR. Nested reactions were performed in 40µl, containing 4µl template, 1X GoTaq Reaction Buffer, 1.5mM MgCl₂, 0.2mM dNTPs, 0.5µM of each primer, and 1.25U of GoTaq DNA Polymerase (Promega).

After sequencing, it was apparent that the sequences from the blood and dung were vastly different, and the different sequences needed to be verified as mitochondrial copies versus nuclear copies or possible PCR recombinants. While a mitochondrial DNA extraction from heart or liver tissue would have been the ideal solution to this problem, no such tissue was obtainable despite extensive efforts. The nested primers C2 and E2 were then used alone to amplify product from the blood samples using the same conditions as for the second step of the nested PCR. These results were used to identify the authentic mtDNA copy (see chapter 3, section 3.2), and the conserved regions in these sequences were then used to design *Oreotragus* mtDNA-specific control region primers.

(iii) Primers blood1 and blood4

Primers blood1 and blood4 (Table 2) were designed using the conserved regions of the authentic mtDNA control region sequences obtained from the blood samples. These primers were used to amplify a 480bp fragment of the control region by PCR. PCR conditions were optimised with respect to MgCl₂ concentration, primer concentration and annealing

temperature, and were performed on a Hybaid Thermal Reactor at the following conditions: denaturation at 94°C for 3min. 5 cycles of denaturation at 94°C for 45sec, annealing at 45°C for 45sec and elongation at 72°C for 1.5min, then 30 cycles of denaturation at 94°C for 30sec, annealing at 49°C for 30sec and elongation at 72°C for 45sec, and a final elongation step at 72°C for 7min. Reactions were performed in 20µl, containing 2µl template, 1X GoTaq Reaction Buffer, 2mM MgCl₂, 0.2mM DNTPs, 0.5µM of each primer, and 1.25U of GoTaq DNA Polymerase (Promega). Samples that failed to amplify under these conditions were amplified under the following conditions: denaturation at 94°C for 3min. 5 cycles of denaturation at 95°C for 30sec, annealing at 47°C for 45sec and elongation at 72°C for 1min, then 35 cycles of denaturation at 94°C for 50sec, annealing at 50°C for 1min and elongation at 72°C for 2min, and finally two cycles of elongation at 72°C for 5min.

(iv) Primers blood2, blood3, blood6 and blood7

For dung and museum samples that failed to amplify under any of the previously described conditions, additional primers were designed. These primers were used in a variety of combinations, with themselves and others, in order to produce two overlapping control region fragments, which could then be combined to obtain the full 423bp target fragment for each sample. The positions of all the primers can be seen in Figure 3. Primers blood1 and blood6 were used to obtain the first part of the fragment, and either blood3 and blood2, or blood7 and blood 2 were used to amplify the second part of the fragment.

PCR conditions were optimised with respect to MgCl₂ concentration, primer concentration and annealing temperature, and were performed on a Hybaid Thermal Reactor at the following conditions: denaturation at 94°C for 3min, 5 cycles of denaturation at 95°C for 30sec,

annealing at 44-50°C for 45sec and elongation at 72°C for 1min, then 35 cycles of denaturation at 94°C for 50sec, annealing at 48-53°C for 1min and elongation at 72°C for 2min, and finally two cycles of elongation at 72°C for 5min. Reactions were performed in 20µl, containing 4µl primerless PCR template, 1X GoTaq Reaction Buffer, 2mM MgCl₂, 0.2mM dNTPs, 0.5µM of each primer, and 1.25U of GoTaq DNA Polymerase (Promega).

Table 2: Details of the primers used for the *Oreotragus* control region amplification

Primer Name	Primer Sequence (5'-3')	Source
C	TGA ATT GGA GGA CAA CCA GT	Wilkinson & Chapman 1991
E	CCT GAA GTA GGA ACC AGA TG	Wilkinson & Chapman 1991
C2	CAA CCA GTT GAA CAC CCA TAC	This study
E2	CCA GAT GTC TGA TAA AAT TCG	This study
Blood1	GCT ATA GCT CCA CTA TCA AC	This study
Blood2	GTG CTA TGT ACG ATC AAT CAG	This study
Blood3	CAC ATG AAT GAG CAC ATA C	This study
Blood4	GCT TAT ATG CAT GGA CTA G	This study
Blood6	CAC CCA TGT GCT AGG TAG	This study
Blood7	GAG CAC ATA CAG TTA ATG TAC	This study

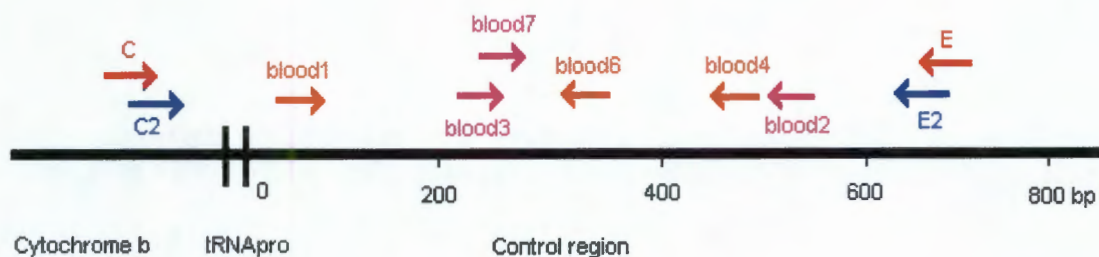


Figure 3: The positions of the control region primers. Colours indicate forward and reverse primer sets.

2.4 PCR purification and sequencing

All PCR products were electrophoresed on 1.5% agarose gels in 1X TBE buffer and visualised by EtBr under UV light. The correct size fragments were identified using a λ PstI DNA ladder and excised from the gel. These bands were then purified using the Promega Wizard PCR Clean-Up Kit, as per the manufacturer's instructions. Cycle sequencing was performed using the BigDye® Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems) in 10 μ l reactions, containing 1-4 μ l of purified DNA, 3.2 μ M primer, 2X sequencing buffer and 2 μ l Big Dye solution. The final volume was made up to 10 μ l with distilled water. The cycle sequencing was performed on an ABI GeneAmp® 2700 PCR System, using the same primers as in the PCR amplification, with the following thermal cycling profile: 25 cycles of 94°C for 30sec, 50°C for 5sec and 60°C for 4min. The automated capillary electrophoresis was performed by the Central Analytical DNA Sequencing Facility at the University of Stellenbosch.

2.5 Data analysis

2.5.1 *Sequence identity confirmation*

Cytochrome *b* sequences were edited by eye in Chromas v2.3 (Technelysium Pty. Ltd. available at www.technelysium.com.au) and aligned using the ClustalW algorithm (Thompson *et al.* 1994) in BioEdit v7.0.5.2 (Hall 1999). The identities of all samples were confirmed using the BLASTN algorithm in the NCBI nucleotide database in GenBank (<http://www.ncbi.nlm.nih.gov/BLAST>).

2.5.2 *Identification of authentic mtDNA control region sequences*

The control region sequences were edited by eye in Chromas v2.3 and aligned using the ClustalW algorithm in BioEdit v7.0.5.2. The sequences appeared to represent three different product sets, namely two from the blood samples and one from the dung samples, identifying the need to determine the source DNA of all three sets. Haplotype diversities were determined for all three sequence sets. Possible Numt sequences would be expected to exhibit less diversity, as a result of a slower rate of mutation (Zhang & Hewitt 1996). These tests were implemented in DnaSP v4.10.9 (Rozas *et al.* 2006). Neighbour-joining phylogenetic trees were constructed in Mega v3.1 (Kumar *et al.* 2004) using Kimura-2-parameter corrected distances (Kimura 1980), and the strength of the phylogenies generated were calculated using bootstrap analysis that consisted of 1000 replications. The first tree constructed consisted of the three sequence sets, and the second consisted of the three sequence sets plus three known control region Numt sequences, numt1bos, numt1lovis and numt2bos, from GenBank

(accession numbers *Bos taurus* AF309773.1, *Ovis aries* AF309771.1 and *Bos taurus* AF309772.1 respectively).

The three sets of sequences were then analysed using the Recombination Detection Program RDP3 (Martin *et al.* 2005) to uncover any recombinant sequences, as well as to identify the possible parental sequences and possible daughter recombinants. The sequence alignment was analysed using the following program algorithms within RDP3: the original RDP method (Martin & Rybicki 2000), SiScan (Gibbs *et al.* 2000), GENECONV (Padidam *et al.* 1999), BOOTSCAN (Salminen *et al.* 1995), MAXCHI (Maynard Smith 1992) and CHIMAERA (Posada & Crandall 2001). In order to minimise the number of false positives, only recombination events identified by two or more methods were considered. Once the authentic mtDNA control region sequences were identified, all the other sequences were excluded and further analysis was performed on the authentic mitochondrial-encoded set only.

2.5.3 Genetic variation and phylogenetic analysis

Haplotype diversity, nucleotide diversity and the average number of pairwise differences were calculated using Arlequin v3.11 (Excoffier 2006), with gaps treated as a fifth state. Phylogenetic analyses were performed using neighbour-joining (NJ), parsimony (MP) and Bayesian methods. Neighbour-joining phylogenetic trees were constructed in SplitsTree4 (Huson & Bryant 2006) for the control region sequences using the best model of DNA substitution. This is necessary as accurate phylogenetic inferences depend on the use of the most appropriate model of DNA evolution for a particular dataset. The model of DNA substitution that best fitted the data was determined based on hierarchical likelihood ratio tests (hLRT) implemented within MODELTEST v3.06 (Posada & Crandall 1998) and run from

PAUP*4.0b10 (Swofford 2002). Node support was tested using 1000 bootstrap replications. The neighbour-joining method is a distance matrix method, which relies on a certain degree of genetic distance between the pairs of taxa under examination. It is often the distance method of choice because it is able to incorporate different mutation rates for different lineages ((Holder & Lewis 2003).

Parsimony phylogenetic trees were constructed in Mega v3.1, with 1000 bootstrap replications. The parsimony method generates a tree that uses the least number of mutational changes to explain the differences between individual lineages (Holder & Lewis 2003). Bayesian posterior probabilities were calculated using MrBayes v3.1.2 (Huelsenbeck & Ronquist 2003) employing random starting trees and the model of evolution selected by MODELTEST. Two different runs were performed in order to confirm that the chains were converging on the same posterior probability distribution. Each run consisted of 2 000 000 Markov Chain Monte Carlo (MCMC) generations with tree sampling every 50 generations and a burn-in of 10 000 trees. Bayesian analysis is advantageous as it produces a tree estimate as well as measures of uncertainty for the groups on the tree, based on the estimated posterior probabilities (Holder & Lewis 2003; Huelsenbeck & Ronquist 2001).

The relationships between the observed haplotypes were depicted using minimum-spanning networks, which were calculated in Arlequin v3.11. Minimum-spanning networks connect each haplotype to its closest neighbour based on pairwise differences, and arrange the haplotypes as nodes of a network rather than the terminal tips of a tree. This can be particularly useful if many of the sequences are equally derived from the same ancestral haplotype (Davison *et al.* 2001).

2.5.4 Population structure

Population genetic structure was investigated by the analysis of molecular variance (AMOVA; Excoffier *et al.* 1992), as implemented in Arlequin v3.11. This analysis estimates the proportion of genetic variation at different hierarchical levels, using information from the geographic distribution of haplotypes and the pairwise differences between them (Excoffier *et al.* 1992). AMOVA was also used to calculate F_{ST} values, in order to assess the level of population differentiation. Population analyses were only performed on localities that consisted of four or more klipspringer samples, or regions where it was appropriate to group smaller subsets of samples.

Chapter 3

Results

3.1 Sequence identity confirmation

The identity of the 251bp cytochrome *b* sequences obtained for each sample were compared against the NCBI nucleotide database, using the BLASTN algorithm in Genbank. All sequences except one were positively identified as of klipspringer origin. The remaining sequence, a dung sample from Karoo National Park, was identified as a mountain reedbuck (*Redunca fulvorufula*) illustrating that an incorrect dung sample was collected. This sample was consequently excluded from the study.

3.2 Identification of authentic mtDNA control region sequences

3.2.1 *Sequence types*

The initial amplification of a 600bp fragment of the mitochondrial control region yielded three sets of sequences which, when aligned, appeared too dissimilar to all be of mitochondrial origin (Figure 4). The first sequence type (S1) was amplified in the blood samples using the conserved primers C and E. Because these primers yielded no product in the dung samples, primers C2 and E2 were designed. In order to obtain a PCR product in the dung samples, a

nested PCR approach using both the primer sets was required. Radical differences between the sequences obtained from the dung (S3) and the blood (S1) prompted the use of the C2 and E2 primers in the blood. This reaction co-amplified S1 and another sequence type, S2. The sequence types resulting from the different primers and tissue types are shown in Table 3.

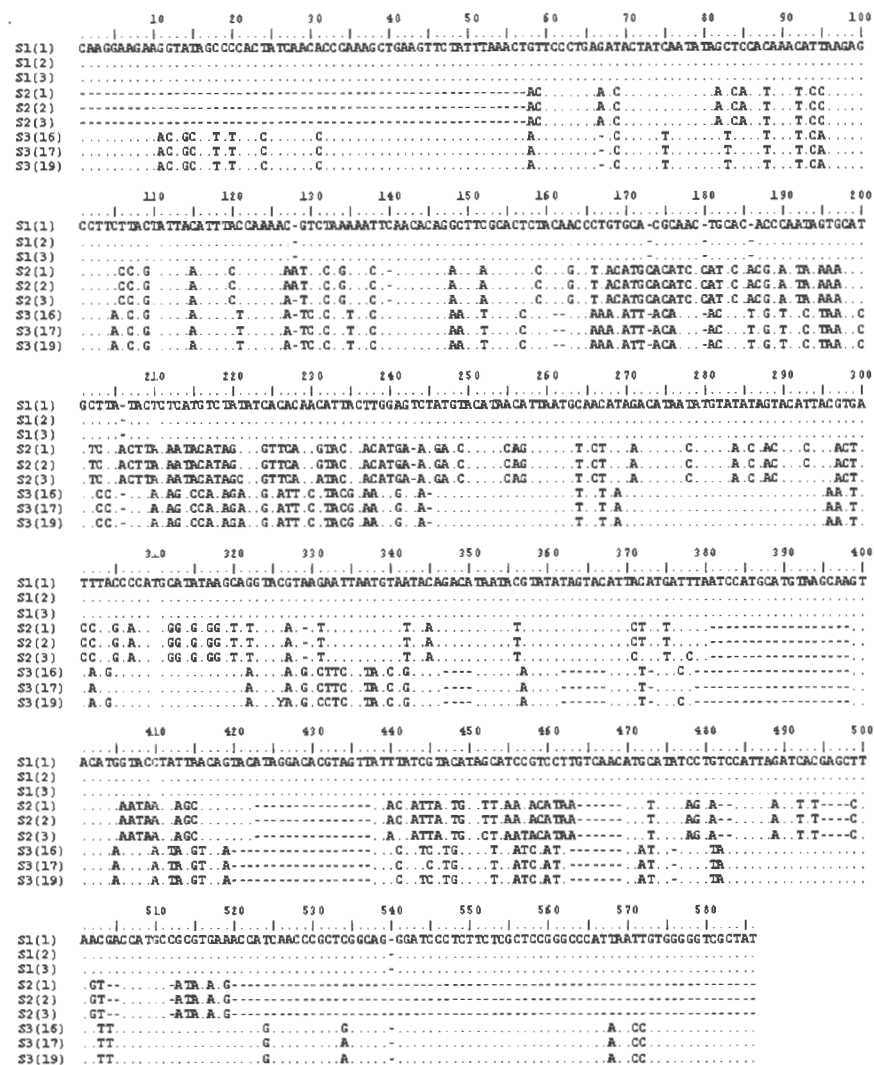






Figure 4: Alignment of representative sequences for sequence sets S1, S2 and S3. The first three sequences represent S1, the next three represent S2 and the last three represent S3.

Table 3: Details of the three sequence types obtained from the amplification of *Oreotragus spp.* control region

Primers Used	Source Material	Sequence/s Obtained
C & E	Blood	S1 
C & E	Dung	No amplification
C2 & E2	Dung	No amplification
C & E, then C2 & E2	Dung	S3 
C2 & E2	Blood	S1  S2 

3.2.2 Haplotype diversity and phylogenetic analysis

The haplotype diversities of S1, S2 and S3 were 0.53, 0.93 and 0.88 respectively. These values suggest S1 as the Numt sequences due to the lower diversity found for this group, but do not adequately differentiate between the origins of S2 and S3. Phylogenetic analysis of the three sequence sets using the neighbour-joining method showed the S2 sequences grouping independently of S1 and S3 (Figure 5a). When the same analysis was performed on the three sequence sets with the addition of three Numt sequences from Genbank, the Numt sequences clustered with S1 and S3 (Figure 5b). This suggests that both S1 and S3 are possibly nuclear in origin, although it does not explain the level of diversity observed in S3.

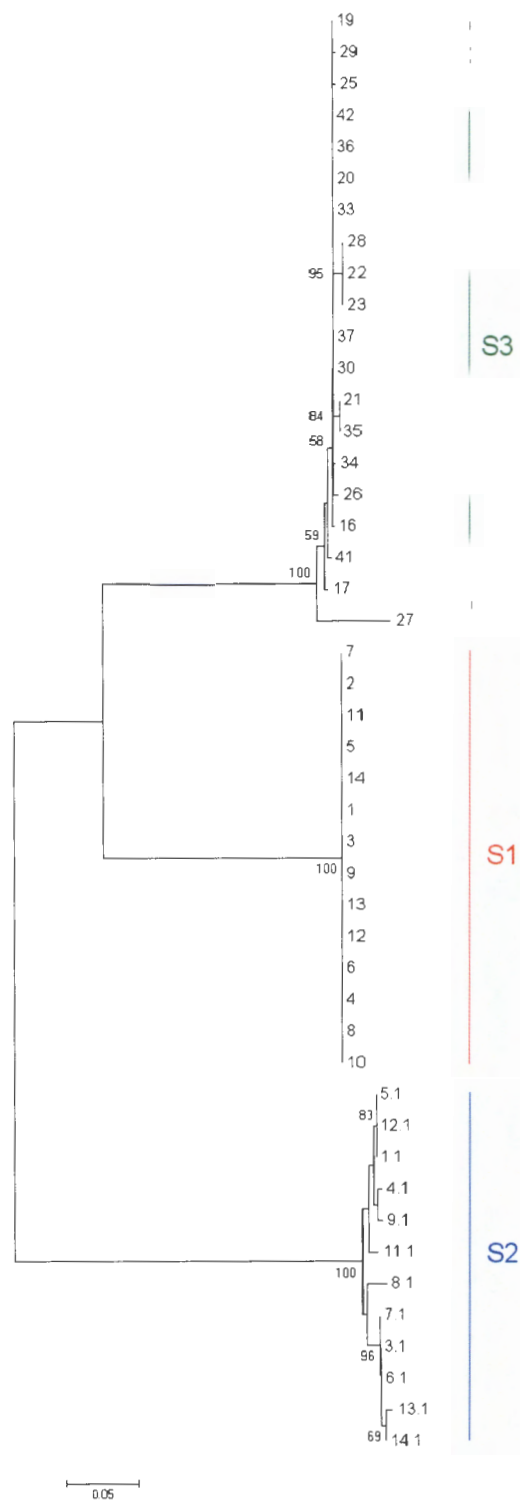


Figure 5(a): Neighbour-joining tree of the three sequence sets S1, S2 and S3. Bootstrap values above 50 are shown.

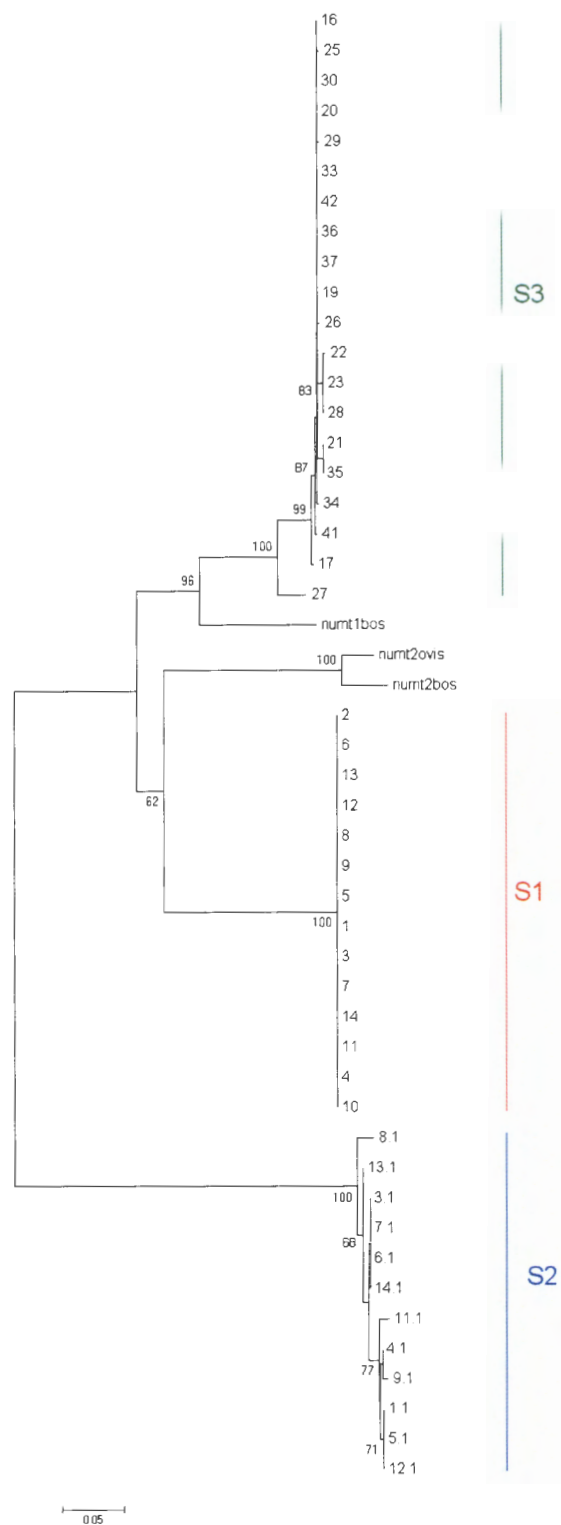


Figure 5(b): Neighbour-joining tree of the three sequence sets S1, S2 and S3 with the addition of three Numt sequences, numt1bos (*Bos taurus*), numt2ovis (*Ovis aries*) and numt2bos (*Bos taurus*). Bootstrap values above 50 are shown.

3.2.3 Identification of recombinants

RDP3 was used to detect possible recombination events between the three sequence types, as well as to identify possible parent sequences and daughter recombinants. The analysis of the dataset showed evidence to suggest that either the S1 or S3 sequence types were possible PCR recombinants of the other two sequence types. This was supported by two of the recombination detection methods, MAXCHI and CHIMAERA. Although the analysis was not entirely conclusive, when seen in conjunction with the haplotype diversity and phylogenetic results it seems most likely that the S3 sequence type is a result of recombination between the S1 and S2 sequence types. The recombination appears to occur at approximately 215-235 base pairs, with the first section of the S3 sequence type derived from S1, and the remainder derived from S2 (Figure 6).

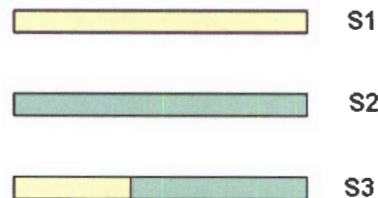


Figure 6: Schematic diagram of the proposed composition of the three sequence sets. Yellow sections represent probable Numt sequence type and green sections represent probable authentic mitochondrial copy sequence type.

The combination of the above results and the neighbour-joining analysis provide sufficient evidence to assign S1 as the most likely Numt sequence set, S2 as the authentic mitochondrial control region copy, and S3 as PCR recombinants of the Numt and authentic mitochondrial copies. Consequently, S1 and S3 were excluded from the remainder of this study. The Numt (S1) sequence alignment can be seen in appendix 1.

3.3 Authentic mtDNA analyses

3.3.1 The cytochrome *b* fragment

The 251bp cytochrome *b* fragment was successfully amplified in 83 of the 89 available samples. A total of 14 variable sites, of which nine were parsimony informative, consisted of 13 transitions, 1 transversion which together defined 12 haplotypes. Haplotype diversity, nucleotide diversity and the average number of pairwise differences with standard deviations were 0.78 ± 0.03 , 0.013 ± 0.01 and 3.25 ± 1.69 respectively. Cytochrome *b* fragment base composition was 31.40% C, 24.96% T, 33.01% A and 10.64% G. The cytochrome *b* sequence alignment can be seen in appendix 2.

The distribution of these haplotypes among the various klipspringer subspecies is shown in Table 4. Of the 12 haplotypes found for the cytochrome *b* region, nine were unique to particular subspecies whilst three were shared among two to seven subspecies. Parsimony phylogenetic analysis suggested the presence of two groups consistent with approximate geographical structuring, although bootstrap support values were low, suggesting that there could be a number of other equally parsimonious branching patterns (Figure 7). These geographic groupings appear to correspond with samples from southern and eastern Africa. An exception is seen in HapG, which is found in Malawi yet clusters with the southern group. Interestingly, four of the subspecies tested showed no unique haplotypes, perhaps indicating that they are more recently derived or that incomplete lineage sorting has occurred.

Table 4: The distribution of cytochrome *b* haplotypes among the nine *Oreotragus* subspecies.

Numbers refer to the frequency of each haplotype in each subspecies.

	<i>O. o tyleri</i>	<i>O. o oreo</i>	<i>O. o transv</i>	<i>O. o centr</i>	<i>O. o acera</i>	<i>O. o schill</i>	<i>O. o aureus</i>	<i>O. o somal</i>	<i>O. o saltatr</i>
HapA		30							
HapB	2	11							
HapC	2	1	2	1	4	1			
HapD		3	2	2	1	8	1		1
HapE		1							
HapF			3						
HapG				1					
HapH		1							
HapI			1						
HapJ			1						
HapK							1		
HapL								1	

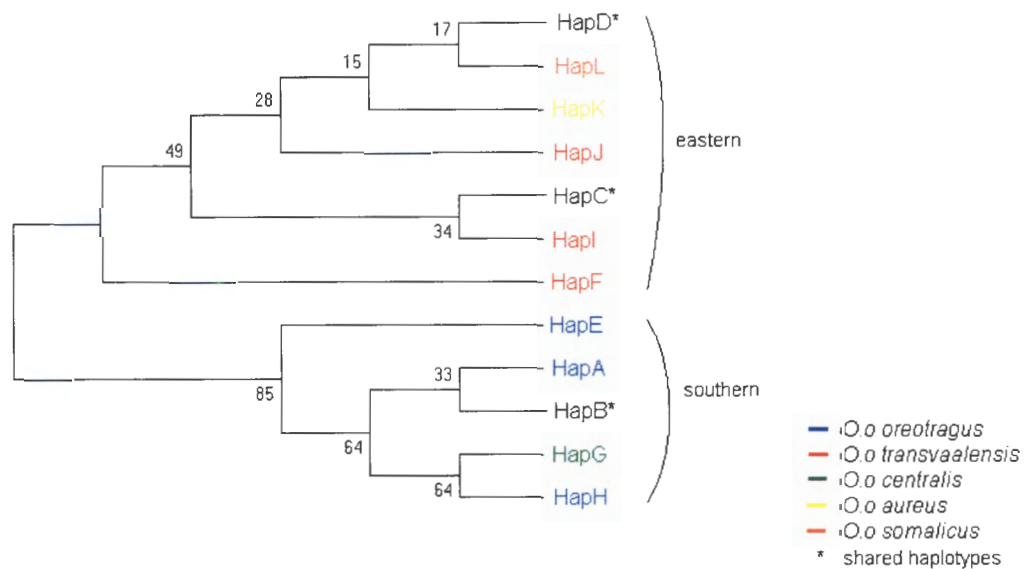


Figure 7: Parsimony tree of the 12 cytochrome *b* haplotypes of *Oreotragus oreotragus*. Bootstrap values are based on 1000 replications. Colours represent haplotypes unique to a particular subspecies.

The construction of a minimum-spanning network of the cytochrome *b* haplotypes supports the topology recovered by the parsimony analysis (Figure 8). More specifically, the presence of two groups is also suggested using this method, although these groups are separated by only three mutational steps. The groups correspond with the same approximate geographic southern and eastern groupings. Most of the haplotypes are just a single mutational step from each other, suggesting that the cytochrome *b* haplotypes of *Oreotragus* diverged fairly recently.

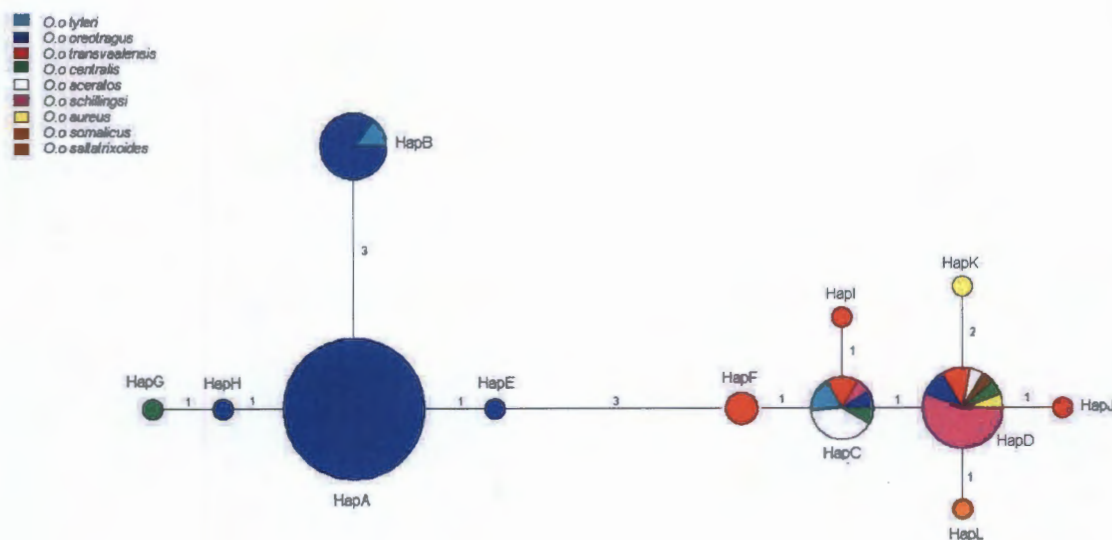


Figure 8: Minimum-spanning network of the 12 cytochrome *b* haplotypes identified. Circles are roughly proportional to haplotype frequency. The number of mutational steps between haplotypes is indicated on the diagram. Colours indicate specific subspecies in which the haplotype was observed.

3.3.2 The control region fragment

(i) Genetic variation and phylogenetic analysis

A final fragment of 423bp was amplified from 60 of the 89 available klipspringer samples. A total of 138 polymorphic sites consisting of 115 transitions, 11 transversions and 16 indels defined 37 haplotypes. Haplotype diversity, nucleotide diversity and the average number of pairwise differences with standard deviations were 0.98 ± 0.01 , 0.08 ± 0.04 and 34.26 ± 15.14 respectively. The control region fragment base composition was 20.67% C, 27.46% T, 40.40% A and 11.47% G. The control region sequence alignment can be seen in appendix 3.

Haplotypes 1-18 were found in *O. o oreotragus*, with Hap6 also occurring in *O. o transvaalensis*, *O. o centralis* and *O. o schillingsi*. Hap33 was found in both *O. o schillingsi* and *O. o aceratos*. All other haplotypes were unique to a particular subspecies. MODELTEST identified the HKY model of substitution with gamma distribution (HKY + G; gamma shape parameter 0.3170) as the best model of substitution for the control region data. Neighbour-joining and parsimony phylogenetic analyses consistently resolved three distinct groups among the klipspringer haplotypes (Figure 9a, 9b). The Bayesian analysis did not clearly elucidate the relationships among the three groups, but did present well-resolved terminal groups (Figure 9c), which were consistent with two of the three groups identified in the other methods. The three main groups appear to be roughly geographic in composition, with two groups representing southern/south-western Africa (S/SW1 and S/SW2) and one group representing eastern/north-eastern Africa (E/NE). Haplotypes from the east of South Africa occur in both the southern and eastern groups. No strong evidence for grouping according to subspecies designations was found based on the control region haplotypes. The S/SW2 and

E/NE groups form well-supported clades, while S/SW1 is more indicative of a group of lineages. The E/NE clade appears to consist of two lineages, the existence of which is reflected in all three phylogenetic analyses. An exception is seen with Hap37 (from Somalia) which clusters with the S/SW samples while being from the E/NE in the parsimony analysis. However, as it clusters with the E/NE samples in both other methods, there is greater support for it being in the E/NE group.

The phylogenetic trees were not rooted because the control region sequences from the proposed sister species, the duiker (*Cephalophus spp.*), found in GenBank were suspected to be Numts (see discussion). While Numts can be used as outgroups in phylogenetic analysis (Zhang & Hewitt 1996), it was not deemed appropriate in this case because the model of evolution that was used for the neighbour-joining analysis is unlikely to be the same model that would best fit the mode of evolution of the Numt. This hypothesis was supported when phylogenetic analyses done using the proposed Numt sequence as an outgroup showed that the Numt was not the most basal group in the tree (data not shown). This suggests that this was either a result of the differences in the evolution of the sequences, or that the Numt was more recently derived than the split of the *Oreotragus* lineages.

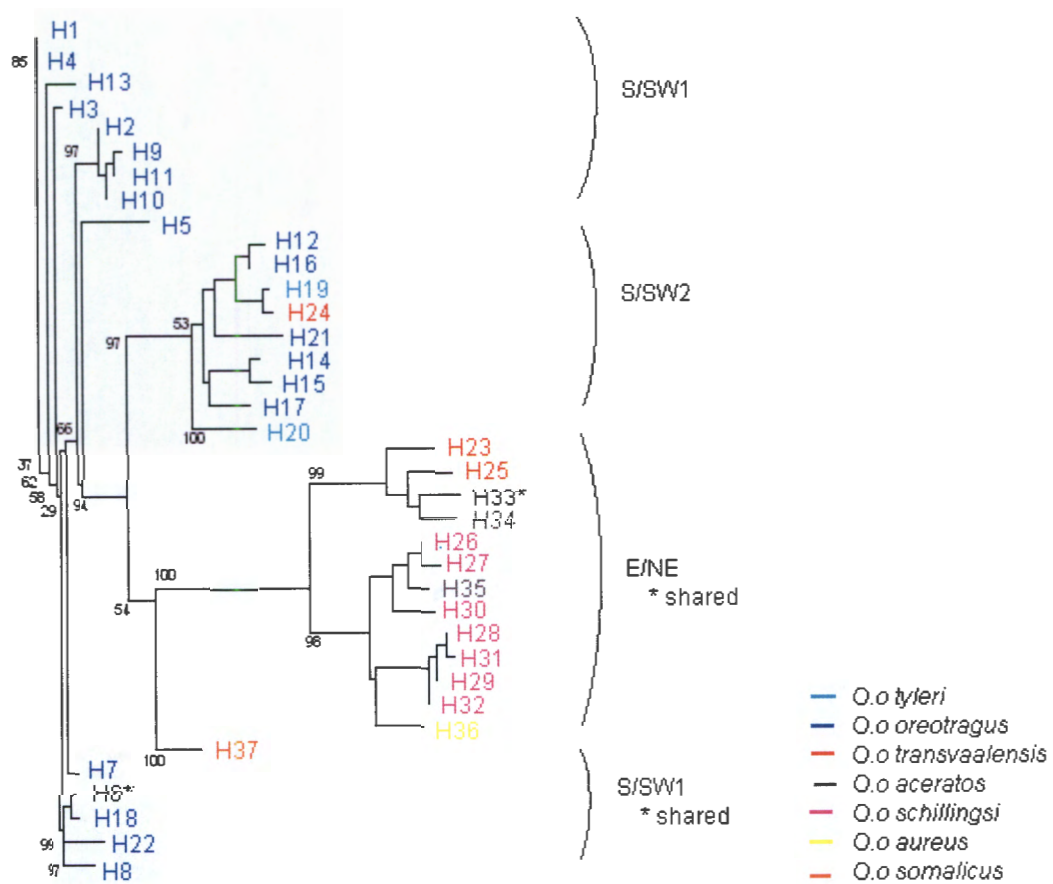


Figure 9(a): Unrooted neighbour-joining tree of the 37 mtDNA control region haplotypes of *Oreotragus oreotragus*, computed using HKY+G distances. Colours indicate haplotypes unique to specific subspecies. Bootstrap values based on 1000 replications are shown.

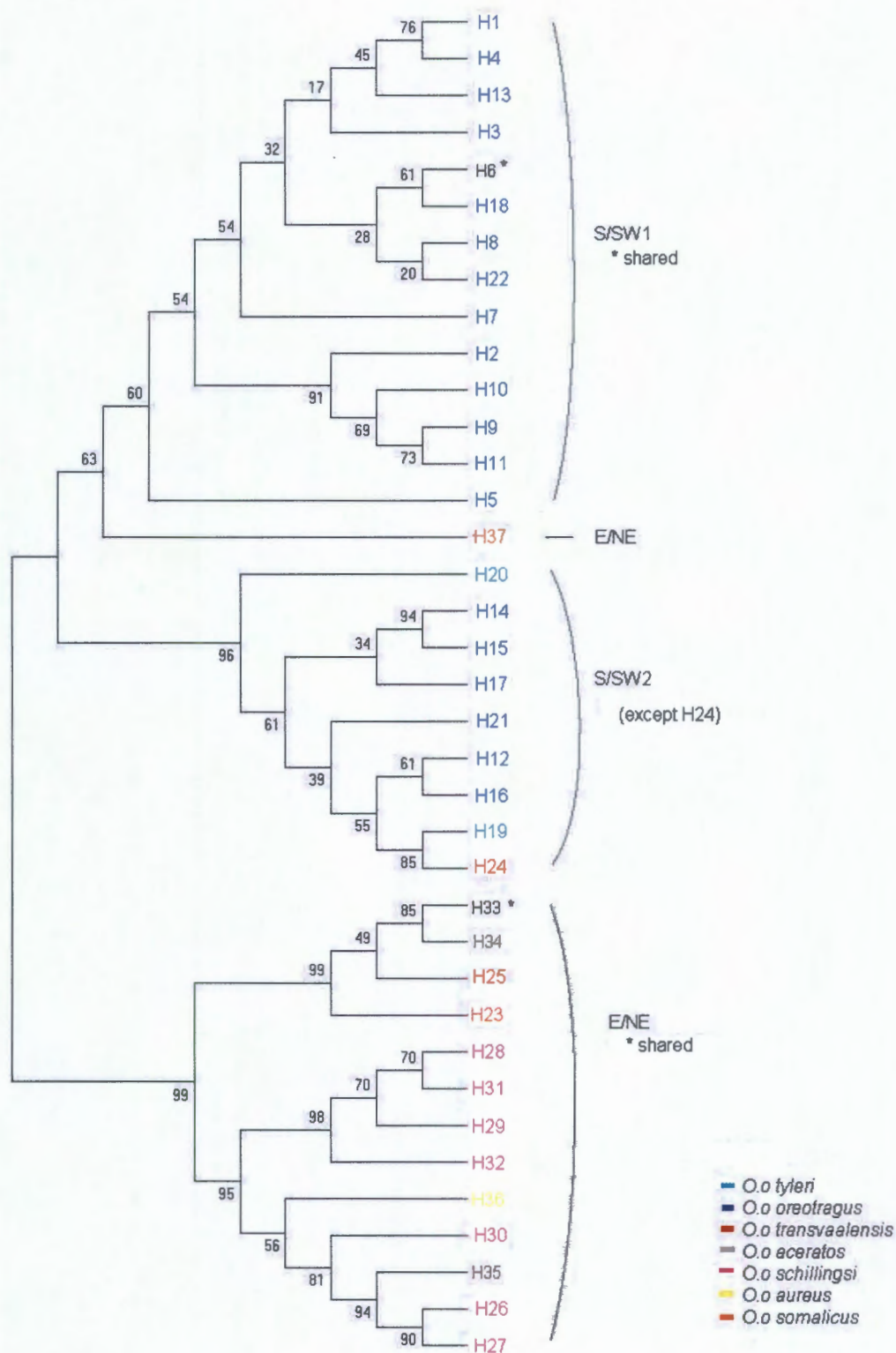


Figure 9(b): Consensus parsimony tree of the 37 mtDNA control region haplotypes. Bootstrap values are based on 1000 replications. Colours indicate haplotypes unique to specific subspecies.

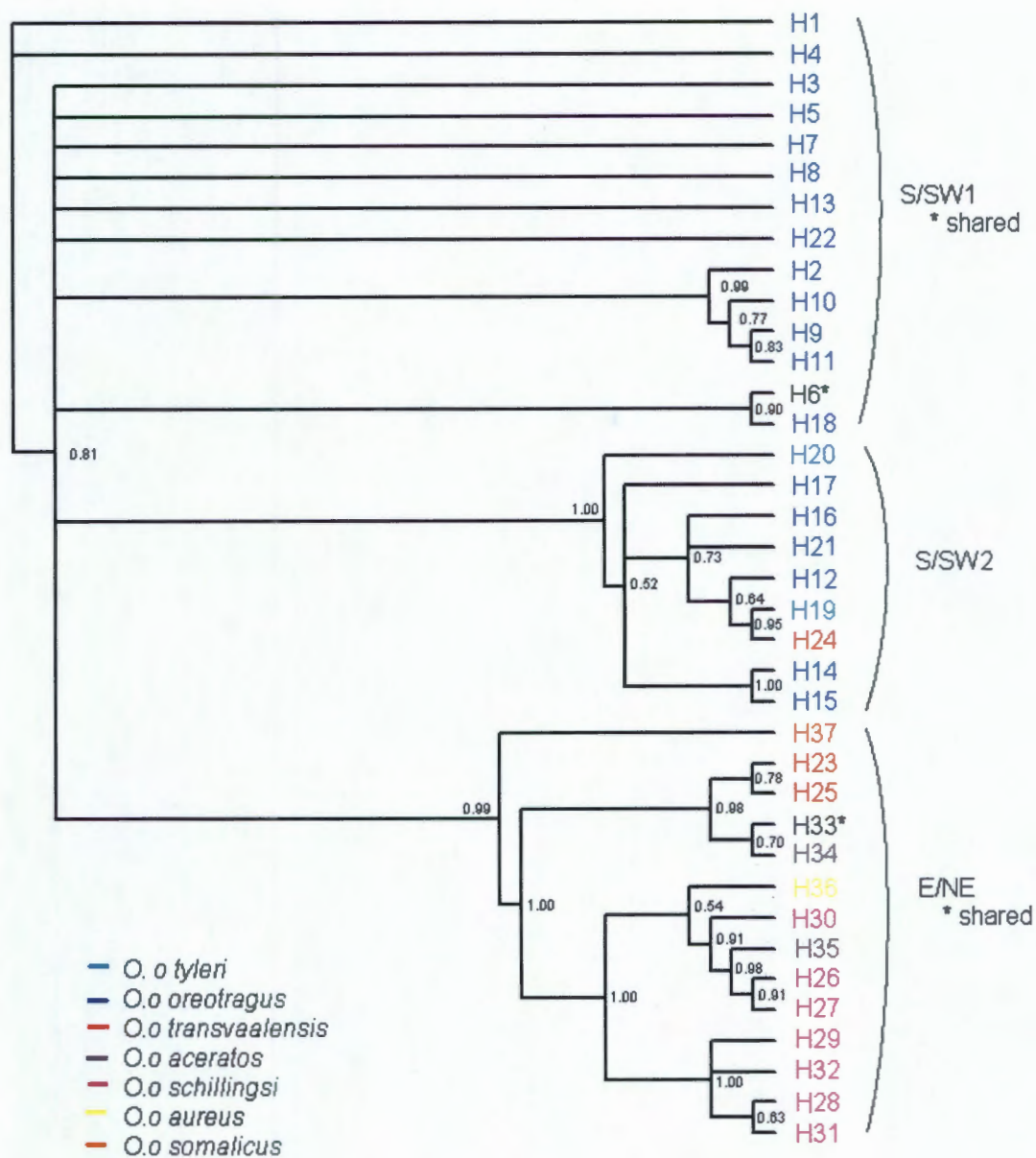


Figure 9(c): Bayesian tree of the 37 mtDNA control region haplotypes, computed using HKY+G distances. Colours indicate haplotype unique to a particular subspecies. Posterior probabilities are indicated.

The three geographic groups were evident in the minimum-spanning network (Figure 10). A sequence divergence of 11.4% separates the E/NE clade from the S/SW groups, and the S/SW1 and S/SW2 groups differ by approximately 5.4%. The occurrence of the two E/NE lineages can also be seen on the far left.

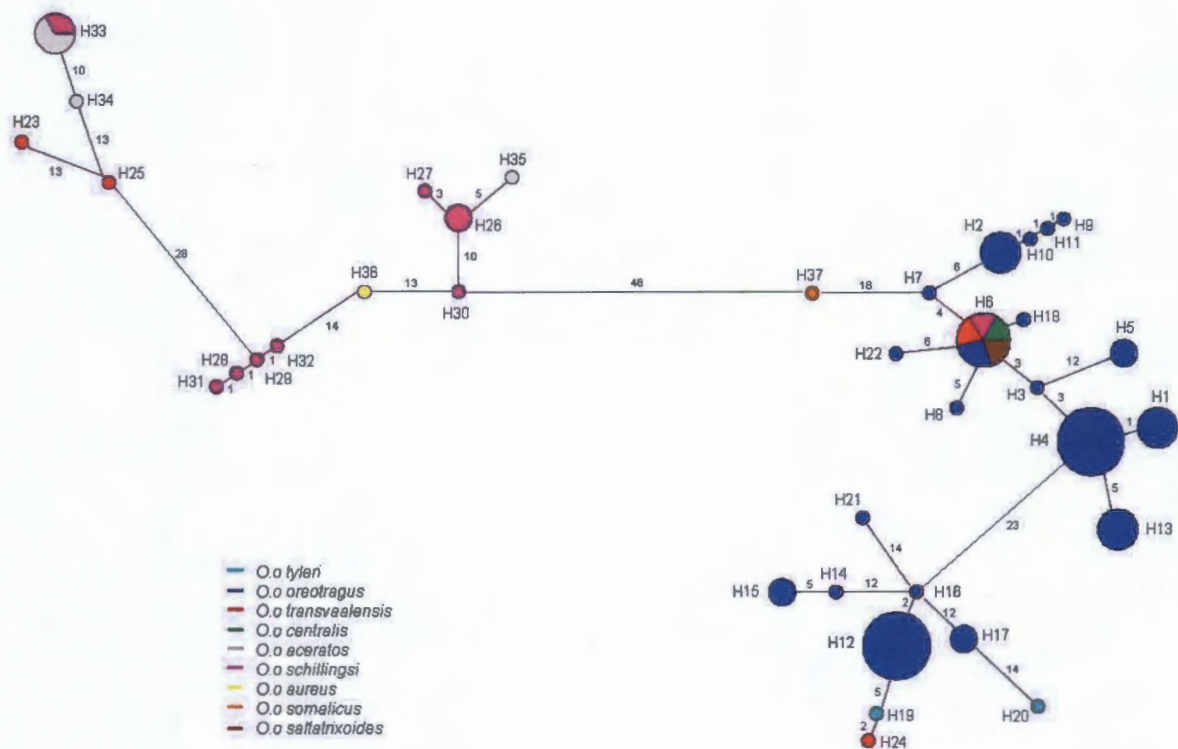


Figure 10: Minimum-spanning network of the 37 control region haplotypes. Circles are roughly proportional to haplotype frequency. Numbers next to branches indicate mutational steps. Colours indicate particular subspecies.

(ii) *Population structure*

Based on the results of the phylogenetic analyses, it seems most likely that the highest degree of population differentiation observed in the klipspringer is a result of the three geographic groups identified. In order to test this hypothesis, a number of possible geographic and subspecies groupings were investigated using AMOVA. In the first scenario, individuals were separated into the two broad geographic groups, S/SW and E/NE, identified by the phylogeographic analyses. The S/SW group comprises individuals from South Africa (S/SW) and Namibia. The E/NE group comprises individuals from South Africa (East), Malawi, Tanzania, Kenya, Somalia and Ethiopia. In the second scenario, the above divisions were maintained with the exception of South Africa (East) which was excluded as its appearance in both the S/SW and E/NE clades makes its placement uncertain. In the third scenario the two S/SW groups were compared. In the fourth scenario, subspecies represented by four or more individuals were compared. These subspecies were *O. o oreotragus*, *O. o transvaalensis*, *O. o schillingsi* and *O. o aceratos*. In all scenarios a reduced subset of the haplotypes of *O. o oreotragus* was used (nine of the original 20 haplotypes), in order to prevent the results being influenced by the disproportionate sample size of this subspecies. Numerous other possible scenarios of population subdivision were considered, but could not be tested due to lack of samples for particular areas, or lack of precision in the exact geographic origins of samples.

The analysis of population differentiation showed similar levels of variation within and among the two geographical groups used in scenarios one and two (Table 5). The highest differentiation was seen in the third scenario, using the three geographic clades, with an F_{ST} value of 0.70, indicating that 70% of genetic variation seen was explained by variation among the groups. The most differentiated pair of groups among these were S/SW1 and E/NE, which

had a significant F_{ST} of 0.74. The next highest level of differentiation occurred in scenario four, at the subspecies level. The percentage of sequence variation found within and among the subspecies was 39.32% and 60.68% respectively, with a significant F_{ST} value of 0.61. Significant differentiation was seen for all pairs of subspecies compared, with the exception of *O. o transvaalensis* and *O. o aceratos* for which no significant differentiation was found. Levels of differentiation appeared to increase, from moderate to high, between pairs of subspecies of increasing geographic distance, with the highest differentiation occurring between *O. o oreotragus* and *O. o schillingsi*.

Table 5: Measures of population differentiation based on the AMOVA approach using mtDNA control region data. The abbreviations are as follows: Oreo = *O. o oreotragus*, Trans = *O. o transvaalensis*, Acera = *O. o aceratos* and Schill = *O. o schillingsi*.

Subdivision	% variation within	% variation among	F_{ST}
2 groups (S/SW & N/NE)	55.35	44.65	0.45**
2 groups (excl. SA - east)	48.15	51.85	0.52**
3 groups S/SW1 vs. S/SW2	30.06 37.34	69.94 62.66	0.70** 0.63**
S/SW1 vs. E/NE	25.99	74.01	0.74**
S/SW2 vs. E/NE	32.62	67.38	0.67**
4 groups	39.32	60.68	0.61**
<i>Oreo</i> vs. <i>Trans</i>	56.95	43.05	0.43*
<i>Oreo</i> vs. <i>Acera</i>	22.50	77.50	0.77**
<i>Oreo</i> vs. <i>Schill</i>	20.90	79.10	0.79**
<i>Trans</i> vs. <i>Acera</i>	84.83	15.17	0.15 ^N
<i>Trans</i> vs. <i>Schill</i>	64.05	35.95	0.36**
<i>Acera</i> vs. <i>Schill</i>	70.97	29.03	0.29*

*significant ($P < 0.05$); ** significant ($P < 0.001$)

^N not significant

3.3.3 Subset analysis: *O. o oreotragus*

(i) Genetic Variation and Phylogenetic Analysis

The control region data for the subspecies *O. o oreotragus* was used for an in-depth phylogeographic analysis in order to further elucidate the recent evolutionary history at the subspecies level. This particular subspecies was used as it was the only subspecies with enough samples for such an investigation. The 36 samples in this group contained a total of 68 polymorphic sites consisting of 65 transitions, 1 transversion and 3 indels, and comprised 20 haplotypes. Haplotype diversity, nucleotide diversity and the average number of pairwise differences with standard deviations were 0.96 ± 0.02 , 0.04 ± 0.02 and 17.49 ± 7.96 respectively. Nucleotide composition was 20.75% C, 27.25% T, 41.05% A and 10.95% G.

The *O. o oreotragus* individuals were grouped according to their approximate sampling location for the population differentiation analysis. These groups were Paarl (n=7), Tulbagh (n=7), Cape Point (n=5), Karoo (n=14) and Northern Cape (n=3). Although the Cape Point group is included, the actual origin of these samples is unknown as they were translocated to the Cape Point area. It is suspected that they are of Porterville origin, although some may form part of the Tulbagh or Paarl groups. Network construction showed the existence of two groups separated by approximately 5% sequence divergence (Figure 11). This mirrors what was found in the entire dataset. Group one consists of haplotypes from all five sampling locations, while group two is dominated by Karoo haplotypes, but also includes one haplotype from the Northern Cape and two haplotypes found in both the Karoo and Cape Point populations. This pattern is suggestive of a movement or colonisation from the south-western Cape (Tulbagh/Paarl/Cape Point) towards the Karoo, as a lower haplotype diversity is expected in

areas that have been colonised more recently. This is consistent with the results of the entire dataset, which suggest a general south to east colonisation pattern, with the divergence of the two S/SW groups.

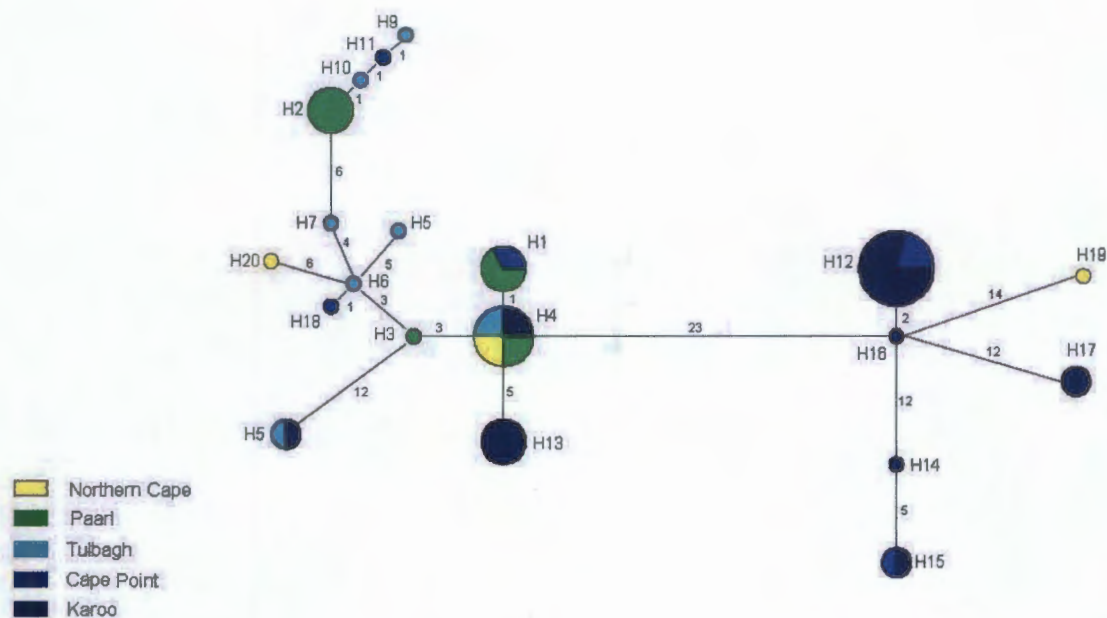


Figure 11: Minimum-spanning network of the 20 *O. o. oreotragus* haplotypes. Circles are roughly proportional to haplotype frequency. Numbers indicate the mutational steps between haplotypes.

(ii) Population structure

Pairwise F_{ST} values showed significant differentiation between only the Karoo and Paarl, and Karoo and Tulbagh populations, with modest F_{ST} values of 0.34 and 0.31 respectively (Table 6). All other population pairwise comparisons yielded non-significant results. This is

consistent with the results of the minimum-spanning network, which suggests a separation between the two areas, and movement from the Cape eastwards towards the Karoo.

Table 6: *O. o oreotragus* population pairwise F_{ST} values.

	Paarl	Tulbagh	Cape Point	Karoo
Paarl	-			
Tulbagh	-0.02 ^N	-		
Cape Point	0.12 ^N	0.09 ^N	-	
Karoo	0.34*	0.31*	-0.03 ^N	-
N Cape	0.19 ^N	0.10 ^N	-0.17 ^N	0.03 ^N

* significant ($P < 0.05$)

^N not significant

Chapter 4

Discussion

4.1 Identification of authentic mtDNA control region sequences

4.1.1 Identification of the Numt sequences

Over the past decade, reports of nuclear mitochondrial-like sequences identified in various taxa have steadily increased in number. The list of identified Numts extends from fungi to molluscs, to insects, birds and mammals (Bensasson *et al.* 2001). Despite increased awareness concerning the prevalence of Numts, they frequently remain undetected and thus confound the results of phylogenetic and phylogeographic studies (Anthony *et al.* 2007; Zhang & Hewitt 1996). In this study, the generation of three distinctly different sequence types, by using combinations of primer pairs, suggested that at least one of these sequence types was possibly nuclear in origin. Tissue type has been shown to be a factor in the preferential or co-amplification of Numt sequences (Bensasson *et al.* 2001), with blood identified as a specific problem in some cases. e.g. in birds (Sorenson & Quinn 1998). In addition, conserved primers have a much higher probability of generating Numt sequences, as because Numts evolve at a slower rate compared to the authentic mtDNA copies, they are more similar to the ancestral form of the gene in question and therefore have a higher homology to conserved primers (Bensasson *et al.* 2001; Zhang & Hewitt 1996). These two factors suggest that the S1

sequence type, which was derived from blood samples using conserved primers, is most likely an *Oreotragus* Numt. Closer investigation into the properties and variation of the different sequence types confirmed that S1 exhibited far less variation than the other two, and this is expected if the sequence is evolving at a slower rate. In addition, when published Numt sequences were included in the phylogenetic analysis, the S1 and S3 sequences formed a monophyletic clade with the Numts. Conclusive proof could be obtained by performing a mtDNA extraction from tissue that is rich in mitochondria, such as heart or liver tissue, but this was not possible in this study due to the lack of availability of fresh or frozen klipspringer tissue of these types. Nonetheless, the combination of the above evidence points conclusively to the S1 sequence type as a Numt.

4.1.2 Identification of PCR recombinants

Following the identification of the Numt S1 sequence type, the origin of both the S2 and S3 sequence types remained unexplained. Both sequence types exhibited high levels of variation, and could therefore be the authentic mitochondrial copy, yet were also highly dissimilar to each other. *In vitro* recombination between different sequence types has been shown to result in the generation of chimaeric PCR products in gorillas (Anthony *et al.* 2007), and this phenomenon is often found to occur in the presence of degraded DNA (Pääbo *et al.* 1990). This suggests that the S3 (dung) sequence type might be a potential PCR recombinant given that the DNA extracted from the dung was extremely degraded. Furthermore, the S2 (blood) amplification was consistent between samples and in repeated PCR experiments, while S3 amplification was fairly unpredictable. Recombination detection analysis performed on the three sequence types conclusively detected recombination by two methods implemented in the software, but the identification of the daughter recombinants was ambiguous. The results

identified either S1 or S3 as the daughter recombinants of the other two sequence types. In conjunction with the evidence for the S1 sequence type as a Numt, it seems most likely that S3 is a PCR recombinant. In addition, the method obtained to produce these sequences (the nested PCR, initially using conserved primers) might increase the chance of producing recombinants when both nuclear and mitochondrial copies of a gene exist. As a final check, sequence-specific primers were designed to S2 and S3 under the assumption that it should be possible to amplify the authentic sequence, but not the recombinant, in the other tissue type. Consistent with all the other evidence, the S3 primers failed to yield results with the blood samples while the S2 primers generated a product using the dung samples, convincingly supporting the S2 sequence type as the authentic mtDNA copy and the S3 sequence type as the PCR recombinant.

4.1.3 Numts, recombinants and sequence databases

In light of the previously mentioned technical difficulties, it seems clear that the problems of nuclear insertions and PCR recombinants can be far more serious than previously anticipated. Results obtained from the use of a single tissue type and/or conserved primers should be viewed with caution and should be double-checked using an alternative tissue type or a full mtDNA extraction if suspicious results are obtained. In the presence of both nuclear and mitochondrial gene copies, the generation of PCR recombinants could be prevented by avoiding nested PCR techniques. Ironically, nested PCR techniques are often the most successful approach when using degraded DNA (Niesen & Arctander 2001). Alternatively, the possibility of recombinants should be checked for using some type of recombination detection analysis. Furthermore, placing confidence in the sequences found in public sequence databases, such as GenBank, should be done with caution. When the three sequence types

obtained in this study were blasted against those in GenBank, all three were identified as mitochondrial control region sequences. This suggests that ostensibly authentic mitochondrial control region sequences in Genbank might actually comprise Numts, PCR recombinants and valid sequences. More specifically, when the S1 Numt sequences were blasted, they returned matches to Tibetan antelope (*Pantholops hodgonsii*), moose (*Alces alces*) and white-bellied duiker (*Cephalophus leucogaster*). In addition, an identified Numt sequence from GenBank, Numt1bos (which was used for comparison in this study), may be a PCR recombinant. The Numt1bos sequence was found to cluster with the recombinant S3 sequence type, rather than with the Numt S1 type, in the phylogenetic analysis. In the absence of verified sequences, future studies could benefit from careful interpretation of the results from BLAST searches until full examination of the existing mitochondrial sequences has been made.

4.2 Mitochondrial DNA analysis

4.2.1 Genetic variation

The level of genetic diversity found within *Oreotragus spp.* is extremely high, with haplotype diversity for the cytochrome *b* and control region at 0.78 and 0.98 respectively. These values, although high, are comparable to the degree of diversity found in a number of other bovid species, including Tibetan gazelle (*Procapra picticaudata*; Zhang & Jiang 2006), Grant's gazelle (*Gazella granti*; Arctander *et al.* 1996), African buffalo (*Syncerus caffer*; Simonsen *et al.* 1998) and kob (*Kobus kob*; Birungi & Arctander 2000). Tibetan gazelle show an astonishing level of variation, with nucleotide and haplotype diversities of 0.08 and 0.98 respectively (Zhang & Jiang 2006). Likewise, Grant's gazelle, buffalo and kob exhibit

haplotype diversities ranging from 0.88 to 0.95 (Simonsen *et al.* 1998), which suggests that some commonality among bovids with varying life history traits could be involved in generating and maintaining this high degree of variation.

High diversity levels can be a function of a historically large population size (Balakrishnan *et al.* 2003), which is consistent with what little is known about the klipspringer's history and distribution. Despite relying on rocky, mountainous areas for safety, the distribution of the klipspringer is extensive, with fair-sized populations existing in many countries within and north of the southern African subregion (Skinner & Chimimba 2005). Much of the current haplotype diversity could be the result of the strong structure identified between subspecies, which was seen with only two out of 37 control region haplotypes shared between subspecies. Structuring at this level, however, must be interpreted with caution due to the vast differences in sample numbers that were available for the different subspecies. Interestingly, *O. o oreotragus* is known to have been through a severe bottleneck around 100 years ago, that resulted in its eradication from various parts of its range, yet it still exhibits a high level of diversity (0.96 ± 0.02). This phenomenon was also seen in buffalo, which were found to show a remarkable level of variability (haplotype diversity = 0.98) despite apparent decimation by rinderpest in the late 19th century (Simonsen *et al.* 1998). Similarly, the Tibetan gazelle has been hunted extensively and has lost a large proportion of its habitat, yet still maintains a haplotype diversity of 0.98 (Zhang & Jiang 2006). This suggests that either the population reduction in *O. o oreotragus* was not as severe as previously thought, given the population size, or that insufficient time has past since the onset of the population reduction to erode levels of genetic variation in this subspecies (Balakrishnan *et al.* 2003).

4.2.2 Phylogeography

Analysis of the control region data within a phylogeographic framework supports the existence of three main lineages within the genus *Oreotragus*. These groups are not comprised of specific subspecies, but rather consist of haplotypes grouped according to approximate geographic location. More specifically, the two most strongly supported clades correspond to a south/south-western Africa clade (S/SW2), and an eastern/north-eastern Africa clade (E/NE), with the S/SW2 clade consisting of haplotypes from the Western Cape of South Africa and Namibia, and the E/NE clade containing haplotypes from Gauteng/KwaZulu Natal in South Africa all the way to Ethiopia. The third group (S/SW1) is comprised exclusively of S/SW haplotypes. This group appears to be more basal than the other two clades suggesting a cluster of more ancient lineages, although this could not be confirmed, as the phylogenetic trees could not be rooted due to the lack of an appropriate outgroup. The S/SW and E/NE lineages appear paraphyletic; that is, the E/NE lineage forms a monophyletic group, which is nested within the lineages of the S/SW group (Arctander *et al.* 1999). This pattern is often characteristic of a colonisation event, and in this situation suggests that klipspringer colonised Africa from the south/south-west, moving eastwards (Arctander *et al.* 1999). Alternatively, this structure could be attributed to variation in population size between the geographical areas. If a larger population existed in the south/south-west, it would be far more likely to retain ancestral polymorphisms than would a significantly smaller eastern population (Avise *et al.* 1984).

The afore-mentioned pattern of geographical structure is further corroborated by the fact that it is not unique to the klipspringer, but is found in a number of other African bovid species such as wildebeest (*Connochaetes taurinus*; Arctander *et al.* 1999), common impala and

greater kudu (Nersting & Arctander 2001). In wildebeest, while no clear clustering pattern between subspecies was found, clear geographical subdivisions were identified. More specifically, these subdivisions were mainly southern and eastern groupings of mitochondrial haplotypes, suggesting that colonisation of new habitats by wildebeest also occurred from the south to the east of Africa (Arctander *et al.* 1999). Similar findings were observed in impala and kudu. The greater kudu exhibits distinct south-western and eastern groupings, in addition to decreasing level of diversity among populations moving from south-western to eastern Africa (Nersting & Arctander 2001). The impala do not exhibit such clearly defined geographic structure, but do show the decrease in genetic diversity along the same south-west to east path. Evidence from studies on the impala and kudu indicates that populations may have been isolated in the south-west for enough time for lineages to sort (Nersting & Arctander 2001). Unfortunately, this pattern of decreasing genetic diversity could not be tested for in the klipspringer due to insufficient sample numbers.

Congruent patterns of genetic differentiation, with respect to geographic location, in numerous African bovid studies suggest that a south/south-west African refugium may have existed during the late Pliocene/early Pleistocene, approximately 2 million years ago (MYA) (Nersting & Arctander 2001). While comparatively little is known about the paleoclimatic history of southern Africa relative to East Africa, fossil record analysis indicates many first and last appearances of bovid lineages between 2.7-2.5 MYA, with secondary waves around 1.8 and 0.7 MYA (deMenocal 2004). Many of the bovid lineages appearing at these times were arid-adapted species, and correlate with three intervals, 2.9-2.4 MYA, 1.8-1.6 MYA and 1.2-0.8 MYA, when climatic shifts resulted in drier conditions throughout Africa (deMenocal 2004). Changes in the composition of the vegetation, specifically the increase of savannah grasslands and the decrease in closed forest and woodland areas during these intervals are

thought to have influenced the expansion and diversification of arid-adapted species (Bobe & Behrensmeyer 2004; deMenocal 2004). While klipspringer may not appear obviously adapted to arid conditions because they are browsers rather than grazers, Antilopinae browsers (to which klipspringers belong) have been shown to occupy arid environments in which the leaves of arid-adapted plants and bushes constitute a large part of their diet (Bobe 2006; Bobe & Behrensmeyer 2004). Although they predominantly inhabit rocky outcrops, klipspringer move down to flatter areas at night to feed, and they are also water-independent, i.e. they do not depend on free water sources for survival (Skinner & Chimimba 2005). All of the above-mentioned information suggests that an expansion from south/south-western Africa to east/north-eastern Africa in the klipspringer may have occurred during the expansion of the savannah grasslands during the late Pliocene/early Pleistocene. It is also possible that fluctuations between wet and dry conditions during the Pliocene, and the resultant expansion and contraction of habitats and populations (Bobe 2006), could have caused the isolation and then secondary contact between the two south/south-western groups seen in this study.

The degree to which the three klipspringer lineages are differentiated is highly significant, with 11% divergence between the S/SW1 and E/NE groups, and 6% divergence between the two S/SW groups. While the level of differentiation between the S/SW and E/NE in particular is extremely high, it is comparable to some of the highest levels of intraspecific variation previously reported in bovids. These include 12-14% in Grant's gazelle (Arctander *et al.* 1996), 10% in kob (Birungi & Arctander 2000) and 9% intraspecific sequence divergence within chimpanzees (*Pan troglodytes*; Morin *et al.* 1994). This degree of sequence divergence is most likely to be observed after mtDNA genotypes have been separated for long periods of time by physical/geographic barriers to gene flow or by large geographic distance (Morin *et al.* 1994); similar levels are also observed after secondary contact between previously

separated populations (Arctander *et al.* 1996). It seems most likely that the two S/SW groups were separated for a period of time due to variable climatic conditions during the Pliocene, and after expansion and colonisation to the east, the large difference has been maintained by the geographic distance between the S/SW and E/NE, as well as the limited dispersal of the species. Mismatch distribution analysis (Rogers and Harpending 1992) did not conclusively show evidence of population expansion in the *O. o. oreotragus* population (data not shown). The data was not included as the strength of this analysis was confounded by the difficulty in assigning individuals into groups to be tested.

4.2.3 Population differentiation and subspecies designations

Of the 11 klipspringer subspecies described by Ansell (1972), nine are represented in this study. Disproportionate sampling across the different subspecies does, however, place limitations on the conclusions of this study regarding molecular support for morphological subspecies designations. Nevertheless, it is interesting to note that 35 out of 37 control region haplotypes and nine out of 12 cytochrome *b* haplotypes were found to be unique to a particular subspecies. In particular, the cytochrome *b* results suggest structuring at a deep level, as the rate of evolution of cytochrome *b* in bovids is approximately 5-10 times slower than that of the control region (Zhang & Jiang 2006). Conversely, the results of the phylogeographic analyses did not support the clustering of haplotypes within designated subspecies, but rather the existence of geographic groups. Subsequent results of the AMOVA analysis reflected that while a high level of differentiation is seen at the subspecies level, with a F_{ST} of 0.61, even greater differentiation overall is seen when partitioning the data into three geographic groups ($F_{ST} = 0.70$). It is also possible that the occurrence of unique and fairly differentiated

haplotypes within different subspecies is simply a function of the lack of samples, both in location and in numbers.

When pairs of subspecies were compared, all appeared significantly differentiated from one another, with the exception of *O. o. transvaalensis* and *O. o. aceratos*, which were not significantly different. While the distributional boundaries of the ranges of these two subspecies are not clearly defined, it is thought that both inhabit parts of Mozambique, and thus it seems possible that they could be considered as one subspecies. In addition, *O. o. aceratos* did not exhibit any unique cytochrome *b* haplotypes, although this was also true of two other subspecies, *O. o. schillingsi* and *O. o. centralis*. The degree of differentiation between other subspecies pairs was found to increase along with increasing geographic distance. Each subspecies was most similar to its nearest neighbour, suggesting that differentiation between the subspecies is a result of an isolation-by-distance mechanism (Slatkin 1993). This corresponds with the life history of the klipspringer, in that the species is known to pair-bond, exhibit strong territoriality and is possibly fairly restricted in dispersal by its need for rocky terrain (Skinner & Chimimba 2005). Furthermore, klipspringer breeding seasons are reported as variable, based on local climatic conditions (Norton 1980). This trait could help to explain the degree of differentiation observed between subspecies, especially those which exhibit local climatic differences, as asynchronous breeding seasons would have a profound effect of the ability of the regionally adapted klipspringer to interbreed and thus reinforce differentiation.

While a more extensive sampling of this species would provide greater insight into its recent evolutionary history, the current results presented in this study that include phylogenetic and phylogeographic analyses, along with measures of population differentiation, suggest that

structuring amongst the klipspringer is shaped by geographic location, at a local and broader level, rather than by subspecies (and thus morphological) designations.

4.2.4 Implications for Conservation

Despite its relative abundance and extensive distribution, the klipspringer has been eradicated in parts of its natural range (Skinner & Chimimba 2005). Of the 11 designated subspecies, only *O. o. porteousi*, from Nigeria, has received threatened status (Ansell 1972). Nonetheless, with the reintroduction of klipspringer back into areas which it previously occupied, such as Table Mountain National Park, it is important to assess the basis of the subspecies designations as well as investigate the determinants of finer-scale population structure within this species. Moreover, the identification of evolutionary significant units (ESUs) and management units (MUs) within the klipspringer could aid in the selection of appropriate individuals for translocations (Moritz 1994).

Based on the criteria set by Moritz (1994), klipspringer could be classified as three management units in accordance with their geographic/genetic groupings, with particular emphasis on the more recently diverged S/SW2 and E/NE clades. The S/SW1 group represents an ancestral population pool comprised of a number of different lineages, but could still be considered as a MU based on differences in haplotype frequency. With respect to the translocation of individuals into Table Mountain National Park, the pattern of isolation by distance observed among the klipspringer lineages, along with certain life history traits, such as local variation in breeding seasons, suggests that individuals should be sourced from neighbouring populations. Increasing levels of genetic diversity of the introduced population need not be a priority, as the level of variation exhibited in the klipspringer is already high.

A number of factors influence the success of translocations, in addition to the source population used, and these include the number of animals released, the quality of the release habitat, the location of the release habitat in relation to the historical range of the species, and whether the original cause of decline has been removed (Rout *et al.* 2007). Examples of successfully translocated species include the white stork (*Ciconia ciconia*; Schaub *et al.* 2004), the alpine ibex (*Capra ibex*; Beck *et al.* 1994), as well as various game species (Griffith *et al.* 1989). However, while translocations are widely utilised in conservation management, there are a number of possible detrimental effects which can be associated with this process. These include the reduction of fitness due to outbreeding depression, genetic introgression, and the introduction of exotic pathogens (Storfer 1999). For example, Galindo-Leal & Weber (1994) showed that the movement of white-tailed deer (*Odocoileus virginianus*) across subspecies boundaries to stock captive breeding programs resulted in the death of the offspring due to difficulties in the birth process. Dystocia, or abnormal or difficult labour, is the most common reproductive problem among artiodactyls in zoos, and frequently results from size or proportion differences between mating individuals (Galindo-Leal & Weber 1994). Thus care must be taken when conceiving and implementing conservation strategies such as translocations, particularly when differentiated subspecies are involved.

4.2.5 Conclusion

In conclusion, the major finding of this study was that two major klipspringer lineages exist; a south/south-western (S/SW) and east/north-eastern (E/NE) group, with the S/SW group being further split into two divergent groups. This phylogeographic structure strongly suggests revealed that the predominant factors influencing the structure observed in this species are

geographic location and an isolation-by-distance mechanism, rather than subspecies differentiation. The data suggests that klipspringer originated in south/south-western Africa and after a period of isolation and population differentiation, probably caused by climatic shifts during the late Pliocene/early Pleistocene, colonised east/north-eastern Africa. Local differentiation between subspecies is most likely a function of isolation-by-distance (or nearest-neighbour mating), as well as specific life history traits, such as limited dispersal and local variation in breeding seasons. The crucial conservation implication of this study is that any translocations should occur within the geographic groups identified, preferably between populations that are as close as possible within these groups, in order to avoid outbreeding depression. Furthermore, the genetic variation found in the klipspringer was high, thus maintaining or increasing diversity in reintroduced populations need not be a major concern at the present time.

References

- Amos, W. & Balmford, A. (2001) When does conservation genetics matter? *Heredity*. **87**, 257-265.
- Amos, W. & Harwood, J. (1998) Factors affecting levels of genetic diversity in natural populations. *Philosophical Transactions: Biological Sciences*. **353**(1366), 177-186.
- Ansell, W.F.H (1971) Artiodactyla. In: *The mammals of Africa: an identification manual* (Eds. Meester, J. & Setzer, H.W.) Smithsonian Institution Press, Washington D.C.
- Anthony, N.M., Clifford, S.L., Bawe-Johnson, M., Abernethy, K.A., Bruford, M.W. & Wickings, E.J. (2007) Distinguishing gorilla mitochondrial sequences from nuclear integrations and PCR recombinants: Guidelines for their diagnosis in complex sequence databases. *Molecular Phylogenetics and Evolution*. **43**, 553-566.
- Antunes, A. & Ramos, M.J. (2005) Discovery of a large number of previously unrecognised mitochondrial pseudogenes in fish genomes. *Genomics*. **86**, 708-717.
- Arbogast, B.S. & Kenagy, G.J. (2001) Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*. **28**, 819-825.
- Arctander, P., Johansen, C. & Coutellec-Vreto, M. (1999) Phylogeography of three closely related African bovids (tribe Alcelaphini). *Molecular Biology and Evolution*. **16**(12), 1724-1739.
- Awise, J.C. (1998) The history and purview of phylogeography: a personal reflection. *Molecular Ecology*. **7**, 371-379.
- Awise, J. C. (2000) *Phylogeography*. Harvard University Press, Cambridge.
- Awise, J.C., Arnold, R.M., Ball Jr, E., Bermingham, T., Lamb, J.E., Neigel, C.A., Reeb, A. & Saunders, N.C. (1987) Intraspecific phylogeography: the mitochondrial DNA bridge

- between population genetics and systematics. *Annual Review of Ecology and Systematics*. **18**, 489-522.
- Balakrishnan, C.N., Monfort, S.L., Gaur, A., Singh, L. & Sorenson, M.D. (2003) Phylogeography and conservation genetics of Eld's deer (*Cervus eldi*). *Molecular Ecology*. **12**, 1-10.
- Ballard, J.W.O. & Dean, M.D. (2001) The mitochondrial genome: mutation, selection and recombination. *Current Opinion in Genetics & Development*. **11**, 667-672.
- Barrowclough, G.F. (1982) Geographic variation, predictiveness, and subspecies. *The Auk*. **99**, 601-603.
- Beebee, T. & Rowe, G. (2004) Introduction to molecular ecology. Oxford University Press.
- Beck, B.B., Rapaport, L.G., Price, M.R.S & Wilson, A.C. (1994) Reintroduction of captive-born animals. In: *Creative Conservation: Interactive Management of Wild and Captive Animals* (Eds. Olney, P.J.S, Mace, G.M. & Feistner, A.T.C). Chapman & Hall, London.
- Bensasson, D., Zhang, D-X. & Hewitt, G.M. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution*. **16**(6), 314-321.
- Bernatchez, L. & Danzmann, R.G. (1993) Congruence in control-region sequence and restriction-site variation in mitochondrial DNA of brook charr (*Salvelinus fontinalis* Mitchill). *Molecular Biology and Evolution*. **10**, 1002-1014.
- Birungi, J. & Arctander, P. (2000) Large sequence divergence of mitochondrial DNA genotypes of the control region within populations of the African antelope, kob (*Kobus kob*). *Molecular Ecology*. **9**, 1997-2008.
- Bobe, R. (2006) The evolution of arid ecosystems in eastern Africa. *Journal of Arid Environments*. **66**, 564-584.

- Bobe, R. & Behrensmeyer, A.K. (2004) The expansion of grassland ecosystems in Africa in relation to mammalian evolution and the origin of the genus *Homo*. *Palaeogeography, Palaeoclimatology, Palaeoecology*. **207**, 399-420.
- Boore, J.L. (1999) Animal mitochondrial genomes. *Nucleic Acids Research*. **27**(8), 1767-1780.
- Brown, G.G., Gadaleta, G., Pepe, G., Saccone, C. & Sbisà, E. (1986) Structural conservation and variation in the D-loop containing region of the vertebrate mitochondrial DNA. *Journal of Molecular Biology*. **192**, 503-511.
- Burton, R.S. (1998) Intraspecific phylogeography across the point conception biogeographic boundary. *Evolution*. **52**(3), 734-745.
- Cook, C.E., Wang, Y. & Sensabaugh, G. (1999) A mitochondrial control region and cytochrome *b* phylogeny of sika deer (*Cervus Nippon*) and report of tandem repeats in the control region. *Molecular Phylogenetics and Evolution*. **12**(1), 47-56.
- Crandall, K.A., Bininda-Emonds, O.R.P., Mace, G.M. & Wayne, R.K. (2000) Considering evolutionary processes in conservation. *Trends in Ecology and Evolution*. **15**, 290-295.
- Davison, A., Birks, J.D.S., Brookes, R.C., Messenger, J.E. & Griffiths, H.I. (2001) Mitochondrial phylogeography and population history of pine martens *Martes martes* compared with polecats *Mustela putorius*. *Molecular Ecology*. **10**, 2479-2488.
- deMenocal, P.B. (2004) African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters*. **220**, 3-24.
- Deuter, R., Pietsch, S., Hertel, S. & Müller, O. (1995) A method for the preparation of fecal DNA suitable for PCR. *Nucleic Acids Research*. **23**(18), 3800-3801.
- Dunbar, R.I.M & Dunbar, E.P. (1974) Social organisation and ecology of the klipspringer (*Oreotragus oreotragus*) in Ethiopia. *Z. Tierpsychol.* **19**, 155-161.

- Eggert, L.S., Rasner, C.A. & Woodruff, D.S. (2002) The evolution and phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. *Proceedings of the Royal Society London B*. **269**, 1993-2006.
- Emerson, B.C., Paradis, E. & Thebaud, C. (2001) Revealing the demographic histories of species using DNA sequences. *Trends in Ecology and Evolution*. **16**(12), 707-716.
- Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. **131**, 479-491.
- Excoffier, L., Laval, G. & Schneider, S. (2006) ARLEQUIN Version 3.01: An integrated software package for population genetics data analysis. Computational and Molecular Population Genetics Lab (CMPG), Institute of Zoology, University of Berne, Bern, Switzerland.
- Fernando, P., Pfrender, M.E., Encalada, S.E. & Lande, R. (2000) Mitochondrial DNA variation, phylogeography and population structure of the Asian elephant. *Heredity*. **84**, 362-372.
- Frankel, O.H. & Soule, M.E. (1981) Conservation and evolution. Cambridge University Press, UK.
- Galindo-Leal, C. & Weber, M. (1994) Translocation of deer subspecies: reproductive implications. *Wildlife Society Bulletin*. **22**(1), 117-120.
- Gemmell, N.J., Metcalf, V.J. & Allendorf, F.W. (2004) Mother's curse: the effect of mtDNA on individual fitness and population viability. *Trends in Ecology and Evolution*. **19**(5), 238-244.
- Gibbs, M.J., Armstrong, J.S. & Gibbs, A.J. (2000) Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics*. **16**, 573-582.

- Griffith, B., Scott, J.M., Carpenter, J.W. & Reed, C. (1989) Translocation as a species conservation tool: status and strategy. *Science*. **245**, 477-480.
- Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. **41**, 95-98.
- Hassanin, A. & Douzery, E.J.P. (1999) The tribal radiation of the family Bovidae (Artiodactyla) and the evolution of the mitochondrial cytochrome *b* gene. *Molecular Phylogenetics and Evolution*. **13**(2) 227-243.
- Holder, M. & Lewis, P.O. (2003) Phylogeny estimation: Traditional and Bayesian approaches. *Nature*. **4**, 275-284.
- Huelsenbeck, J.P. & Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics Applications Note*. **17**(8) 754-755.
- Huson, D.H & Bryant, D. (2006) Application of phylogenetic networks, in evolutionary studies. *Molecular Biology and Evolution*. **23**(2), 254-267.
- IEA (Institute of Applied Ecology). (1998) *Oreotragus oreotragus*. In *African Mammals Databank – A databank for the conservation of the African mammals Vol 1 and 2*. Bruxelles: European Commission Directorate. Available online at <http://gorilla.bio.uniroma1.it/amd/amd184b.html>
- Inger, R.F. (1961) Problems in the application of the subspecies concept in the vertebrate taxonomy. In: *Vertebrate speciation* (Ed. Blair, W.F.) University Texas Press, Austin, Texas.
- Keller, L.F. & Waller, D.M. (2002) Inbreeding effects in wild populations. *Trends in Ecology and Evolution*. **17**(5), 230-241.
- Kimura, M. (1980) A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*. **16**, 111-120.

- Kohn, M.H. & Wayne, R.K. (1997) Facts from feces revisited. *Trends in Ecology and Evolution*. **12**(6), 223-227.
- Kreader, C.A. (1996) Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Applied and Environmental Microbiology*. **62**(3) 1102-1106.
- Kumar, S., Tamura, K. & Nei, M. (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*. **5**, 150-163.
- Lecis, R. & Norris, K. (2004) Population genetic diversity of the endemic Sardinian newt *Euproctus platycephalus*: implications for conservation. *Biological Conservation*. **119**, 263-270.
- Lopez, J.V., 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.
- Manier, M.K. (2004) Geographic variation in the long-nosed snake, *Rhinocheilus lecontei* (Colubridae): beyond the subspecies debate. *Biological Journal of the Linnean Society*. **83**, 65-85.
- Martin, D.P. & Rybicki, E. (2000) RDP: detection of recombination amongst aligned sequences. *Bioinformatics*. **16**, 562-563.
- Martin, D.P., Williamson, C., Posada, D. (2005) RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics*. **21**, 260-262.
- Mathee, C.A. & Robinson, T.J. (1999) Cytochrome *b* phylogeny of the family Bovidae: resolution within the Alcelaphini, Antilopini, Neotragini, and Tragelaphini. *Molecular Phylogenetics and Evolution*. **12**(1), 31-46.
- Maynard Smith, J. (1992) Analysing the mosaic structure of genes. *Journal of Molecular Evolution*. **34**, 126-129.

- Mayr, E. (1982) Of what use are subspecies? *The Auk*. **99**, 593-595.
- Monroe, B.L. Jr. (1982) A modern concept of the subspecies. *The Auk*. **99**, 608-609.
- Morin, P.A., Moore, J.J, Chakraborty, R., Jin, L., Goodall, J. & Woodruff, D.S. (1994) Kin selection, social structure, gene flow and the evolution of chimpanzees. *Science*. **265**, 1193-1201.
- Moritz, C. (1994) Defining “evolutionary significant units” for conservation. *Trends in Ecology and Evolution*. **9**, 1539-1547.
- Mullenbach, R., Lagoda, P.J.L. & Welter, C. (1989) An efficient salt chloroform extraction of DNA from blood and tissues. *Trends in Genetics*. **5**, 391.
- Nersting, L.G. & Arctander, P. (2001) Phylogeography and conservation of impala and greater kudu. *Molecular Ecology*. **10**, 711-719.
- Nielsen, K.K. & Arctander, P. (2001) Recombination among multiple mitochondrial pseudogenes from a passerine genus. *Molecular Phylogenetics and Evolution*. **18**(3), 362-369.
- Norton, P.M. (1980) The habitat and feeding ecology of the klipspringer *Oreotragus oreotragus* (Zimmerman 1783) in two areas of the Cape Province. M.S. thesis, University of Pretoria, South Africa.
- Pääbo, S., Irwin, D.M. & Wilson, A.C. (1990) DNA damage promotes jumping between templates during enzymatic amplification. *Journal of Biological Chemistry*. **265**, 4718-4721.
- Padidam, M., Sawyer, S. & Fauquet, C.M. (1999) Possible emergence of new geminiviruses by frequent recombination. *Virology*. **265**, 218-225.
- Posada, D. & Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*. **14**, 817-818.

- Posada, D. & Crandall, K.A. (2001) Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proceedings of the National Academy of Sciences of the USA*. **98**, 13757-13762.
- Reed, J.Z., Tollit, D.J., Thompson, P.M. & Amos, W. (1997) Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal faeces. *Molecular Ecology*. **6**, 225-234.
- Rebholz, W. & Harley, E. (1999) Phylogenetic relationships in the bovid subfamily Antilopinae based on mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*. **12**(2), 87-94.
- Ricklefs, R.E. & Schluter, D. (1993) Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago.
- Roberts, S.C. (1998) Behavioural responses to scent marks of increasing age in klipspringer, *Oreotragus oreotragus*. *Ethology*. **104**, 585-592.
- Rogers, A.R. & Harpending, H.C. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*. **9**, 552-569.
- Rohland, N., Siedel, H. & Hofreiter, M. (2004) Non-destructive DNA extraction method for mitochondrial DNA analyses of museum specimens. *Biotechniques*. **36**(5), 814-821.
- Rozas, J., Sánchez-DelBarrio, J.C., Messeguer, X. & Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*. **19**, 2496-2497.
- Rout, T.M., Hauser, C.E. & Possingham, H.P. (2007) Minimise long-term loss or maximise short-term gain? Optimal translocation strategies for threatened species. *Ecological modelling*. **201**, 67-74.

- Salminen, M.O., Carr, J.K., Burke, D.S. & McCutchan, F.E. (1995) Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *Aids Research and Human Retroviruses*. **11**, 1423-1425.
- Schaub, M., Pradel, R., & Lebreton, J.D. (2004) Is the reintroduced white stork (*Ciconia ciconia*) population in Switzerland self-sustainable? *Biological Conservation*. **119**, 105-114.
- Simonsen, B.T., Siegismund, H.R. & 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.
- Skead, C.J. (1985) Historical mammal incidence in the Cape Province: The western and northern Cape. The Department of Nature and Environmental Conservation of the Provincial Administration of the Cape of Good Hope, Cape Town.
- Skinner, J.D. & Chimimba, C.T. (2005) The mammals of the southern Africa subregion. Cambridge University Press.
- Slatkin, M. (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*. **47**, 264-279.
- Sorenson, M.D. & Quinn, T.W. (1998) Numts: a challenge for avian systematics and population biology. *The Auk*. **115**(1), 214-221.
- Stemmer, W.P.C. (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proceedings of the National Academy of Sciences of the USA*. **91**, 10747-10751.
- Storer, R.W. (1982) Subspecies and the study of geographic variation. *The Auk*. **99**, 599-601.
- Storfer, A. (1999) Gene flow and endangered species translocation: a topic revisited. *Biological Conservation*. **87**, 173-180.

- Su, B., Wang, Y., Lan, H. & Zhang, Y. (1999) Phylogenetic study of complete cytochrome *b* genes in musk deer (genus *Moschus*) using museum samples. *Molecular phylogenetics and Evolution*. **12**(3), 241-249.
- Swofford, D. (Eds.) (2002) PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, MA.
- Tang, Q., Liu, H., Mayden, R. & Xiong, B. (2006) Comparison of evolutionary rates in the mitochondrial DNA cytochrome *b* gene and control region and their implications for phylogeny of the Cobitoidea (Teleostei: Cypriniformes). *Molecular Phylogenetics and Evolution*. **39**, 347-357.
- Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*. **22**, 5673-4680.
- Tilson, R.L. & Norton, P.M. (1981) Alarm duetting and pursuit deterrence in an African antelope. *The American Naturalist*. **118**(3), 455-462.
- Venkatesh, B., Dandona, N. & Brenner, S. (2006) Fugu genome does not contain mitochondrial pseudogenes. *Genomics*. **87**, 307-310.
- Weber, D.S., Stewart, B.S., Garza, J.C. & Lehman, N. (2000) An empirical genetic assessment of the severity of the northern elephant seal population bottleneck. *Current Biology*. **10**, 1287-1290.
- Westerman, M., Loke, S. & Springer, M.S. (2004) Molecular phylogenetic relationships of two extinct potoroid marsupials, *Potorous platyops* and *Caloprymnus campestris* (Potoroinae: Marsupialia). *Molecular Phylogenetics and Evolution*. **31**, 476-485.
- Whittier, C.A., Dhar, A.K., Stem, C., Goodall, J. & Alcivar-Warren, A. (1999) Comparison of DNA extraction methods for PCR amplification of mitochondrial cytochrome *c* oxidase

- subunit II (COII) DNA from primate fecal samples. *Biotechnology Techniques*. **13**, 771-779.
- Wilkinson, G.S. & Chapman, A.M. (1991) Length and sequence variation in evening bat d-loop mtDNA. *Genetics*. **128**, 607-617.
- Wilson, E.O. & Brown, W.L. (1953) The subspecies concept and its taxonomic application. *Systematic Zoology*, **2**(3), 97-111.
- Wilson, V.J. & Child, G. (1965) Notes on klipspringer from tsetse fly control areas in eastern Zambia. *Arnoldia* (Rhodesia). **1**(35), 1-9.
- Zhang, F. & Jiang, Z. (2006) Mitochondrial phylogeography and genetic diversity of Tibetan gazelle (*Procapra picticaudata*): implications for conservation. *Molecular Phylogenetics and Evolution*. **41**(2), 313-321.
- Zhang, D-X. & Hewitt, G.M. (1996) Nuclear integrations: challenges for mitochondrial DNA markers. *Trends in Ecology and Evolution*. **11**(6), 247-251.
- Zusi, R.L. (1982) Intraspecific geographic variation and the subspecies concept. *The Auk*. **99**, 606-608.

APPENDICES

Appendix 1: Sequence alignment of the identified Numt sequences

	10	20	30	40	50	60	70	80	90	100	110	120	
1	T	A	A	G	A	C	T	C	A	A	G	A	A
2	G	T	A	T	A	A	A	G	T	G	A	A	A
3	C	C	C	C	A	C	A	T	A	T	C	A	A
4	C	A	A	C	C	A	A	A	G	T	G	A	A
5	G	A	A	G	T	G	A	A	A	G	T	G	A
6	A	A	G	T	G	A	A	A	G	T	G	A	A
7	A	A	G	T	G	A	A	A	G	T	G	A	A
8	A	A	G	T	G	A	A	A	G	T	G	A	A
9	A	A	G	T	G	A	A	A	G	T	G	A	A
10	A	A	G	T	G	A	A	A	G	T	G	A	A
11	A	A	G	T	G	A	A	A	G	T	G	A	A
12	A	A	G	T	G	A	A	A	G	T	G	A	A
13	A	A	G	T	G	A	A	A	G	T	G	A	A
14	A	A	G	T	G	A	A	A	G	T	G	A	A
	130	140	150	160	170	180	190	200	210	220	230	240	
1	C	A	T	T	A	C	C	A	A	A	G	T	G
2	G	T	G	A	A	A	G	T	G	A	A	A	A
3	C	C	C	A	C	A	T	A	T	C	A	A	A
4	C	A	A	C	C	A	A	A	G	T	G	A	A
5	G	A	A	G	T	G	A	A	A	G	T	G	A
6	A	A	G	T	G	A	A	A	G	T	G	A	A
7	A	A	G	T	G	A	A	A	G	T	G	A	A
8	A	A	G	T	G	A	A	A	G	T	G	A	A
9	A	A	G	T	G	A	A	A	G	T	G	A	A
10	A	A	G	T	G	A	A	A	G	T	G	A	A
11	A	A	G	T	G	A	A	A	G	T	G	A	A
12	A	A	G	T	G	A	A	A	G	T	G	A	A
13	A	A	G	T	G	A	A	A	G	T	G	A	A
14	A	A	G	T	G	A	A	A	G	T	G	A	A
	250	260	270	280	290	300	310	320	330	340	350	360	
1	G	A	G	T	C	T	A	T	G	T	A	A	C
2	A	A	C	A	T	A	A	T	G	C	A	A	C
3	A	A	C	A	T	A	A	T	G	C	A	A	C
4	A	A	C	A	T	A	A	T	G	C	A	A	C
5	A	A	C	A	T	A	A	T	G	C	A	A	C
6	A	A	C	A	T	A	A	T	G	C	A	A	C
7	A	A	C	A	T	A	A	T	G	C	A	A	C
8	A	A	C	A	T	A	A	T	G	C	A	A	C
9	A	A	C	A	T	A	A	T	G	C	A	A	C
10	A	A	C	A	T	A	A	T	G	C	A	A	C
11	A	A	C	A	T	A	A	T	G	C	A	A	C
12	A	A	C	A	T	A	A	T	G	C	A	A	C
13	A	A	C	A	T	A	A	T	G	C	A	A	C
14	A	A	C	A	T	A	A	T	G	C	A	A	C
	370	380	390	400	410	420	430	440	450	460	470	480	
1	A	T	A	G	T	A	C	A	T	T	A	A	T
2	C	C	A	T	G	C	A	T	G	T	A	A	C
3	C	C	A	T	G	C	A	T	G	T	A	A	C
4	C	C	A	T	G	C	A	T	G	T	A	A	C
5	C	C	A	T	G	C	A	T	G	T	A	A	C
6	C	C	A	T	G	C	A	T	G	T	A	A	C
7	C	C	A	T	G	C	A	T	G	T	A	A	C
8	C	C	A	T	G	C	A	T	G	T	A	A	C
9	C	C	A	T	G	C	A	T	G	T	A	A	C
10	C	C	A	T	G	C	A	T	G	T	A	A	C
11	C	C	A	T	G	C	A	T	G	T	A	A	C
12	C	C	A	T	G	C	A	T	G	T	A	A	C
13	C	C	A	T	G	C	A	T	G	T	A	A	C
14	C	C	A	T	G	C	A	T	G	T	A	A	C
	490	500	510	520	530	540	550	560	570	580	590		
1	T	G	T	C	C	A	T	T	A	C	G	A	C
2	T	G	T	C	C	A	T	T	A	C	G	A	C
3	T	G	T	C	C	A	T	T	A	C	G	A	C
4	T	G	T	C	C	A	T	T	A	C	G	A	C
5	T	G	T	C	C	A	T	T	A	C	G	A	C
6	T	G	T	C	C	A	T	T	A	C	G	A	C
7	T	G	T	C	C	A	T	T	A	C	G	A	C
8	T	G	T	C	C	A	T	T	A	C	G	A	C
9	T	G	T	C	C	A	T	T	A	C	G	A	C
10	T	G	T	C	C	A	T	T	A	C	G	A	C
11	T	G	T	C	C	A	T	T	A	C	G	A	C
12	T	G	T	C	C	A	T	T	A	C	G	A	C
13	T	G	T	C	C	A	T	T	A	C	G	A	C
14	T	G	T	C	C	A	T	T	A	C	G	A	C

Appendix 2: Sequence alignment of 251 base pairs of *Oreotragus oreotragus* cytochrome *b*

[illegible]

14969T.....C.....A.....T.....
11720T.....C.....A.....T.....
14971C..C.....G.....
14972C.....
PB1T.....C.....A.....T.....
3062T.....C.....a.....t.....
3063T.....C.....a.....t.....
KNP1T.....C.....A.....T.....
3134T.....C.....A.....
13376T.....C.....A.....C
14976T.....C.....A.....
14977T.....C.....A.....T.....
2524T.....C.....A.....
3108T.....C.....A.....
3164T.....C.....A.....
3174T.....C.....A.....
3223T.....C.....A.....
3259T.....C.....A.....
1688T.....C.....A.....
3147T.....C.....A.....
1781T.....C.....A.....T.....
1666T.....C.....A.....T.....
2509T.....C.....A.....T.....
3344T.....C.....A.....t.....
3083T.....C.....A.....
654610T.....C.....A.....T.....
663861T.....C.....A.....
113121	...C.....T.....C.....A.....
127107T.....C.....A.....A.....
4451437T.....C.....A.....

	140	150	160	170	180	190	200	210	220	230	240	250
1_	GGAGACCCAGATAACTACACCCAGCAAACCCACTCAACACTCCCCCTCACATTAACCAGAAATGATATTTCTATTTCGCATACGCAATCCTACGATCAATCCCCAATAAACTaggAGGCG											
2_												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												
26												
27												
28												

29	
30	
31	
32	
33	
34	
35	
36	
37	
39t.....
40t.....
41	
42	
43t.....
44	
45	
46T.....
15729t.....
15730	
38382	
17195t.....
40431	
40467	
Bia1	
Bia2	
14969T.....
11720T.....
14971	
14972	
PB1C.....T.....
3062t.....
3063t.....
KNP1	
3134T.....
13376T.....
14976T.....
14977T.....
2524T.....
3108T.....
3164t.....
3174T.....
3223T.....
3259T.....
1688T.....
3147T.....
1781t.....
1666T.....
2509T.....
3344t.....
3083t.....
654610T.....
663861T.....
113121G.....T.....
127107T.....
4451437T.....

Appendix 3: Sequence alignment of 423 base pairs of *Oreotragus oreotragus* control region

	10	20	30	40	50	60	70	80	90	100	110	120	130	140
1	ACTCCCTGAAACACTATCAATATAACCACATAAAATACCAAGAGCCTTCCCAGTATTAAATTTCCCAAAAAATTCAGAACT-AACACAGACTTAGCACTCCACAGCCTTACATG-----CACATCC-CATCCCAGG													
2													
3													
4													
5													
6													
7													
8							-G					TA	
9													
10													
11										C			
12													
13													
14													
15													
16				T	T	-C				C GC		C	
17				T		-C				C GC		C	
20							A					T C	
21													
23							A					T C	
24				T	T	-C				C GC		C	
27				T		-C				C GC		C	
28							A					T C	
30				T	T	-C				C GC		C	
31							-G					TA	
32	ACTCCCTGAAACACTATCAATATAACCACATAAAATACCAAGAGCCTTCCCAGTATTAAATTTCCCAAAAAATTCAGAACT-AACACAGACTTAGCACTCCACAGCCTTACATG-----CACATCC-CATCCCAGG													
33				T	T	-C				C GC		C	
34				T	T	-C				C GC		C	
35				T	T	-C				C GC		C	
36				T	T	-C				C GC		C	
39													
42				T	T	-C				C GC		CT	
38382													
40431													
40467													
Bia1				T	T	-C				C GC		C	
Bia2													
14969	T.....T.....						-C	A	T		A		A C T	
14971				T	T	-C				C GC		C	
14972													
PB1	T.....				C		T	-C	T	G		TGCAACAACG	A TT TA	
3062				T	T	-C				C GC		C	
3063													
KNP1	T.....				C		T	-C	T	G		TGCAACAACG	A TT TA	
14976													
2524	T.....				T		-C		TG	G		C	T AT C T T	
3108	T.....				T		-C		TG	G		C	T AT C T T	
3164	T.....								T	G		C	T TT C T T	
3174	T.....								T	G		C	T TT C T T	
3223	T.....								TG	G		C	T AT C T T	
3259	T.....					A			T	G		C	T TT C T T	
1688	T.....				T				TG	G		C	T AT C T T	
3147	T.....								T	G		C	T TT C T T	
1781	T.....								T	G		C	T TT C T T	

	150	160	170	180	190	200	210	220	230	240	250	260	270	280
1	CAATAAAACATGTCTAACTTACAATACATAGTATGTTCAAAGTACACACATGAATGAGCACATACAGTTAATGTACTATAAACATAACATGTAACAACACACTACACTACCTAGCACATGGGTGTGGGTATGTACATATA													
2														
3				C		A					T			
4														
5														
6				C		A					T			
7				C		A					T			
8				A							G		A	
9														C
10														
11											T			
12														
13				C		A					T		G	
14				C		A					T			
15				C		A					T			
16					TG		T G		G C		C		T	
17					TG		T G		G C		C		T	
20														
21														
23														
24					T	C	T G		G		T C		T	
27					TG		T G		G C		C		T	
28														
30					TG		T G		G C		C		T	
31				A										
32					T	C	T G		G		T C		T	
33					TG		T G		G C		C		T	
34					TG		T G		G C		C		T	
35					T	C	T G		G		C		T	
36					T	C	T G		G		C		T	
39														
42					T	C	T G		G		T C		T	
38382														
40431														
40467														
Bla1					TG		T G		G C		C		T	
Bla2					T		T G		G		C		T	
14969				G		C		C						
14971					TG		T G		G C		C		T	
14972														
PB1				C	G		A		G	TG		A		G
3062					TG		T G		G C		C		T	
3063														
KNP1				C	G		A		G	TG		A		G
14976														
2524				C	G		A		G	T		A		G
3108				C	G		A		G	T		A		G
3164				C	G		A		G	T		A		G
3174				C	G		A		G	T		A		G
3223				C	G		A		G	T		A		G
3259				C	G		A		G	T		A		G

```

1688 . . C . G . . . A . . . . G . T . . . . A . . . . . TT . . . C . . C . . . G . . . . . C . . . . . G . T . . . T . . . G . . T . . . A . . . A . . GC . . . .
3147 . . C . G . . . A . . . . G . T . . . . A . . . . . TT . . . C . . C . . . G . . . . . C . . . . . G . T . . . TC . . . G . . T . . . AA . . A . . GC . . . .
1781 . . C . G . . . A . . . . G . TG . . . . A . . . . . G . . . . T . . . C . . C . . . - G . . C . . . C . . . . . G . T . . . T . . . A . . T . . . AA . . A . . CG . . . GC . .
1666 . . C . G . . . A . . . . G . TG . . . . A . . . . . G . . . . T . . . C . . C . . . - G . . C . . . C . . . . . G . T . . . T . . . A . . T . . . AA . . A . . CG . . . GC . .
2509 . . C . G . . . A . . . . G . TG . . . . AC . . . G . . . . T . . . C . . C . . . - G . . C . . . C . . . . . G . T . . . T . . . G . . T . . . AA . . A . . CG . . . GC . .
3344 . . C . G . . . A . . . . G . TG . . . . A . . . . . G . . . . T . . . C . . C . . . - G . . C . . . C . . . . . G . T . . . T . . . A . . T . . . AA . . A . . CG . . . GC . .
3083 . . C . G . . . A . . . . G . T . . . . A . . . . . TT . . . C . . T . . . G . . . . . C . . . . . G . T . . . T . . . G . . T . . . AA . . A . . GC . . . .
663861 . . C . G . . . A . . . . G . T . . . . A . . . . . TT . . . C . . C . . . G . . . . . C . . . . . G . T . . . T . . . G . . T . . . AA . . A . . GC . . . .
127107 . . C . G . . . A . . . . G . T . . . . A . . . . . TT . . . C . . C . . . G . . . . . C . . . . . G . . . . . T . . . G . . T . . . A . . A . . C . . . . .
4451437 . . C . G . . . A . . . . G . T . . . . A . . . . . TT . . . C . . C . . . G . . . . . C . . . . . G . . . . . T . . . G . . T . . . A . . A . . C . . . . .

```

```

290 300 310 320 330 340 350 360 370 380 390 400 410 420
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1 ATTAATGTATTAAAGACATAATATGTATATAGTACATTCTATTATTTAGTACATGAATAATAAGCACAGTACATACAATTAAATGTATTAAACACATAAATGTATATAGTACATTAAATTATCTAGTCCATGCATATAAG
2
3 . . . . . C . . C . . . . . T . . . . . C . . T . . . . .
4 . . . . . T . . . . . C . . . . .
5
6 . . . . . C . . C . . . . . T . . . . . C . . T . . . . .
7 . . . . . C . . C . . . . . T . . . . . C . . T . . . . .
8 . . . . . TC . . C . . . . . T . . . . . T . . . . . C . . T . . . . . C . . . . . G . . . . .
9
10 . . . . . C . . . . . G . . . . . T . G . . . . C . . T . . . . .
11 . . . . . T . G . . . . CC . . . . . C . . . . .
12
13 . . . . . C . C . C . . . . . T . . . . . C . . T . . . . . C . . . . .
14 . . . . . C . C . C . . . . . T . . . . . C . . T . . . . .
15 . . . . . C . C . C . . . . . T . . . . . C . . T . . . . . C . . . . .
16 G . . . . . G . . . . . C . . T . . . T . . . . . G . . . . .
17 G . . . . . G . . . . . C . . T . . . T . . . . . G . . . . .
20 . . . . . C . . . . . T . . . . .
21
23 . . . . . C . . . . . T . . . . .
24 G . . . . . G . . . . . T . . . . .
27 G . . . . . G . . . . . C . . T . . . T . . . . . G . . . . .
28 . . . . . C . . . . . T . . . . .
30 G . . . . . G . . . . . C . . T . . . T . . . . . G . . . . .
31 . . . . . TC . . C . . . . . T . . . . . C . . T . . . . . C . . . . . G . . . . .
32 . . . . . G . . . . . C . . . . . T . . . . .
33 G . . . . . G . . . . . C . . T . . . T . . . . . G . . . . .
34 G . . . . . G . . . . . C . . T . . . T . . . . . G . . . . .
35 . . . . . TC . . C . . . . . T . . . . . C . . . . . T . . . . .
36 . . . . . G . . . . . TC . . C . . . . . T . . . . . C . . . . . T . . . . .
39 . . . . . C . . . . . T . G . . . C . . . . .
42 . . . . . G . . . . . C . . . . . T . . . . .
38382
40431
40467
Bia1 G . . . . . G . . . . . C . . . . . T . . . . . C . . . . . G . . . . .
Bia2 . . . . . G . . . . . C . . C . . . G . . . . . T . . . . . GT . . . . . T . . . . . G . . . . .
14969 . . C . G . . . G . . . . . C . . . . . C . . T . . . GT . . . . .
14971 . . C . . . C . . . . . CTC . . . . . T . . . . . T . . . . . GT . . . . . C . . . . .
14972 . . . . . G . . . . . T . G . . . C . . . . . T . . . . .
PBI1 G . . . . . G . . . . . TC . . G . CC . . . . . TG . . . . . C . . G . . .
3062 G . . . . . TG . . . . . C . . . . . T . . . . . T . G . . . C . . . . .
3063 . . . . . T . G . . . C . . . . .
KNP1 G . . . . . G . . . . . TC . . G . CC . . . . . TG . . . . . C . . GT . . . . . G . . . . .
14976 . . . . . T . G . . . C . . . . .
2524 G . . . . . C . . G . . . C . . C . . . T . . . G . G . . C . . . . .
3108 G . . . . . C . . G . . . C . . C . . . T . . . G . G . . C . . . . .
3164 G . C . . . G . . G . G . . . . . C . . C . . . C . . GT . . . T . . . . . GT . . . . .

```

3174	G.C....G...G.....C.....C...GT.....T.....GT.....
3223	G.....C..G.....C..C.....G..T.....GT.....C..T.....
3259	G.C....G..G.G.....C.....C.....C...GT.....T.....GT.....
1688	G.....C...G.....C..C.....T.....G..G.....C.....
3147	G.C....G...G.....C.....C...GT.....T.....GT.....
1781	G..G...G..G.G.....TC...G..C.....T.....GT.....G.....G.....
1666	G..G...G..G.G.....TC...G..C.....T.....GT.....G.....G.....
2509	G.....G...G.....TC...G..C.....G.....T.....T.....C...GT.....G.....
3344	G..G...G..G.G.....TC...G..C.....T.....GT.....G.....G.....
3083	G.....C..G.....C..C..C.G.....T.....GT.G.....C.....
663861	G.....G.G.....C.....C..C.....T.....GT.....GT.....G.....
127107	G.....G.....G.....TC...C.....G.....G..T.....GT.....C..G.....
4451437T.G.....C.....