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.

University of Cape Town

Phylogeny and Phylogeography of the Chacma Baboon *(Papio ursinus):*

The role of landscape in shaping contemporary genetic structure in the southern African baboon

Riashna Sithaldeen

University of Cape Town

Supervisors:
Assoc. Prof. Rebecca Ackermann, Department of Archaeology
Dr. Jacqueline Bishop, Department of Zoology

Thesis Presented for the Degree of
DOCTOR OF PHILOSOPHY
in the Department of Archaeology

Faculty of Science
UNIVERSITY OF CAPE TOWN
March 2011

The Big Baboon

The Big Baboon is found upon
The plains of Cariboo:
He goes about with nothing on
(A shocking thing to do).
But if he dressed up respectably
And let his whiskers grow,
How like this Big Baboon would be
To Mister So-and-so!

Hilaire Belloc

University of Cape Town

DECLARATION

This thesis reports the results of the original research I conducted under the auspices of the Department of Archaeology in the Faculty of Science at the University of Cape Town, between 2004 and 2011. All the assistance that I received has been acknowledged. This work has not been submitted for a degree at any other university.

Signed by candidate

Signature Removed

_____

Riashna Sithaldeen

University of Cape Town

Author: Riashna Sithaldeen

Thesis title: Phylogeny and Phylogeography of the Chacma Baboon *(Papio ursinus):*
The role of landscape in shaping contemporary genetic structure in the southern African baboon

Date: 04 March 2011

ABSTRACT

Baboons (genus Papio*) have evolved over the course of the Plio-Pleistocene, and remain widely distributed throughout Africa. This thesis contributes to our understanding of the role of climate and landscape change in structuring diversity within chacma baboons (*Papio ursinus*). The data set comprises molecular sequences from two mitochondrial DNA markers: the Brown region and the hypervariable D-loop. Faecal samples were collected from 261 free living chacma baboons across southern Africa from which DNA was extracted and processed. Phylogenetic and phylogeographic techniques, including coalescent modeling, are used to examine past and present population dynamics of chacma baboon populations. Bayesian tree constructions provide a timeline of diversification for the sample. Results recover a major diversification event at ~1.6 Ma which divides chacma baboons into two major mitochondrial lineages, a northern lineage (NL) and a southern lineage (SL). This is coincident with the aridification of southern Africa and a period of genetic isolation between populations, due to the expansion of the Kalahari Desert. Two monophyletic clades are nested within these lineages. The first is a clade of Namibian baboons which diversified at ~1.0 Ma, most likely in response to fluctuations in levels of the Orange River. The second is a northeastern clade (NeC) which diversified at ~420 ka in response to a warmer and wetter climates. Although the phylogenetic placement and geographic distribution of NeC relative to NL suggests that it was founded by a minority fragment of NL, the limited sample cannot confirm this. A sudden expansion event is recovered for SL, most likely out of the southwestern Cape, which was used as a glacial refuge for the species. This expansion is dated to ~15 ka which coincides with the end of the last glacial maximum. Estimates of the contribution of the geographic parameters shaping modern genetic structure were inconclusive. The results of this study support a model of climate driven diversification for* Papio *in which periods of isolation drive genetic and phenotypic differentiation. Although the ecological drivers of ongoing differentiation remain unclear, it has been shown that population contractions and expansions have also played a significant role in driving regional genetic structure within the species*

University of Cape Town

ACKNOWLEDGEMENTS

I will be eternally grateful for the support and encouragement of my supervisors Assoc Prof. Becky Ackermann and Dr. Jacqui Bishop. Your wisdom and patience in guiding me though the preparation of this thesis has been invaluable. Thank you.

I would also like to thank:

The Palaeoanthropological Scientific Trust, the South African National Research Foundation, the Wenner-Gren Foundation and the University of Cape Town for financial support.

Professor Raj Ramesar and the staff and students of the Department of Human Genetics (UCT Medical School) who provided lab space, resources and advice.

Professor Judy Sealy and the Department of Archaeology at UCT for creating a supportive and stimulating environment to work in.

South African National Parks for providing permits to collect samples and for rangers and conservation officers always ready to lend a hand.

Ministry of Environment and Tourism, Namibia, for collection permits.

Professor Maarten De Wit and Dr. Woody Cotterill of the African Earth Observatory Network (AEON) unit at UCT for the use of their computational facilities to perform some of the analyses presented here.

For samples I am grateful to:

Nicolas Fourie, for going above and beyond the call of friendship.

Kurt Ackermann, Shabnum Behari, Jacqui Bishop, Woody Cotterill, Angela Layman, Lieutenant Moton (South African Police Services) and Anthony Roberts (Lapalala Wilderness Game Reserve) for much appreciated help in the field.

Becky Ackermann, the Archer family, Esme Beamish (Baboon Research Unit, UCT), Seth Eiseb (National Museum of Namibia), Helena Fitchat (Centre for Rehabilitation

University of Cape Town

of Wildlife), Jacinta Beehner, Thore Bergman, Dorothy Cheney, and Robert Seyfarth of the Dorothy Cheney and Robert Seyfarth 'Lab', Matthew Bond, Geraldine Frasier (Namibia National parks), Gareth Hempson, Eric Hermann, Hannes Marais (Mpumalanga Parks Board), Erin Roberts (Lapalala Wilderness School).

Special thanks to:

Andy Burrell (New York Consortium in Evolutionary Primatology) and Dietmar Zinner (Deutsche Primatenzentrum) who kindly provided me with *P. ursinus* sequence data.

Ms. Lynn Cable (Department of Archaeology) you have been a pillar of strength to me over these many years. Thank you for assistance with logistics, and for always being willing to listen to the crazy ramblings of put upon graduate students.

To my fellow graduate students, for a sympathetic ear, and to all my friends for forgiving my neglect of you, thank you.

To my family (Sithaldeen and Archer), I thank you for your unwavering belief that I could finish this.

And finally to my partner Will Archer, your assistance in the field has made a significant contribution to the completion of this project and your personal pursuit of excellence has been a continuous inspiration. For all your support, you have my eternal gratitude.

University of Cape Town

Ethical clearance was obtained from the University of Cape Town Science Faculty Animal Experimentation Committee (AEC #2006/v2/RS) and field samples were collected under Western Cape Nature Conservation Board Permit #001-201-00004. Collection permits were also obtained from the Ministry of Environment and Tourism, Namibia (993/2005) and the Tsitsikamma National Park.

TABLE OF CONTENTS

University of Cape Town

University of Cape Town

CHAPTER 6

TESTING MODELS OF DIVERSIFICATION USING STATISTICAL PHYLOGEOGRAPHIC METHODS

University of Cape Town

LIST OF FIGURES

CHAPTER 2

University of Cape Town

CHAPTER 3

CHAPTER 4

CHAPTER 5

University of Cape Town

CHAPTER 6

University of Cape Town

University of Cape Town

LIST OF TABLES

University of Cape Town

LIST OF APPENDICES

University of Cape Town

CHAPTER 1

GENERAL INTRODUCTION

Baboons are large, primarily terrestrial, old world monkeys. They are highly adaptable and are generally considered to be one of the most successful and versatile of non-human primates. These primates have been studied extensively for over 70 years (Altmann and Altmann 1970) and have contributed significantly to our understanding of a wide array of primate behavioural, physiological, developmental and evolutionary traits (Barrett et al. 2000; Bergman et al. 2003; Cheney and Seyfarth 1995, 1997, 1999; Cheney et al. 1996, 2004 ; Engh et al. 2006; Fischer et al. 2000, 2002, 2004; Henzi and Barrett 2002, 2003; Kitchen et al. 2004,2005a, 2005b ; O'Connell and Cowlishaw 1994; Palombit et al. 1997, 2001; Weingrill et al. 2004). Studies of the evolutionary history of baboons can contribute to our understanding of ecologically linked diversification in large bodied primates (Jablonski 2002; Jolly 2001), and provide insight into mammalian diversification more broadly.

The diversification of *Papio* into five distinct forms -- hamadryas, chacma, olive, yellow and guinea -- has been linked to range shifts as climates fluctuated in the Plio-Pleistocene (Jolly 2001; Newman et al. 2004; Zinner et al. 2009). Based on Hewitt's model (1996, 2001) for diversification patterns in European mammalian lineages, Jolly (2001) proposed that as climates fluctuated and environments cycled in Africa, these processes shaped genetic structure in baboons. In this model, habitat shifts prompted baboon populations to respond by shifting their ranges, either expanding or contracting, in accordance with local resource availability. A decrease in the size of populations would represent genetic bottlenecks and surviving groups would then act as founding populations in times of range expansion. Geographic discontinuities across the genus as a whole would have resulted in reduced gene-flow between populations. This combination of factors greatly increased the likelihood for the fixation and proliferation of novel morphological and behavioural autapomorphies in the core populations. Repeated cycles would have led to the strengthening of morphological differences between populations, eventually resulting in the emergence of five distinct but reproductively compatible *Papio* morphotypes (Jolly 2001).

University of Cape Town

Jolly's model is elegant and very neatly explains the genetic and phenotypic structure that is observed in baboons today (Newman et al. 2004; Wildman et al. 2004, Zinner et al. 2009). However in the years since this model was proposed little has been done to statistically test the hypotheses proposed within it, or to develop it further. Instead studies aimed at phylogenetic reconstructions of the genus tend to fit the model to their data (e.g. Newman et al. 2004; Zinner et al. 2009). In this dissertation, phylogenetic and phylogeographic methods are employed to test the mechanisms proposed within Jolly's (2001) model, focusing on the role played by climate and landscape change in driving genetic structure and diversity in chacma baboons (*Papio ursinus*).

The two chapters following this general introduction (**Chapter 1**) are intended to provide a background to the thesis. **Chapter 2** summarises the evolutionary history of Papionins and *Papio* in southern Africa, and introduces each of the five major species that comprise *Papio*. **Chapter 3** is a description of the chacma baboon with particular focus on the ecological flexibility displayed by the species. Chacma baboons have been evolving independently for almost as long as modern *Papio* and therefore offer a comparable temporal record of genetic subdivisions and population responses to climate change. Chacma baboons are also southern African baboons and ecological generalists, which is the most likely ancestral state of all *Papio.* Chacma baboons are therefore offered here as a suitable proxy for modelling the response of the earliest baboon populations to climate change and habitat fragmentation. Some relevant information regarding the biogeography of the sub-continent, past and present, is also provided.

Chapters 4 through 6 are written as stand-alone chapters which present the methods and results of a series of analyses. To test the hypothesis of climate driven diversification for baboons phylogenetic methods are used. First a timeline of diversification for the chacma baboon is generated in **Chapter 4**. Bayesian and Maximum Parsimony methods are used to analyse a sample of mitochondrial Brown region sequence data, and evolutionary relationships together with node ages are estimated. These ages are then assessed for temporal correlation between genetic structuring and climate and landscape change events within southern Africa. Results support Jolly's original model which proposes that climate driven landscape change has played a significant role in driving structure within baboons.

As evident in many African lineages, genetic structure within a species is also driven by biotic factors such as population responses to habitat shifts. **Chapter 5** therefore investigates how chacma baboon populations respond to climate and landscape change. This is done using phylogeographic techniques to analyse a set of mitochondrial D-loop

University of Cape Town

data, which allows for the reconstruction of population histories. Results hint at a complex history of population contractions and expansions that may be significantly shaped by the climate and topography of the southern African landscape.

The results of Chapters 4 and 5 identify potentially important mechanisms driving population structure within baboons. The statistical probabilities related to these proposed population histories are assessed in **chapter 6**, using statistical phylogeographic methods and a coalescent framework. The results of the thesis are then discussed in light of our understanding of the role of climate and landscape change in shaping genetic structure within baboons.

The thesis concludes with a brief summary (**Chapter 7**).

University of Cape Town

CHAPTER 2

AN INTRODUCTION TO BABOONS (GENUS *PAPIO*)

Evolutionary relationships among the African Papionins

Climate fluctuations and habitat change have played a significant role in the evolutionary chain of events leading from the ancestor of all Papionini[1] to modern baboons. The extant Old World monkey tribe Papionini is a monophyletic group comprising the Eurasian macaques (genus *Macaca*) and the African Papionins which is made up of mangabeys (*Cercocebus* and *Lophocebus*), mandrills and drills (*Mandrillus*), and baboons (*Papio* and *Theropithecus*) (Delson 1975; Strasser and Delson 1987; Szalay and Delson 1979). This geographically defined split within Papionins is attributed to the formation of the Sahara Desert (Jablonski 2002) ~7.0 Ma (Schuster et al. 2006; Vignaud et al. 2002). The formation of this desert began with the gradual aridification of North Africa in the Miocene. First habitats shifted as tropical trees (Jacobs 2004; Wolfe 1985) were replaced with open grassland (Cerling 1992; Jacobs 2004), and finally at the Miocene/Pliocene boundary, the Sahara Desert formed, representing a physical barrier to gene flow in many mammal species (Douady et al. 2003). Climate driven habitat change and allopatric fragmentation of the Papionini thereby gave rise to the African papionins (Papionina).

Phylogenetic reconstructions of the lineages within Papionina have yielded three possible branch arrangements of the subtribe. The first two are based on morphological studies and the tree configurations are shown in Fig 2.1a and Fig 2.1b. These are presented with Eurasian macaques as an outgroup to Papionina. Although both trees show *Mandrillus* and *Papio* to be monophyletic and sister to *Cercocebus* and *Lophocebus*, they differ in the placement of *Theropithecus*. The first phylogeny (Fig.2.1a) places *Theropithecus* as the most divergent lineage in Papionina (Jolly 1972; Strasser and Delson 1987; Szalay and Delson 1979), while the second (Fig 2.1b) places *Theropithecus* as sister to *Mandrillus* and *Papio* (Delson and Dean 1993). More recently, a third phylogeny (Fig.2.1c) has been produced, based on the integration of morphological (Fleagle and McGraw 1999, 2002; Gilbert 2007, 2008; Groves 1978) and molecular (Cronin and Sarich 1976 ; Disotell 1994 ; Disotell et al. 1992; Harris and Disotell 1998; Hewett-Emmett and Cook 1978; Page and Goodman 2001; Tosi et al. 1999, 2003) evidence. This tree suggests that *Cercocebus* is

[1] Here I use the term Papionini interchangeably with Papionin as in Jolly (2001).

University of Cape Town

more closely related to *Mandrillus*, while *Lophocebus*, *Papio*, and *Theropithecus* form a second group (Singleton 2002). Due to a high degree of conservation in papionin craniometric traits across the tribe (Singleton 2002), it is difficult to determine relative ancestry within it using morphological data alone. It is therefore likely that the integration of morphological and molecular data provides the most accurate representation of true evolutionary relationships within the subtribe.

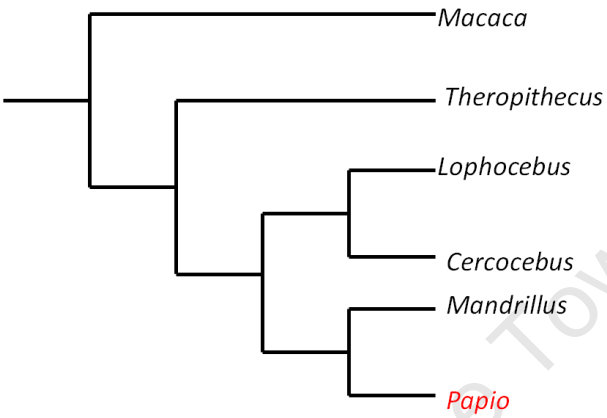

Figure 2.1a- A phylogenetic tree of the extant papionins, based on morphological data, shows* Theropithecus *as the most divergent African lineage. (Jolly 1972; Strasser and Delson 1987; Szalay and Delson 1979).

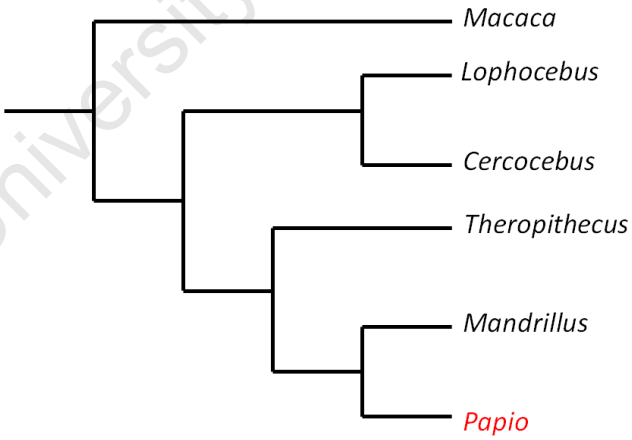

Figure 2.1b- A second phylogenetic tree of the extant papionins from morphological data is shown. Here* Theropithecus *is sister to a monophyletic clade of* Mandrillus *and* Papio *(Delson and Dean 1993).

University of Cape Town

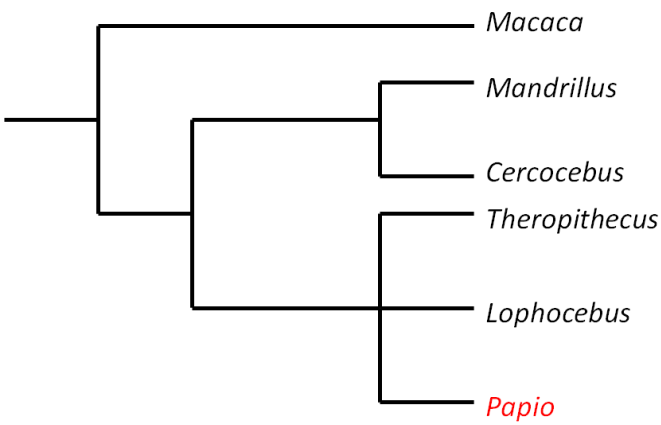

Figure 2.1c- A phylogenetic tree of the extant papionins generated from an integration of molecular and morphological data, shows a monophyletic clade of* Madrillus *and* Cercocebus *which is sister to an unresolved clade of* Theropithecus*,* Lophocebus *and* Papio *(Singleton, 2002).

University of Cape Town

The southern African Papionins

There are two early fossil papionin forms in Africa, *Parapapio* and *Pliopapio*. *Parapapio* is the oldest form, known from 7.4 - 2.0 Ma (Delson 1984,1988; Leakey et al. 2003). Although it is represented in East Africa (Brain 1981; Frost and Delson 2002), *Parapapio* is significantly more abundant in southern Africa (Brain 1981; Freedman 1957, 1976). Papionins in the South African Plio-Pleistocene are represented by up to five extinct and extant genera: *Papio*, *Parapapio*, *Dinopithecus*, *Gorgopithecus* and *Theropithecus* (Brain 1981; Delson 1975; Freedman 1976; Jolly 1972; Maier 1971; McKee and Keyser 1994).

Living Papionina are generally characterized morphologically by a steep anteorbital drop, facial fossae, and pronounced sexual dimorphism (Delson 2000). *Papio* has a particularly long muzzle, generally deep facial fossae, and pronounced sexual dimorphism in both cranial size and dentition (Fig. 2.2) (Szalay and Delson 1979). In contrast, *Parapapio*, which was predominantly arboreal, had a generally straighter face, with facial fossae and supraorbital tori less developed than in *Papio*. Dental sexual dimorphism in this taxon is slight (Jablonski 2002). *Pliopapio* is a Plio-Pleistocene Papionin known only from East Africa (Frost 2001). This genus is considered more derived than *Parapapio* as it has a slightly more flexed face, however the muzzle is short and it lacks facial fossa.

In addition to *Parapapio* two other forms quite similar to *Papio* were present in the southern African Plio-Pleistocene. *Dinopithecus sp.*, found in southern Africa from sediments aged ~3.0 - 1.5 Ma, is a large and rugged papionin (Jablonski 2002) and the muzzle lacks

maxillary fossae (Freedman 1957). *Gorgopithecus major* (1.9 – 1.5 Ma), found only in South Africa, is also large and much like *Papio*. However it has a shorter, narrower muzzle, deeper facial fossae and displays less dental sex dimorphism than *Papio* (Freedman 1957 ; Szalay and Delson 1979). *Papio* and *Theropithecus* are also present in the southern African fossil record, as will be discussed in detail in the next section.

Figure 2.2-A photograph of adult male (left) and female (right) chacma baboons skulls shows that* Papio *has a long muzzle, deep facial fossae, and pronounced sexual dimorphism in both cranial robusticity and dentition size (Szalay and Delson 1979).

Analyses based on various craniofacial traits have struggled to resolve the relationships of these fossil Papionins to extant taxa. The concave facial profile of *Pliopapio* suggests that this genus may be more closely related to the extant forms than *Parapapio* (Burrell 2008; Burrell et al. 2009). *Dinopithecus* has in turn been assigned to *Papio* (Delson and Dean 1993), *Mandrillus* (Delson 2000) and *Theropithecus* (Freedman 1957; Jolly 1965), while *Gorgopithecus* could be related to *Theropithecus* (Freedman 1957) or to *Papio* (Delson 2000). More recently Gilbert (2008) used three dimensional geometric morphometric techniques to analyse the basicranium and ultimately construct a phylogeny of extinct and extant Papionins. The two most parsimonious trees recovered suggest that *Parapapio*, *Pliopapio* and *Dinopithecus* are stem African Papionins, and that *Theropithecus* is the most primitive crown African Papionin taxon. This result is congruent with molecular phylogenies based on extant taxa. It was also shown that *Gorgopithecus* is most closely related to *Papio* and *Lophocebus* (Gilbert 2008).

University of Cape

The emergence of *Papio*

Based on fossil evidence for the *Theropithecus*, *Lophocebus*, *Papio* radiation the emergence of *Papio* has been dated to 3.5 - 4.0 Ma (Jablonski 1993). Morphological data alone has been unable to resolve the trichotomy between these taxa (Davenport et al. 2006; Harris and Disotell 1998), however molecular studies which support the split between *Papio* and *Theropithecus* at 3.5 - 4.0 Ma (Harris and Disotell 1998; Page and Goodman 2001) firmly place *Lophocebus* as sister to a monophyletic *Theropithecus* / *Papio* clade (Davenport et al. 2006; Harris and Disotell 1998; Page and Goodman 2001). A graphic summary of the timeline of events leading to the emergence of chacma baboons at ~2.0 Ma is presented in Fig. 2.3.

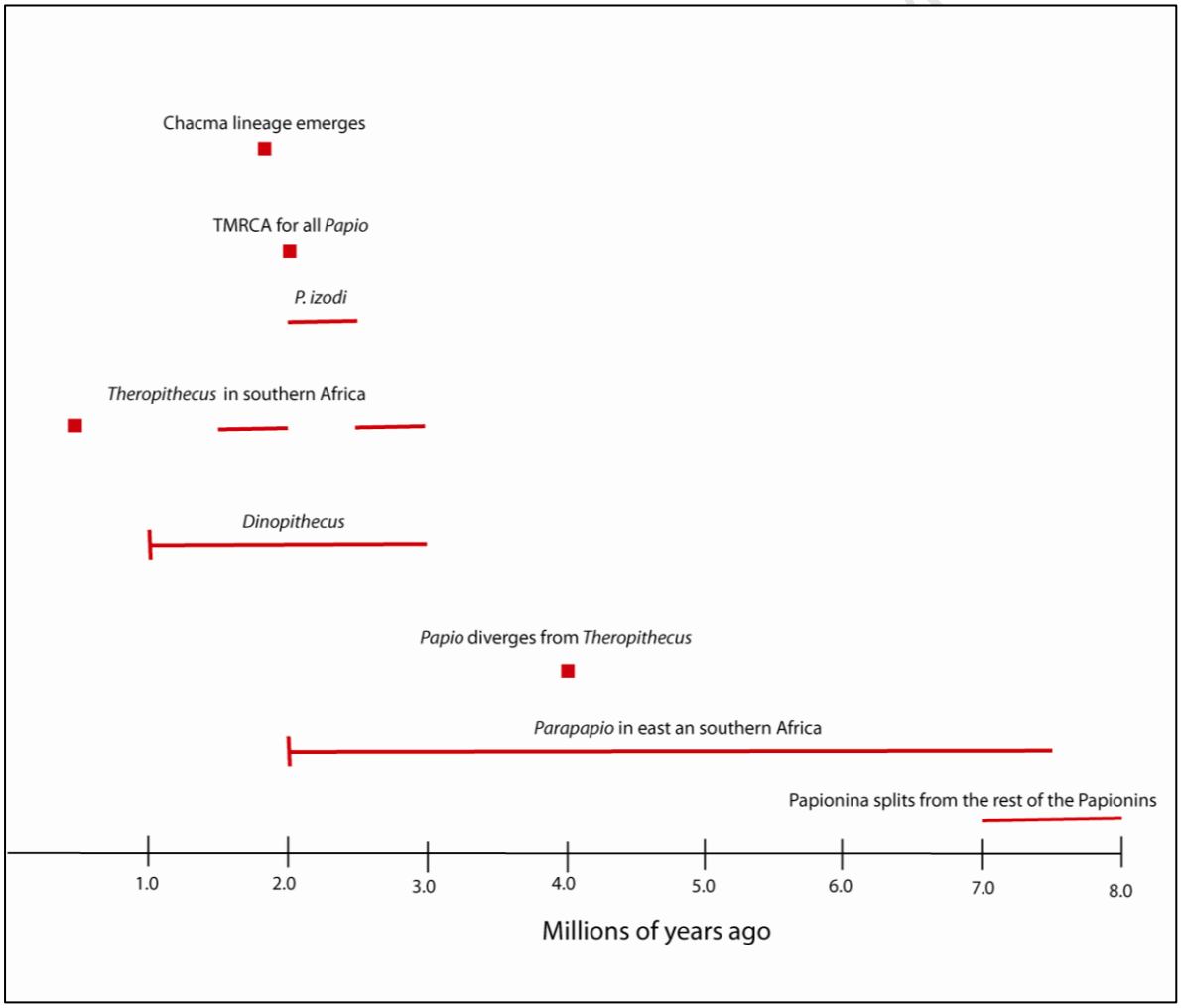

Figure 2.3- This is a graphic summary of the timeline of events leading to the emergence of chacma baboons at ~2.0 Ma. These dates are drawn from various sources that are referenced in the text above.

Both *Papio* and *Theropithecus* are well represented in the Plio-Pleistocene fossil record (Jablonski 1993). Although today modern *Theropithecus* is limited to a single species, confined to the Ethiopian highlands, where it specialises on eating grass blades, seeds, corms and rhizomes; following the emergence of the genus circa 4 Ma (Eck and Jablonski 1987) at least four species have existed, with wide distributions that include the range of present day savannah baboons (Pickford 1993). Indeed, populations of *T. oswaldi* are known from sites in northwestern Africa and even India and Spain (Jablonski 2002).

While the record of *Theropithecus* in East Africa is essentially continuous, it is much less so in southern Africa (Delson 1984). Instead, *Theropithecus* is present in South Africa as *T.darti* at ~3.0 - 2.5 Ma and then again from 2.0 to 1.5 Ma as *T.oswaldii* (Delson 1984). *Theropithecus* only appears again in South Africa at ~400ka. Although this pattern may be an artifact of an incomplete fossil record, the apparent gap in distribution of the species has been interpreted as the result of multiple cycles of local extinction and recolonisation of *Theropithecus* in southern Africa (Pickford 1993). It has been suggested that the global and regional climatic fluctuations characteristic of the Plio-Pleistocene, coupled with predation and competition pressures, may have resulted in the extinction of a number of *Theropithecus* species in southern Africa. This would potentially have left an opportunity for the diversification of savannah-adapted *Parapapio* and ultimately, may have influenced current distributions of the gelada and the savannah baboons (Jolly 1972; Delson 1993; Jablonski 1993).

The abundance and diversity of extant and extinct *Papio* in southern Africa together with the absence of *Papio* fossils in Pliocene deposits in East Africa (Frost and Delson 2002), is often taken as evidence of the southern origin of the genus (Jolly 2001). The fossil record suggests that the modern *Papio* radiation in southern Africa began approximately 3.0 Ma (Delson 1984; Jablonski 2002).The earliest described *Papio* is *P. izodi* [2.6 – 2.3 Ma] (Delson 1988; Klein 1999; Pickering et al. 2004; Pickering and Kramers 2010). This species displays typical *Papio* features particularly in the fronto-nasal region and has a well developed torus and maxillary fossae (Heaton 2007). The first post -- 2.0 Ma baboons are represented by *P. angusticeps* and *P. robinsoni* (Heaton 2007). *P. angusticeps* [2.0 - 1.5 Ma] (Delson 1988; Klein 1999) lacks the derived traits of *P.izodi* and is more similar to the other forms of *Papio*, in particular *P. kindae* (Delson 1984,1988; Williams et al. 2007). *P. robinsoni* from South Africa is dated to ~2.0Ma - 1.5 to 1Ma (Delson 1984). If samples from Sterkfontein, which may be intrusive (Delson 1988), are included, then *P. robinsoni* extends as far back as 2.6 Ma. It is not clear how *P. angusticeps* and *P. robinsoni* relate to the extant baboon species; however some fossil and isotopic data indicate that *P. robinsoni*, which is

University of Cape Town

the most common cercopithecoid in South African Pleistocene deposits, is the most likely ancestor of modern baboons (Delson 1992; Heaton 2007; Lee-Thorp and van der Merwe 1993).

Taxonomy and Phylogeny of modern *Papio*

Baboons have historically been divided into five distinct morphological taxa originally named: the chacma baboon (*Papio ursinus)* (Kerr 1792); the yellow baboon (*P. cynocephalus*) (Linnaeus 1766); the olive or anubis baboon (*P. anubis*) (Lesson 1827); the hamadryas, mantled or sacred baboon (*P. hamadryas*) (Linnaeus 1766); and the Guinea baboon (*P. papio*) (Linneaeus 1766). These five taxa are parapatrically distributed with varying degrees of interbreeding at range boundaries (Jolly 1993). The contradiction between phenotypic distinctiveness and genotypic continuity has resulted in an extended debate over the taxonomic status of the five baboon morphotypes (Jolly 1993). This conflict produces two taxonomic classification systems for baboons. In the first, all taxa are ascribed full species status as indicated above, and in the second, taxa are ascribed as subspecies of *P. hamadryas* (hamadryas has priorty as it was the first to be described in the literature). Under the second system chacma baboons would be *P. hamadryas ursinus*. Here I choose to refer to each major morphotype as a separate species as is suggested by recent genetic data (Zinner et al. 2009). Also conflating all taxa into a single species does not acknowledge the independent evolutionary histories that would have been necessary for the fixation of such distinct morphological and behavioural differences between taxa. I therefore refer to all five morphotypes as full species for the rest of this thesis.

Although these five major morphotypes or species have been historically recognised as (fairly) distinct, the divisions between these groups are not. This is due, in large part, to hybridization at contact zones between taxa (Alberts and Altmann 2001; Jolly 1993; Nagel 1973; Phillips-Conroy et al. 1991; Samuels and Altmann 1991; Sugawara 1979) resulting in high levels of phylogenetic complexity within the genus (Zinner et al. 2009). Alternative scenarios describing the evolutionary relationships among extant *Papio* species have been proposed based on morphological and behavioural data. These are presented below.

1. The savannah hypothesis holds that the unique synapomorphies in behaviour, phenotype and ecology of savannah baboons, which are absent in hamadryas baboons, set the two lineages apart (Fig. 2.4a), thereby suggesting a phylogenetic tree in which hamadryas baboons are sister to all other baboons (Buettner-Janusch 1966; Thorington and Groves 1970).

University of Cape Town

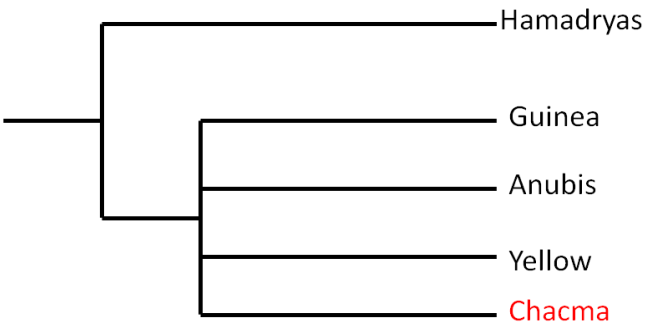

Figure 2.4a- The savannah hypothesis places hamadryas baboons as the most divergent species while all other baboons are paraphyletic (Buettner-Janusch 1966; Thorington and Groves 1970).

2. A second phylogeny uses morphological characteristics to set up a north - south dichotomy within *Papio* (Frost et al. 2003; Jolly 2001; Jolly 1965) (Fig. 2.4b). This interpretation is based on a suite of shared cranial morphometric and pelage characters, with northern baboons being the "maned" baboons (anubis,hamadryas and guinea) while the southern group are "unmaned" (chacma and yellow).

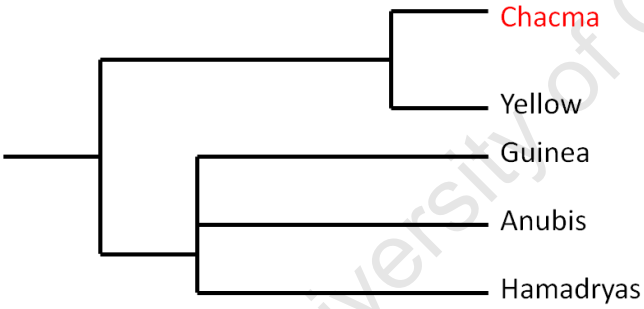

Figure 2.4b- This phylogeny separates northern and southern baboons into "maned" (anubis, hamadryas and guinea) and "unmaned" (chacma and yellow) types (Jolly 1965, 2001; Frost et al. 2003).

University of Cape Town

Recent studies using molecular data have also revealed a number of alternative scenarios in describing phylogenetic relationships among living baboons. Early molecular work using blood protein data (Williams-Blangero et al. 1990) and DNA sequence data from the mitochondrial cytochrome oxidase II gene (Disotell 1992) found *P. papio* to be the most divergent of the baboons, with little differentiation between the other four subspecies (Fig 2.5a). Later studies by Newman et al. 2004 and Wildman et al. 2004 using mitochondrial DNA sequence data, revealed at least four distinct maternal haplogroups within their datasets, with chacma baboons being the most divergent (Fig. 2.5b, 2.5c). Wildman et al. (2004) also recovered a taxonomically unresolved clade containing hamadryas, yellow and olive baboons.

More recently, a phylogeny of *Papio* (Zinner et al. 2009) based on the same mitochondrial marker as used by Newman et al. (2004) and Wildman et al. (2004), but with a more representative geographic sample across the range of *Papio,* revealed geographic rather than taxonomic clustering of mitochondrial haplotypes across *Papio* (a phenomenon first described by Wildman et al., 2004). This phylogeny is represented in Fig. 2.5d. Zinner et al. (2009) suggest that introgression and hybridization may have played a significant role in the evolutionary history of *Papio,* an idea initially proposed by Jolly (Jolly 1975, 2001) and demonstrated at several contact zones across east and southern Africa (Alberts and Altmann 2001; Burrell 2008; Nagel 1973; Jolly and Brett 1973; Jolly 1993; Phillips-Conroy et al. 1991, Samuels and Altmann 1986; Sugawara 1979; Zinner et al. 2009). This molecular result is highly congruent with recent morphological findings which, using landmark based geometric morphometrics and multivariate analysis, show how the cranial morphology of extant baboons differs in a stepped cline, with the greatest difference between northern and southern populations (Frost et al. 2003). This supports geographic rather than taxonomic clustering of individuals within a phylogeny.

University of Cape Town

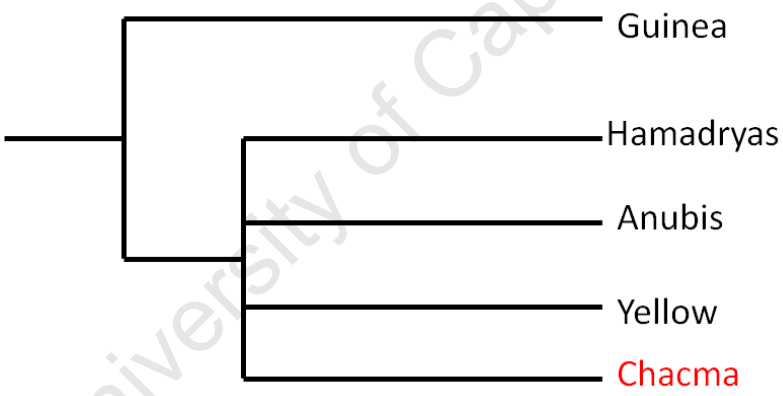

Figure 2.5a- Blood protein (Williams-Blangero et al. 1990) and cytochrome oxidase II gene (Disotell 1992) data place* P. papio *as the most divergent of the baboons, with little differentiation between the other four species.

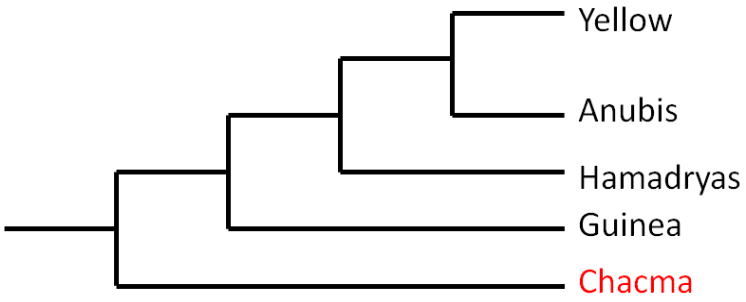

Figure 2.5b- Phylogenetic reconstruction based on mitochondrial Brown region sequence data places chacma baboons as the most divergent and basal to all other species (Newman et al. 2004).

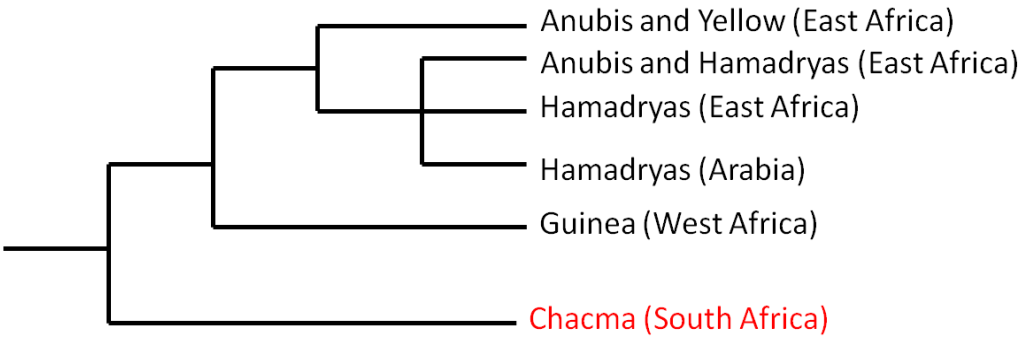

Figure 2.5c- Phylogenetic reconstruction based on mitochondrial Brown region sequence data places chacma baboons as the most divergent and basal to all other species and recovers a taxonomically unresolved clade containing hamadryas, yellow and olive baboons. In brackets are the geographic regions from which the samples were sourced (Wildman et al. 2004).

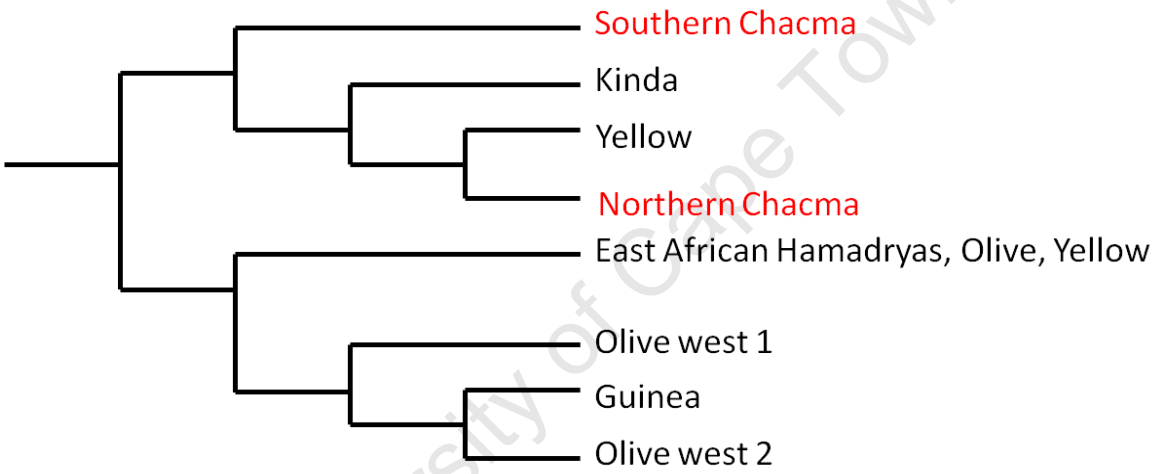

Figure 2.5d- A phylogeny of* Papio *based on on mitochondrial Brown region sequence data the but with a denser sample than previous studies revealed geographic rather than taxonomic clustering of mitochondrial haplotypes (Zinner et al. 2009).

University of Cape Town

It is important to note that a recently described Papionin, the kipunji (*Lophocebus kipunji* or *Rungwecebus kipunji*) (Davenport et al. 2006), might have some bearing on the phylogeny of the genus *Papio*. Indeed, there is some controversy with regards to the evolutionary placement of this new taxon (Zinner et al. 2009a, Burrell 2009). The species has an overall mangabey-like appearance including very deep suborbital fossae, a short rostrum, and relatively short and wide scapulae, and is thought to be predominantly arboreal (Jones et al. 2005). These observations suggested that *kipunji* is more closely related to *Lophocebus*. However, more recent work has documented that these animals are more terrestrial than previously thought (Davenport et al. 2008) and molecular studies suggest a closer relationship to *Papio*. It remains unclear where *kipunji* falls within the *Papio* phylogeny; it may be sister to (Zinner et al. 2009a) or fall within the genus *Papio* (Burrell et al. 2009).

Clearly the phylogenetic reconstruction of relationships of species within *Papio* requires further assessment. While Zinner et al. (2009) have produced the most statistically robust tree arrangement to date, this reconstruction is based solely on the maternal (mitochondrial) history of species. Further analysis using nuclear DNA sequence data and possibly an integration of morphological and molecular data, as was achieved for Papionins (Singleton 2002), is likely to provide the most accurate representation of the evolutionary history of the genus.

Baboon distribution and ecology

Baboons are large, terrestrial primates. Excluding the somewhat atypical hamadryas baboon, they exhibit the conserved Old World monkey behavioural pattern of male-biased dispersal whereby young adult males leave their natal group to seek out potential mates in neighbouring groups (Rasmussen 1979) although one might add, that many populations, including Guinea baboons have yet to be investigated. Highly adaptable and omnivorous, they are distributed across Africa (Fig. 2.6) where they occupy a variety of habitats, including open grassland, savannah, delta, montane, semi-desert, and closed forests (Aldrich-Blake and Richtsmeier 1999; Altmann and and Altmann 1970; Cowlishaw and Hacker 1997; Henzi and Weingrill 1999; Rowell 1966; Whiten et al. 1991). All members of *Papio* are considered dietary generalists (Altmann and Altmann 1970; Norton et al. 1987; Norton and Ashley 2004; Rowell 1966).

Below is a summary of the distinguishing ecological characteristics of each of the five traditional *Papio* species. It is important to note that this is a highly generalized description of each species as extensive geographic variation has been reported *within* taxa, adding to the phenotypic diversity and phylogenetic complexity across the genus:

- Also known as desert baboons, ***hamadryas baboons*** are confined to the semi-desert, eastern lowlands of Somalia, Ethiopia and the eastern highlands of Saudi Arabia and Yemen up to 2600m altitude. Here they forage in arid sub-desert and savannah and woodland vegetation and eat grass seeds, roots, tubers, animal prey and invertebrates eg termites. Unlike other baboons in the genus *Papio*, hamadryas baboons have adopted female-biased dispersal (Jolly 1993; Kingdon 1997). Hamadryas baboons are quite distinct from the other baboons in that they have pinkish faces and display extreme sexual dimorphism in colour. Adult males are grey, while females are olive brown. Males develop prominent mantles at adulthood (Hill 1970).

University of Cape Town

- The ***guinea baboon*** has a relatively small range in the extreme west of Africa, with parts of its range extending into Senegal, Mauritania, Guinea and Sierra Leone. These baboons live in evergreen gallery forest and woodland savannah and will eat fruit, seeds, flowers and animal prey. Guinea baboons are observed to sleep in palm and kapok trees (Jolly 1993, Kingdon 1997). The guinea baboon is a smallish baboon, reddish brown in colour. Males develop an elongated mantle on the shoulders by adulthood (Hill 1970).

- Between the distributions of hamadryas, guinea and yellow baboons is the range of the ***olive baboon***, which extends across equatorial Africa between the Sahara Desert, to about 13° north of the equator. Olive baboons live in semidesert, thorn scrub, savannah, woodland and rainforest up to 4500m. They eat fruit, seeds, tubers, roots, leaves, flowers, animal prey (Jolly 1993; Kingdon 1997). The olive baboon has a large black face and is a dark olive brown colour. Although not as distinct as in hamadryas, these baboons are also considered to have mantles (Jolly 1993). They are distinguished from yellow baboons by being much stouter. In external form they are most similar to chacma baboons (Hill 1970).

- South of the range of olive baboons are the ***yellow baboons*** in southern equatorial and East Africa, distributed from the Benguela Province in Angola across to the northern region of Mozambique, and extending along the east African coast northwards to the Webe Shebeli valley in south-eastern Ethiopia and Somalia. They occur in thorn scrub, savannah, woodland and rainforest up to 1000m and eat fruits, seeds, leaves, flowers, tubers, roots, bulbs leaves, flowers, animal prey (Jolly 1993; Kingdon 1997). The yellow baboon is relatively gracile when compared to the other baboons and has a smaller head and long limbs. As the name suggests, these baboons appear yellow in colour. The have whitish hairs on the face and hairs on the shoulders do not form a mantle.

- The southernmost baboon is the ***chacma baboon*** which occupies most of southern Africa (Jolly 1993). It is widely distributed and is the only baboon south of the 33rd parallel with fifty percent of its range below the tropic of Capricorn (Anderson 1982). Chacma baboons are found south of the range of yellow baboons (Jolly 1993) extending into southwest Angola, southern Zambia, southern Mozambique, Namibia,

University of Cape Town

Botswana, South Africa, Swaziland and Lesotho (Nowark 1999). They inhabit woodland, grassland, acacia scrub, semi-desert, coastal zones and mountains up to 2980m altitude. They eat fruits, seeds, flowers and animal prey and also shellfish. (Jolly 1993; Kingdon 1997). Chacma is the largest and darkest of the baboons and is unmantled (Hill 1970).

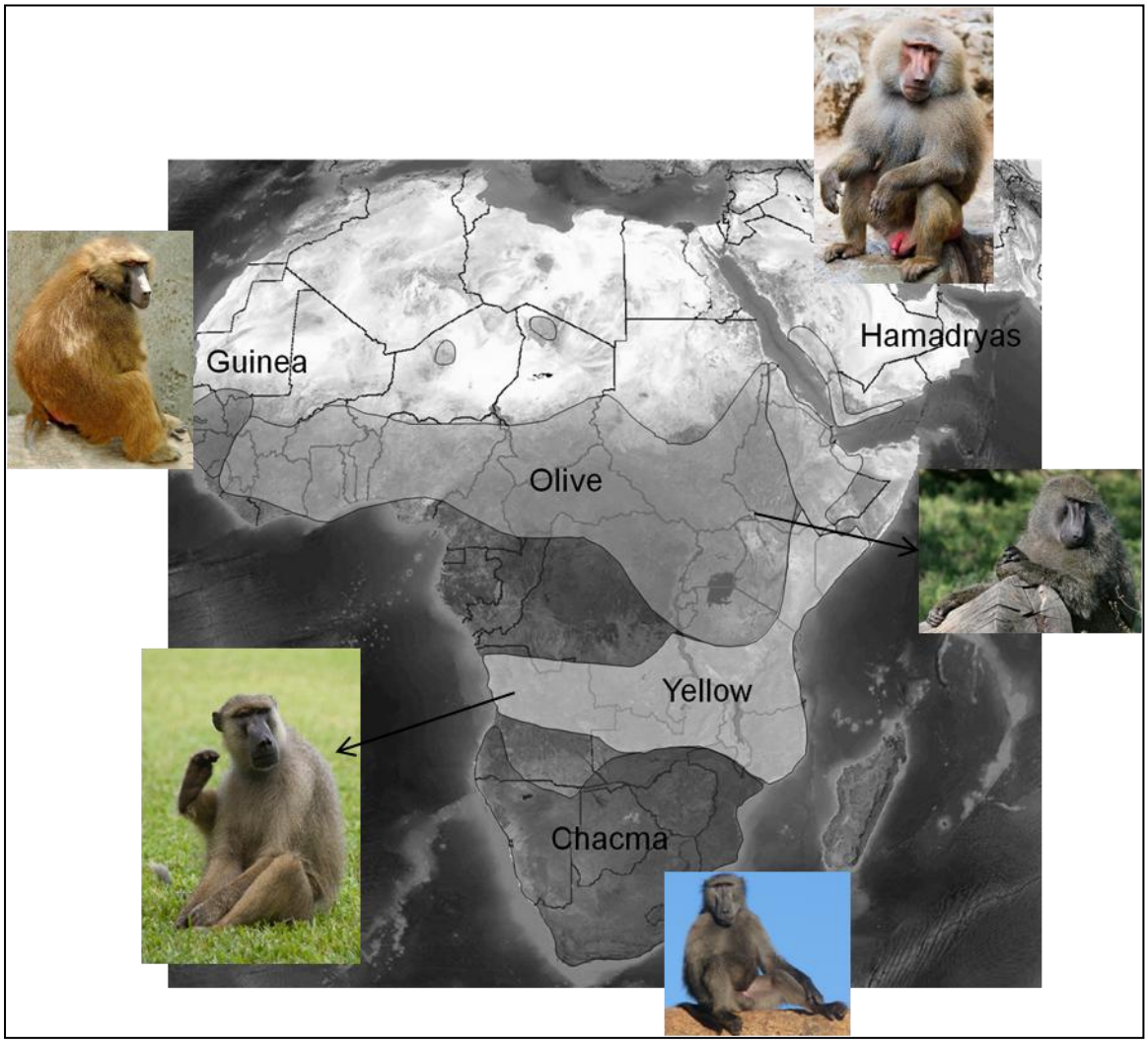

Figure 2.6- An artistic representation of the general distribution of each of the five Papio taxa. This is based on a synthesis of distribution information in Jolly (1993) and Kingdon (1997). Images of each of the five species are included to show the distinctiveness of each phenotype.

Summary

Southern Africa has played an important role in the emergence and evolution of baboons. Numerous Papionins have extensive fossil records in the region, with the genus *Papio* likely emerging in southern Africa circa 3.0 Ma, suggesting a southern origin of the modern Papio taxa. Within the genus *Papio*, there is considerable variation among the extant species in terms of morphology, geographic range, ecology, and behaviour. Although a number of different baboon phylogenies have been proposed, recent evidence suggests that chacma baboons are among the oldest of the baboon lineages. The next chapter will focus on variation within this deep lineage.

.

University of Cape Town

CHAPTER 3

A DESCRIPTION OF THE SOUTHERN AFRICAN CHACMA BABOON (*P. URSINUS*) AND THE HABITATS ACROSS WHICH THE SPECIES IS DISTRIBUTED

Introduction

Chacma baboons have been evolving independently for almost as long as modern *Papio* has been in existence. Mitochondrial evidence suggests that the genus *Papio* appeared in Africa by at least ~2.09 Ma and the emergence of proto chacma baboons is dated to at least ~1.80 Ma (Newman et al. 2004; Sithaldeen et al. 2009; Zinner et al. 2009). Chacma baboons therefore offer an almost equivalent temporal record of genetic subdivisioning and population responses to climate change as ancestral *Papio*. Chacma is a southern African baboon and an ecological generalist, which is arguably the most likely ancestral state of all *Papio.* Fossil evidence (El-Zaatari et al 2005; Frost and Delson 2002) suggests a southern African origin for the genus (Jolly 2001). Chacma baboons could therefore serve as a suitable proxy for modelling the earliest baboon populations in many ways (Henzi and Barret 2003) including responses to climate change and habitat fragmentation. They may also provide insight into understanding diversification in widely distributed, medium to large mammals in southern Africa more broadly. Here I summarise some relevant background to the species, its habitat, and southern African climate/habitat change over the course of the Pleistocene. I also provide some context fo later discussions by briefly summarizing evidence for climate-driven lineage diversification in other southern African species.

University of Cape Town

Ecological variation across the distribution of chacma baboons

Although the distribution of chacma baboons is largely continuous, only a subset of populations have been systematically observed and reported on. The sites of these observations are plotted in Fig 3.1 and presented along with an insert of Mucina and Rutherford (2006) biome classifications to show the diversity of habitats across which chacma baboons are distributed. What follows is a summary of the available literature documenting the distribution of chacma baboons and include my personal observations (pers. obs.) as well as personal communications (pers. comm.) from a number of informed

sources. Museum records are not included. The purpose of this summary is to show the range of habitats within which chacma baboons are currently distributed.

The chacma baboon populations found on the Cape Peninsula in the Western Cape region of South Africa represent the "typical" or nominate chacma phenotype and are unique for two reasons. Firstly, they are the southernmost baboon and non-human primate population in the world, and secondly, they are the only baboon known to forage along the marine littoral zone where they have been observed consuming shellfish (Davidge 1978). Baboons are a common sight in the Western Cape, much of which is dominated by the winter rainfall *fynbos* biome, and can often be found foraging along main thoroughfares (pers. obs.). Baboon troops of varying sizes can be found throughout the Cape Fold mountain range, at sites such as Klein Klip Huis (KKH) just outside Clanwilliam, in the Cedarberg mountains (pers. obs.) and along the Pakhuis Pass (Will Archer, pers. comm.). Other troops have been observed and noted on the west coast at Elandsbaai (pers. obs.) and at Dwarsrivier (Will Archer, pers. comm.). Between one and two troops reside in the Kagga Kamma National Park (pers. obs.). East of the Cape peninsula near Bredasdorp, resides an intensively studied habituated baboon troops within the De Hoop Nature reserve (Barrett et al. 2002; Henzi and Weingrill 1999; Henzi et al. 2003).

The south coast of South Africa receives annual rainfall that is bimodally distributed and supports forest thicket in which baboon troops can be found, e.g. Natures Valley in Tsitsikamma (pers. obs.). Chacma baboon troops are also found all the way along the eastern coast of South Africa and as far north as the St. Lucia wetlands (pers. obs.) and Kosi Bay close to the Mozambiquan border (CROW[2] records). Here they are supported by coastal vegetation and savannah resources. Inland, baboons are distributed across the short montane grasslands of the Drakensberg mountain range (Henzi et al. 1997; pers. obs.) and the Lesotho highlands (Charlie Arthur, pers. comm.). These distributions are not necessarily continuous, particularly across farmland.

The range of chacma baboons extends across the northernmost border of South Africa in the Limpopo province, and there are many troops within the Kruger National Park (pers. obs.) which is covered by deciduous lowveld savannah vegetation (Codron 2003). Baboon troops have also been observed at high altitude in the Waterberg mountains (pers. obs.). Here the savannah is dominated by sour bushveld (tall trees with sour grassveld) and is a temperate area of high productivity (Codron 2003).

[2] CROW is the Centre for Rehabilitation of Wildlife based in Kwa-Zulu Natal, South Africa

University of Cape Town

In the northern central part of South Africa, baboons have been identified in Sandmansfontein (Marais 1939) and other areas of the highveld (Brain 1981). They are also known to live in the less densely occupied areas around the towns of Rustenberg, Kimberley (pers. obs.) and Bloemfontein (Nick Wiltshire, pers. comm.). Baboon populations have been documented in Suikerbosrand National Park (Antoinette Kruger, pers. comm.), Loskop Dam (Nic Fourie, pers. comm.) and Blyde River Canyon, Mpumalanga (Hannes Marais, pers. comm.). Suikerbosrand (Anderson 1992) is a high altitude mountain range which can experience frost for more than 150 days per year. Like their predecessors, baboons still occupy habitats in the Cradle of Humankind, such as the Makapansgat Valley (pers. obs.) and the Buxton Limeworks site in Taung (McGrew et al. 2003).

Troops can also be found in the southern parts of the South African interior often coinciding with farming areas such as Graaf Reinet (pers. obs.) and Baviaanskloof and the Karoo towns of Willowmore and Steytlerville where Nama Karoo is the natural vegetation type (Acocks 1988). There is a conspicuous break in baboon distribution to the north in the Kalahari Desert where South Africa borders Botswana, and there are no baboons in the Kgalagadi Transfrontier Park (Dudu Job, pers. comm.).

West of the Kalahari Desert the coastal region is dominated by succulent karoo vegetation and baboon populations can be found up the South African coast, across arid Namibia (Schalk van Wyk, Harnas Wildlife Rest, pers. comm.), and as far north as the lush Caprivi strip (pers. obs.) and into southern Angola (Nowark 1999). Most Namibian baboons, like the resident troop at the Ais Ais hot spring resort at the southern end of the Fish River Canyon (pers. obs.), survive in an extremely arid environment. In central Namibia, thornbush-savannah is dominant, with extensive grasslands and acacia bush. Toward the north-east, where there is a higher rainfall, the thornbush savannah slowly turns into mopane savannah and there are a greater number of trees. In the relatively humid Caprivi, the dominant vegetation form is the woodland savannah, interspersed with single baobabs, wild figs and makalani palms.

Two populations in Namibia have been reported to survive in the most arid of conditions, clearly illustrating the extreme adaptability of these animals. The Tsoabis Leopard Park, situated in the Namib Desert, Namibia, is home to a population of baboons who subsist on transition vegetation (Giess 1971). The Namib Desert is a narrow strip of land between 80 – 150 km wide that stretches for almost 2000 km along the south-west coast of Africa (Davies and Cowlishaw 1997). Here, rainfall is scarce with an average of 85mm per annum (Cowlishaw and Davies 1997). This area is rocky, largely barren, and is dominated by

University of Cape Town

mountains and ravines. The vegetation is riparian woodland and very sparse. There is little food and shelter here and baboons have been observed to supplement their diet with vertebrate prey, including herpetofauna, birds and mammals (Davies and Cowlishaw 1996).

The second extremely arid adapted baboon population lives in the Kuiseb canyon. This area has an average annual rainfall of 18mm (Anderson 1982). For most of the year, there is no surface water, and although some dense acacia groves dot the landscape, there are generally very few types of trees (Hamilton et al. 1975). Baboons here cope with water stress, and, although they are obligate drinkers, have survived (and thrived) in very water restricted environments. They do this by employing several strategies; they may use food resources to supplement their water intake (Devore and Hall 1965; Hamilton III 1985) or baboons may "steal" water by exploiting *Oryx* dug holes (Hamilton 1985). Finally, baboons will often reduce their activity to conserve water. Brain (1992) observed that baboons could survive for up to 11 days without drinking water.

In the north east of Namibia and across the border into Botswana chacma baboons occupy a completely different environment from their southern counterparts as they forage along the Okavango River with its fertile floodplain (Cheney et al. 2004). These baboons look different from the typical chacma and may represent a gradation into the yellow baboon (Becky Ackermann, pers. comm.). One troop which lives 100km north of Maun in Botswana, live on a heavily vegetated floodplain with woodland islands surrounded by grasslands. This area is very humid with many fruit trees and a high diversity of other foods available to the baboons (Hamilton et al., 1975).Chacma baboons are also known from Zimbabwe, Zambia, Mozambique and Angola (C. Jolly, pers. comm.)

To summarise, chacma baboons occupy a wide range of environments in southern Africa, from coastal *fynbos* regions to frost-prone mountainous terrain to extremely arid deserts. There appear to be few environments that they are not capable of occupying, with only the Kalahari Desert conspicuously bereft of baboons. Reflecting this wide ecological range, chacma baboons are considered to be feeding generalists. Their diets are dominated by fruits, leaves and subterranean items, while flowers and animal matter constitute a much smaller proportion of the diet. Clearly some groups are also able to survive for considerable amounts of time without water, often by adapting their diets appropriately. However, it is worth noting that among thousands of plant species within their home range, a baboon group may use only a hundred species, and use only certain parts of each (e.g. only young leaves, only roots, etc (Alberts et al. 2005; Altmann 1998; Byrne et al. 1993; Norton et al.

University of Cape Town

1987) so that within a set of resources, baboon feeding can actually be highly selective (Altmann and Altmann 1970; Whiten et al. 1991).

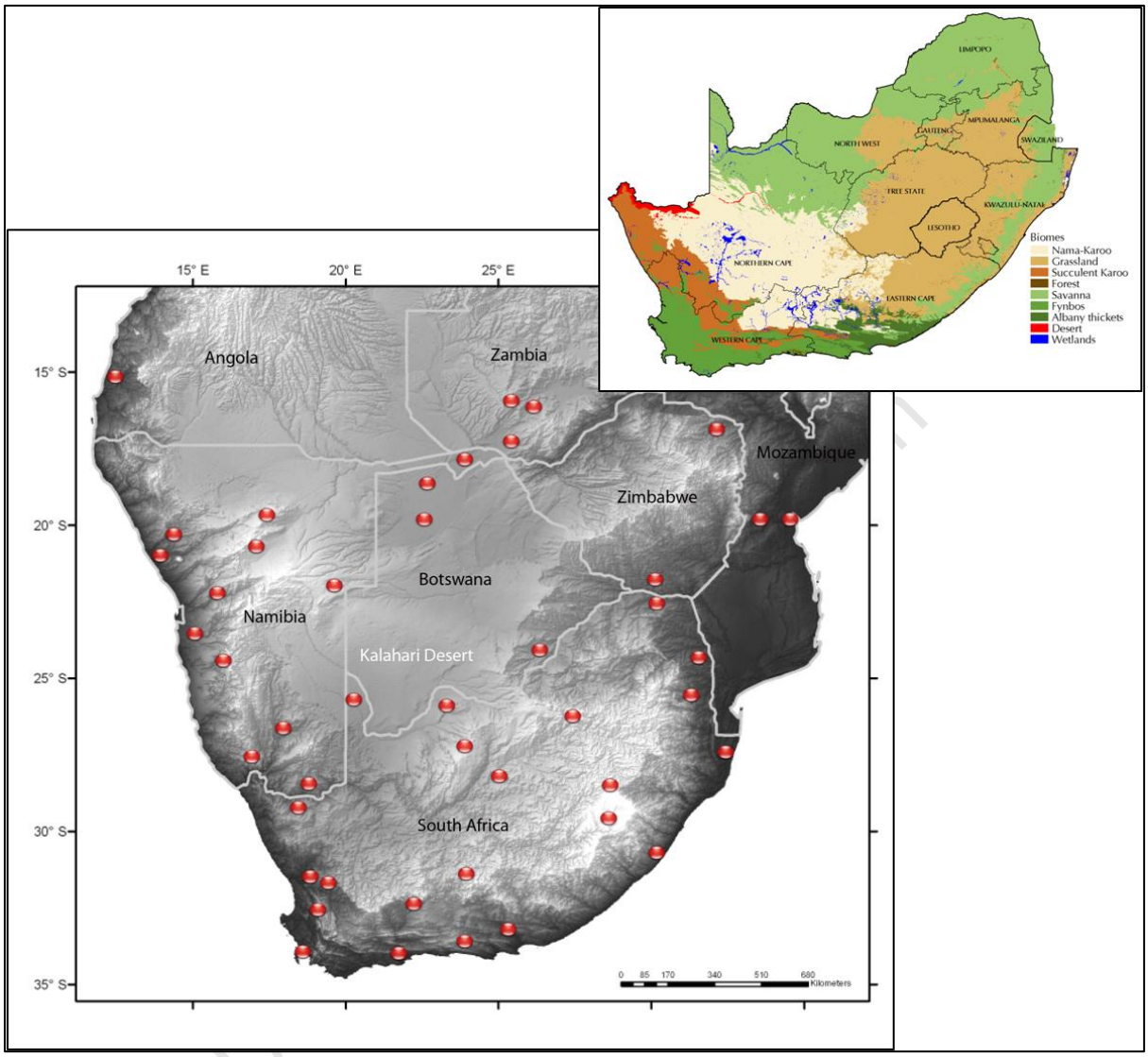

Figure 3.1- Map shows a summary of sites in which baboons have been observed and or documented, not including museum records. The insert is a biome classification map of South Africa (Mucina and Rutherford 2006) which serves to illustrate the diversity of habitats in this region across which chacma baboons are distributed.

Phenotypic variation within chacma baboons

The ecological variability of chacma baboon populations is accompanied by significant phenotypic variation, particularly in size and pelage colouring (Groves 2001; Jolly 1993). The nominate or typical chacma phenotype (*P. ursinus ursinus)* is shown in Fig. 3.2 and is described below. This is the variant against which all others are compared.

"*Chacma baboons are the largest and darkest of the baboons and generally have an almost black pelage The male does not have a mane but will usually have long tufts of hair along the nape. The chacma has no greenish tinge in its pelage and no heavy side whiskers on its face. The tail has a distinctive kink at about 1/3 from the base that makes it look broken. The face, ears and naked parts of hands and feet are deeply pigmented. The face is sparsely covered with short whitish hair. Body hair is coarse and predominantly dark with variations in speckling*" (Hill 1970; pg. 319).

Variation in pelage is the basis of several classification schemes that have been used to identify up to eight distinct chacma forms (Hill 1970; Appendix 3A). The distributions of these forms are represented in Fig. 3.3. Today only three forms are generally recognised:

1. *Papio ursinus ursinus* (Kerr 1792), the typical chacma, is a large baboon with black nape fringes, dark brown fur, black fur on hands and feet and a relatively short tail. This variant occurs in the more southerly and westerly part of the chacma range, including South Africa and some parts of Botswana This group incorporates Hill's (1970) *ursinus*, *orientalis* and *occidentalis* subspecies.

Figure 3.2 - Adult male chacma baboon from Pringle Bay, Western Cape, South Africa represents the typical or nominate chacma variant (*P. ursinus ursinus*).

2. *P. u. griseipes* (Pocock 1911), the grey-footed baboon, is more fawn coloured and found in southwestern Zambia, Zimbabwe, in Mozambique south of the Zambezi and in parts of the Limpopo province region of South Africa and the Okavango Delta, Botswana (Jolly 1993). These are smaller than the typical chacma and have grey hands and feet the same colour as their limbs, and a longer tail. This group incorporates Hill's *griesipes*, *ngamiensis*, *chobiensis* and *jubilaeus* subspecies. Today it is clear that *jubilaeus* is in fact a yellow baboon and not a chacma at all.

3. *P. u. ruacana* is a small black footed baboon that is darker than *P. u griseipes* and smaller than *P. u. ursinus*. They are found in Namibia and southwestern Angola (Groves 2001).

Phenotypic variation within the species is readily observable by eye and suggests that accompanying genetic variation and structure is likely.

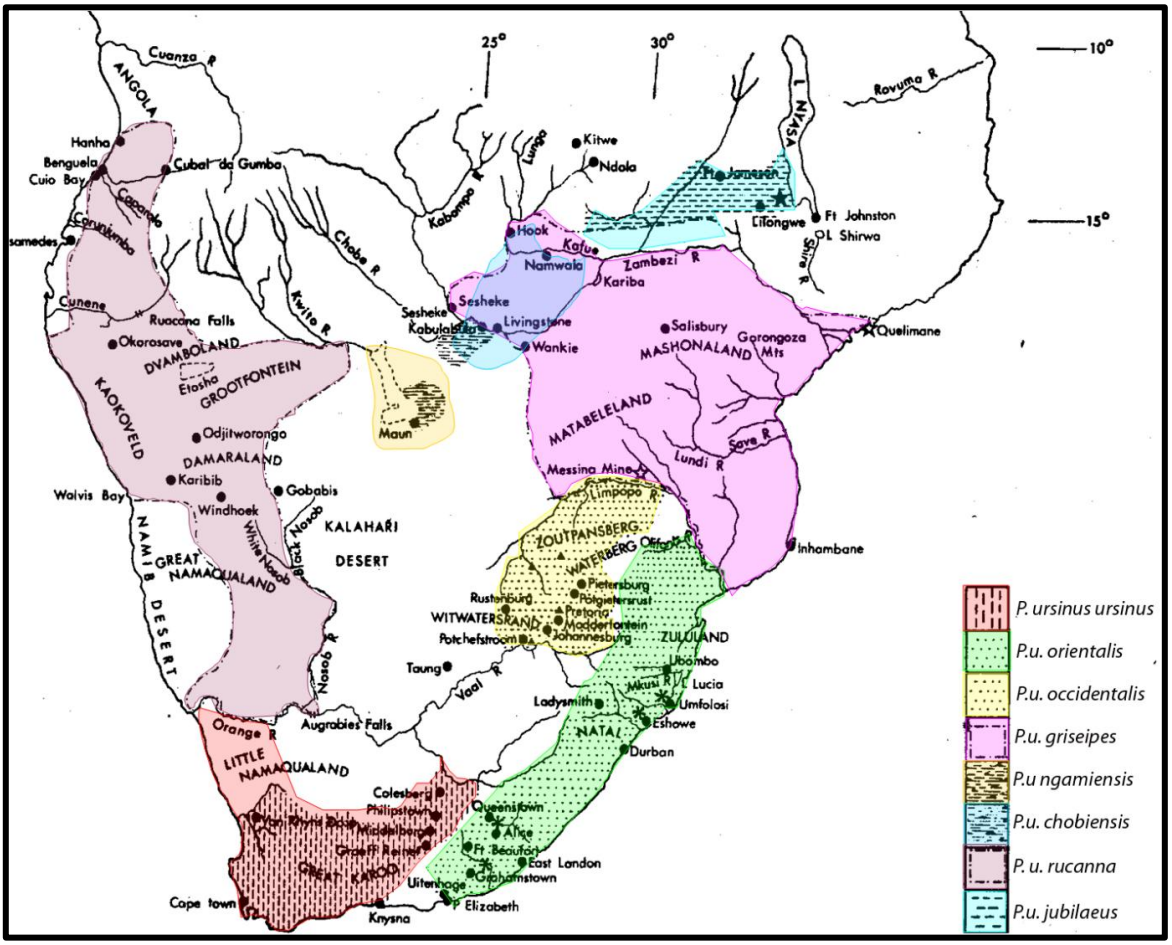

Figure 3.3: Map showing the distribution of chacma morphotypes as summarised by Hill (1970). The map is scanned directly from source and adapted with colour overlays.

Modern habitats of the southern African (chacma) baboon

As shown in Fig 3.1, and not taking into account localised extinctions through anthropological interference, chacma baboons are essentially continuously distributed throughout southern Africa. Southern Africa is defined as the portion of the African continent spanning from 17.8° south of the equator to Cape Aghulas at the tip of Africa. The subcontinent includes the countries of Angola, Botswana, Lesotho, Mozambique, Namibia, South Africa, Swaziland and Zimbabwe (Meadows 2001). Most of the southern African landscape is made up of a highland plateau at a mean elevation of about 1000m above sea level. The plateau is divided into a north western and south eastern terrace by the low lying area of the Kalahari Basin. The central plateau is surrounded by the main fold mountain ranges to the east, west and south. This rim of mountains falls steeply to the coasts. The major mountain chains of Southern Africa are not continuous.

The climate of the sub-continent is strongly influenced both by the latitudinal position of the landmass as well as the nature of ocean currents along its coastlines. Together these variables create two strongly seasonal precipitation regimes. The majority of the subcontinent in the north and east is a summer rainfall zone (SRZ), which means that it receives the majority of its rainfall in the austral summer between October and March. A portion of the southern and western coast, however, receives the majority of its precipitation in the austral winter, between April and September. This is the winter rainfall zone (WRZ). Between these two primary regions is a zone of year round rainfall (YRZ) (Chase and Meadows 2007). Evidence suggests that the seasonal rainfall regimes of southern Africa have been in place at least since the late Pleistocene, although their relative positions may have shifted (Chase and Meadows 2007).

Southern Africa is considered to be generally dry. The west coast in particular, which is bordered by the cold Benguela current, is very arid, while the east and south coasts are bordered by the warm Mozambique and Aghulas currents, respectively, which brings more rainfall to the eastern part of the subcontinent (Deacon and Lancaster 1988). This sets up a pattern of increasing rainfall as one moves from southwest to northeast across the subcontinent and this is reflected in the shifting vegetation biomes across the region.

South Africa is made up of nine (Mucina and Rutherford 2006) distinct vegetation biomes which are shown in Fig 3.1. The distinction between these is based on the dominant plant forms in the region and the prevailing climatic factors. As chacma baboons are primarily vegetarian, the boundaries of these biomes can also serve as boundaries to distinguish

University of Cape Town

between baboon populations with different feeding ecologies. The seven biomes in which chacma baboons are observed are summarised from Ackocks (1988) and Mucina and Rutherford (2006) and presented below:

1. The *savannah* is the largest biome. Here the vegetation is dominated by herbaceous woodland. This biome experiences the largest temperature ranges of all six biomes and the most severe winters – i.e. it is highly seasonal. Mean annual rainfall is 235 mm.
2. The *grassland* biome is found in the high central plateau. It is dominated by grasses and frost is common. Mean annual rainfall is 400-2000 mm.
3. The *Nama-Karoo* is dominated by Karoo shrub and grass and receives a mean annual rainfall of 100-520 mm.
4. The *succulent Karoo*, which as its name suggests is characterised by the dominance of very small, largely endemic succulents, is in the winter rainfall region of the Karoo and receives a mean annual rainfall of 20-290 mm.
5. The *desert* biome has a drainage system that is usually dry and is characterized by mobile dunes. Vegetation is dominated by therophytes. Temperatures do not generally drop below zero. Mean annual rainfall is 13 mm in the west to 70 mm upland. This is an area of extreme water stress.
6. The *forest* biome is the smallest biome in South Africa, located only in a small region on the central southern coast. It is primarily composed of evergreen canopy and receives a mean annual rainfall of between 1000-2000mm.
7. The *fynbos biome* dominates the WRZ. It is a Mediterranean-type heathland, and the major component of the Cape Floristic Region (Tolley et al. 2006). Although this region does experience some frost, especially in the higher areas, winters are not particularly severe. Mean annual rainfall is between 210-3000 mm.

University of Cape Town

Two geographic regions in southern Africa are of particular interest to the population fragmentation patterns discussed later in this thesis and so are described below:

The Kalahari Desert

In southern Africa, many depositional coastal landforms such as dunes, estuaries and the aeolian features of the semiarid and arid regions formed during the last 2myr, including The Kalahari Desert (Meadows 2001). This desert forms a significant portion of the interior of central southern Africa and it is described as a 2.5 million km^2 sand sea (Shaw and Thomas 1996). Although it crosses vegetational biomes, The Kalahari Desert is often defined as a single ecological unit that currently extends from 22.1^{o}S to the Orange River at 29.1^{o}S. Although dominated by the linear dunes, this desert is largely stabilized and occupied by

savanna woodland vegetation (Meadows 2001) although it does become progressively more arid in a southwesterly direction (Shaw and Thomas 1996). Dune activity today is therefore restricted to the south western corner of the desert. Evidence from the late Pleistocene show that cycles of aridity in the past also resulted in periods of dune crest activation in the north-west. Together with the dunes in the south west, these formed the Mega Kalahari Sands (Stokes 1998). During the last interglacial-glacial cycle, four periods of reactivation are documented at 95-115kya, 41-46kya, 20-26kya and post-20 kya (Stokes 1998). These periods of aridity are separated by depositional hiatuses that have been interpreted as wetter periods (Stokes 1998). The Kalahari region is also significant in that topographically, it is relatively flat. Today the interior of the Kalahari represents a barrier to the distribution of chacma baboons.

The Cape Floristic Region (CFR)

Unlike the rest of southern Africa, the southwestern tip of Africa is a winter rainfall zone. The combination of Mediterranean type climate, local geology, topography, soils and environmental history, has given rise to a unique vegetation type known as the *fynbos* (Meadows and Baxter 1999). It is thought that the *fynbos* became established as a major biome in southern Africa between 3.0 and 5.0 Ma (Cowling et al. 1992; Linder et al. 1992; Scott 1999). The *fynbos* is known for its incredible plant diversity, comparable even to the tropical rain forests (Cowling et al. 1992; Galley and Linder 2006; Goldblatt and Manning 2002), and a high degree of endemism. This suggests that the southwestern Cape environment has followed a distinctive course of evolution (Meadows and Baxter 1999). It has been proposed that the CFR represented a region of relative stability during the rapid and dramatic environmental shifts occurring elsewhere in southern Africa in the Pleistocene (Barraclough 2006; Dynesius and Jansson 2000; Meadows and Baxter 1999). The topographical complexity of the region may also have provided organisms the opportunity to find refuge from glacial conditions by simply shifting their altitudinal range (Linder and Vlok 1991; Midgley et al. 2003). Today chacma baboon troops are known throughout this region.

The habitats of southern Africa have been shifted, fragmented, reduced and expanded since at least the beginning of the Pleistocene. These changes in habitat have produced significant genetic structure in many African lineages as populations have been forced to respond to these changes. A summary of what is known about climate and habitat change in the southern African Pleistocene and its effect on mammal diversification is presented below.

University of Cape Town

Climate and habitat change in the southern African Pleistocene

The morphological and ecological diversity seen within chacma baboons is likely to have emerged in response to the same climate and habitat fluctuations that have affected other southern African mammals during the Pleistocene. The primary goal of this thesis is to understand the responses of ancestral baboon populations to these environmental forces and the role of these phenomena on shaping diversity and structure within the species.

Fluctuations in temperature had an enormous impact on glaciations in the temperate latitudes, but for largely semiarid southern Africa, where conditions were not conducive to glaciations, the major impact was from changes in precipitation (Jürgens 1997; Meadows 2001). It is generally assumed that the decrease in temperatures in glacial periods led to increased aridity as an expanding cryosphere held more fresh water out of the hydrological cycle. This resulted in average lower rainfall leading to reduced forests and more-widespread grasslands. There are indications, however, that this was not a uniform pattern across southern Africa. Over the last 2 myr the Earth has experienced roughly twenty glacials and interglacials each lasting approximately 100 000 years. This suggests roughly 20 cycles of aridification and increased precipitation for southern Africa.

A period of major cooling intensification, some time between 0.9 Ma and 0.7 Ma, led to a second glacial maximum with the dominant periodicity of glacial-interglacial cycles shifted from 41ka to 100 ka (e.g Flagstad et al. 2001; Hooghiemstra et al. 1993). The final glacial period that proceeded the current warm epoch is termed the Last Glacial Maximum (LGM). The LGM was a time of maximum global ice volumes between 24kya and 10kya. The exact timing of expression of the LGM varies locally and in southern Africa; it can generally be defined as the period between 18 ka and 11.5 ka. African evidence ranging from the East African low latitudes (Hamilton and Taylor 1991) all the way to the South African Cape (Vogel 1983), shows that LGM was substantially colder than today by 5-6°C (Chase and Meadows 2007; Heaton et al. 1983 ; Kulongoski and Hilton 2004; Stute and Talma 1997) In the mountains of the Western Cape Province, in the modern WRZ, periglacial geological features such as ice wedge casts (Boelhouwers and Meiklejohn 2002) indicate winter temperatures for the region that were 8-10 °C colder than today (Chase and Meadows 2007).

The distribution of humidity in the LGM for southern Africa suggests decreasing rainfall from south-west to north-east across the subcontinent, the opposite of the modern pattern. This indicates a wetter LGM in the WRZ (e.g. Eland's Bay Cave and Cederberg) (Meadows and

University of Cape Town

Baxter 1999; Shaw and Thomas 1996). There is also support for a wetter Kalahari out of phase with other subtropical deserts for this time period (Meadows 2001). The LGM ended with renewed warming after about 13.5 -11.5 ka

Palaeoenvironmental change and species distributions in southern Africa

Habitat[3] theory as described by Vrba (1992) states that when a habitat changes or shifts, a species can either go locally extinct or must respond to ensure survival. There are two major survival responses: (i) a species can simply move with the habitat, or (ii) the species can stay where it is and adapt to the new habitat. Each response affects genetic structuring within the species differently. The first usually leads to significant geographic structuring within the species and possibly eventual speciation, this is largely due to population fragmentation as groups fission under stress and seek out suitable habitat. The second response tends to limit structuring (Vrba 1992).

Genetic analysis of lineage differentiation and estimates of divergence dates in bovid species such as impala (*Aepyceros melampus*), greater kudu (*Tragelaphus strepsiceros*), blue wildebeest (*Connochaetes taurinus*), topi (*Damaliscus lunatus*) and hartebeest (*Alcelaphus buselaphus*) clearly show a correlation between the distribution and timing of emergence of lineage diversifications and climatic fluctuations (Arctander et al. 1999; Flagstad et al. 2001; Nersting and Arctander 2001). This correlation suggests that diversification within these lineages was driven by the cycles of warming and cooling and the resultant changes in the distributions of rainforest and savannah habitats (deMenocal 1995; Flagstad et al. 2001; Hamilton and Taylor 1991; Muwanika et al. 2003; Van Hooft et al. 2002); this strongly supports the hypothesis that large-scale climatic fluctuations have been a major determinant for the evolutionary history of specialized mammal species in Africa.

The relative abundance of grazing mammalian species, which indicate more open woodlands and grasslands, fluctuates in East Africa until approximately 1.8 Ma (Reed 1997), when grazers begin to dominate the fossil record. Although South African hominid localities give evidence of having been more dry and bushy than contemporaneous eastern African sites, percentages of grazers follow the same pattern (Reed 1997). This has been interpreted within a palaeoenvironmental framework as a time of increased grass cover and less bush and woodland. This supports Cerling's (1992) hypothesis of an expansion of modern type C4 (secondary) grasslands in Africa, and deMenocal's (1995) evidence of drier

[3] Here I use the term "habitat" to mean a geographic region that contains all the resources necessary for particular species to survive and reproduce.

University of Cape Town

environments and seasonality at this time. More recent work in southern Africa has confirmed that C_4 grasses were only a minor part of the environment until the late Pliocene and early Pleistocene (Segalen et al. 2007) and that there has been a general trend towards more open environments since 3.0 Ma and a marked change to open grassy habitats at 1.8 Ma (Lee-Thorp et al. 2007).

This expansion of grasslands also drove lineage diversification in less specialised species. It has been shown that the bushbuck (genus *Tragelaphus*) which today is Africa's most widely distributed ungulate, split into two major lineages – *T. sylvaticus* and *T. scriptus* -- circa 2.5 Ma. (Moodley and Bruford 2007). This pattern could result if *Tragelaphus* was split north and south of the hypothetical 'arid corridor' which stretched from the Horn of Africa to the Cape of Good Hope during drier periods. This corridor would have provided a route by which xeric flora and fauna dispersed from central to southern Africa (Knoch and Schulze 1956; van Zinderen Bakker 1969, 1978), but would have fragmented non-xeric species to the northwest and southeast. Climate fluctuations and habitat change have clearly played a significant role in shaping diversity for many mammal species within southern Africa. Recent work on the phylogeny of *Papio* (Zinner et al. 2009) separates the genus into northern (maned) and southern (unmaned) lineages, a pattern of diversification that closely parallels that of *Tragelaphus* both temporally and spatially.

Climate driven cycles of temperature and humidity are also tied to population fragmentations south of the hypothetical "arid corridor" (Jürgens 1997; Matthee and Flemming 2002; Matthee and Robinson 1997; Paulo et al. 2001). A general pattern of lineage diversification has been identified for several southern African vertebrate lineages. While the details may differ according to the habitat needs and dispersal ability of each species, there is a general pattern of divergence into northern and southern lineages. This is true for the ground squirrel, *Xerus inauris*, (Paulo et al. 2001), the rock lizard, *Agama atra*, (Matthee and Flemming 2002) and the rock hyrax, *Procavia capensis* (Prinsloo and Robinson 1992) as well as for the Smiths red rock rabbit, *Pronolagus rupestris* (Matthee and Robinson 1996). These studies all suggest that the patterns of genetic structure observed in modern taxa are the result of habitat fragmentation and subsequent population isolation in glacial refugia.

Recent work on the Cape rock elephant shrew in South Africa recovers a past fragmentation event occurring approximately 1.7 Ma which again separates a northern lineage from a southern one (Smit 2007). The date of this event and the distribution of the northern and southern shrew lineages bears a striking resemblance to the emergence and placement of a southern clade within *Tragelaphus (T.sylvanus)* at ~ 1.7 Ma. For both the shrew and

University of Cape Town

bushbuck lineages this diversification event has been linked to one of three major peaks of aridification in Africa in the Plio-Pleistocene epoch (deMenocal 1995) and, for *Tragelaphus*, with the expansion of grasslands during the glacial period of the early Pleistocene also around 1.7 Mya (Hewitt 2000; Lee-Thorp et al. 2007; Moodley and Bruford 2007).

Summary

The divergence of an independent chacma mitochondrial lineage baboons occured circa 1.8 Ma. Modern chacma baboons are widely distributed across a range of habitats in southern Africa, including extreme habitats in places such as Namibia and the Drakensburg mountains. They are, however, absent from the Kalahari Desert and the Kgalagadi Transfrontier Park. The ecological and habitat variability seen in chacma baboons is mirrored by significant morphological variability, suggesting similar levels of variation in the genome. Correlations between the distribution and timing of emergence of lineage diversifications and climatic fluctuations have been shown for a number of small and medium bodied mammals in southern Africa, and it is proposed that genetic diversity in chacma baboons has been similarly shaped by these factors. The following three chapters will examine this hypothesis.

University of Cape Town

Appendix 3A- Phenotypic variation in size and pelage across the distribution of chacma baboons as summarised by Hill (1970) and including habitat data from the literature.

Phenotype	General	ursinus[1]	orientalis[1]	occidentalis[1]	griseipes[2]
Reference	Kerr 1792	Kerr 1792	Goldblatt 1926	Goldblatt 1926	Pocock 1911
Common name	Chacma	Cape chacma	Eastern cape chacma		Rhodesian chacma
Type locality	Cape of Good Hope	Cape of Good Hope	Undesignated	Rustenburg	Potchefstroom
Distribution	Most of southern Africa.	Western parts of the western Cape with eastern limit at Graaf Reinet and northern limit at the Orange River.	East of Knysna along the coastal tract probably as far north as the lower Limpopo. Includes Mpumalanga, Swaziland, Kwa-Zulu Natal and parts of the Eastern Cape.	Between upper Vaal and upper Limpopo Rivers. Northern limit is Zoutpansberg. Is confined to the western Drakensberg and extends to the edge of the Kalahari.	North of Limpopo, northwards through Zimbabwe into southern Zambia. Meets in the west with *chobiensis*. Meets *jubilaeus* in the east.
Habitat	Seven biomes.	*Fynbos* biome, winter rainfall, succulent karoo and nama karoo.	Savanna and grasslands. Also coastal thicket and forest and the scarp forest.[+]	Savanna and grasslands and mountainous regions of the Waterberg, SA.[+]	Savannah.
Pelage and hair	Almost black, with no greenish element, generally coarse hair with obscure speckling.	Hair dark brown with black tips and a narrow, subterminal yellow band.	Shaggy coated with evident buff speckling due to broader yellow bands on hair than nominate.	Mixed buffy and greyish.	More yellow than *ursinus*, resembling *p. cynocephalus*. Dorsum is grey. Grizzled olive-yellow but darker along spine and crown. Hairs ringed with black and yellow with grey bases and yellow tips.
Ventral surface	Generally lighter than dorsal surface.	Dull greyish, slightly darker on upper chest with a dark band joining shoulders.	Black along spine and sides of nape. grizzling on chest but not on abdomen.		Pale greyish-white, chest and belly annulated.
Limbs	Robust and sturdy.	Distal parts darker than body.		Forearms black.	
Flank	Dark coloured.	Dark coloured.	Silvery.		Dark coloured.

University of Cape Town

Appendix 3A- Phenotypic variation in size and pelage across the distribution of chacma baboons as summarised by Hill (1970) and including habitat data from the literature.

Phenotype	General	ursinus	orientalis	occidentalis	griseipes
Head	Absence of heavy side whiskers.	Absence of heavy side whiskers.			
Hair	Black tufts around the head.	Black tufts around the head.			
Face	Sparsley covered with short whitish hairs. Darkly pigmented.	Sparsley covered with short whitish hairs. Darkly pigmented.			Chin and lower cheeks pale greyish-white.
Eyelids	Flesh coloured.	Flesh coloured.			
Ears	Dark coloured.	naked, brownish-black.			
Hands	Dark coloured.	Dark coloured.	Black	Black	Grey
Feet	Dark coloured.	Black	Black	Black, with brown sides and toes.	Grey
Nails	Black	Black			
Palms/soles	Dark coloured.	Brownish black .			
Size	Largest of the baboon taxa.	Smaller than eastern forms.			
Tail	Appears broken due to kink in tail at approximately 1/3 form the base.			Mostly black.	No dark tip.
Comparison			Generally brighter and more yellowish than *ursinus.*	Underparts, including throat darker than ngamiensis but cheeks not as pale. Lighter and more yellow than orientalis, narrower shorther muzzle than *ursinus*.	On average darker than *orientalis* and *occidentalis.*

University of Cape Town

Appendix 3A- Phenotypic variation in size and pelage across the distribution of chacma baboons as summarised by Hill (1970) and including habitat data from the literature.

Phenotype	ngamiensis[2]	chobiensis[2]	ruacana[3]	jubilaeus[4]
Reference	Roberts 1932	Roberts 1932	Shorthridge 1942	Schwarz 1928
Common name				Dwarf chacma
Type locality	Maun	Kabulabula, Chobe River, Bechuanaland .	Ujiwau, Kaokoveld	Misale, north-eastern Zimbabwe (14^{O} S, 33^{O} 10E)
Distribution	Bechuanaland, especially around lake Ngami. Meets *chobiensis* in the north.	On both banks of the upper Zambezi valley. Meets *griseipes* in the north and east along the Kafue River. Meets *ngamiensis* in the south-west.	Damaraland, Kaokaveld and southwest Angola (west of the Cunene River)	Zambia, eastern province plateau and Luangwa Valley and adjacent pats of Malawi and Mozambique.
Habitat	Aquatic grassland or wetland. #	Dry deciduous forests that is mopane dominated .#	Found across a range of savannah subtypes, including mopane, mountain, thornbush, highland, dwarf shrub and mixed tree. Also desert and semi-desert. *	Floodplains and savanna more open conditions in the drier south to tall dense woodlands in the north and north-west of the distribution. **
Pelage and hair	Deep buffy-grey with a little yellow suffusion, mane of long hairs dark grey-brown at the base with a broad bar of buffy and the rest black.	More yellow than *ngamiensis*, nearer to *griseipes* but with blackish annulation more prevalent above, long mane extending all the way along the back and makes a median crest at the nape.	More ochraceous than southern types but dark along the spinal tract, shoulder hairs basal half sepia, wide straw band and black tip.	Long and lax with crest on nape, moderately long on shoulders, upper parts tan speckled with black, paler on flanks.
Ventral surface	Throat greyish, chest yellowish-grey.			
Limbs	Shoulders buff coloured with longer hairs banded with dark grey and blackish extending to forearms which are more black.	Forearms annulated blackish.	Black forearms with speckling on the distal half increasing towards the hands, thighs pure straw below callosities.	Hair on arms and thighs long and thick.

University of Cape Town

Appendix 3A- Phenotypic variation in size and pelage across the distribution of chacma baboons as summarised by Hill (1970) and including habitat data from the literature.

Phenotype	ngamiensis	chobiensis	ruacana	jubilaeus
Flank	Yellowish on the flanks hind limbs especially thighs.		Lumbar region speckling olive-brown, lighter lateral to callosities.	
Head	Yellowish on the crown.		Hairs on crown Black with straw tips long. Black hair on nape.	Yellow-brown on the crown.
Hair				
Face	Chin greyish.		Short, fine, sparse silvery-white hairs, longer and darker (grey) on cheeks to black around the ears.	Whitish hair on cheeks extends toward the rostrum with only dorsal nose and snout naked. Light hair on chin and throat.
Eyelids			Black eyelashes.	
Ears				
Hands	Not uniformly black.	Uniformly grizzled olive-grey.	Black.	Brown.
Feet	Not uniformly black.	Uniformly grizzled olive-grey, lighter than hands.	Black, above very dark brown hair with some golden annulations .	

University of Cape Town

Appendix 3A- Phenotypic variation in size and pelage across the distribution of chacma baboons as summarised by Hill (1970) and including habitat data from the literature.

Phenotype	ngamiensis	chobiensis	ruacana	jubilaeus
Nails				
Palms/soles				
Size				
Tail	Becomes greyer at tip.	Equal black and yellow at base becoming uniformly brown-grey to the tip.	Base olive-brown speckling with a very dark brown distal half	Hair dense and no terminal tassle.
Comparison			More pale zones than *ursinus*.	Intermediate between *ursinus* and *P. cynocephalus*, long limbed with brown hands.

[1] Incorporated into sub-species *P. u. ursinus*

[2] Incorporated into sub-species *P. u. griseipes*

[3] Incorporated into sub-species *P. u. ruacanna*

[4] Now known to be a yellow baboon.

[+] Mucina and Rutherford 2006

*Wardell-Johnson, G. (2000). Biodiversity and Conservation on Namibia into the 21st Century. pp 17-45 in B. Fuller and I. Prommer, *Population-Development-Environment in Namibia: Background Readings.* Laxenburg, Austria: IIASA, IR-00-031.

[+] http://www.botswanatourism.us/experiencebotswana/wildlifehabitatsmap.html (Created 04 March 2009)

**http://www.zambiatourism.com/travel/hisgeopeop/geograph.htm

University of Cape Town

PREFACE TO CHAPTER 4

This chapter expands on a published, multi-authored paper*. The research presented in this manuscript was funded by grants held by Sithaldeen and Ackermann. Sithaldeen, Bishop and Ackermann were responsible for project design, conceptual contributions and manuscript preparation. Sithaldeen collected and prepared samples and generated and processed data. Sithaldeen and Bishop analysed the data. Sequences have been published in GenBank.

The December 2008 submission of Sithaldeen et al. (2009), preceded the publication of a phylogeny of the baboon genus (*Papio)* by Zinner et al. in April 2009. The published sequences from Zinner et al. are included in the analyses presented here. This chapter therefore expands on the already published results of Sithaldeen et al. (2009).

A subsequent paper by Keller et al., published in May, 2010, is related to this work but made use of a different molecular marker to Sithaldeen et al. (2009). The findings of Keller et al. (2010) were therefore not factored into the original design of this project, nor the 2009 publication, and the results are referenced here only as they relate to the conclusions of this chapter.

University of Cape Town

***Sithaldeen, R.**; Bishop, J. M. and Ackermann, R. R. (2009). Mitochondrial DNA analysis reveals Plio-Pleistocene diversification within the chacma baboon. *Molecular Phylogenetics and Evolution*. 53 (3): 1042-1048.

CHAPTER 4

PHYLOGENY OF THE CHACMA BABOON (PAPIO URSINUS)

Abstract

The results of recent molecular studies support a hypothesis of climate linked diversification for the genus Papio*, however no attempt has been made to test this idea robustly. Here I use phylogenetic techniques to analyse evolutionary relationships within a high resolution sample of* P. ursinus*. These results are then used to correlate lineage diversification within this species to specific climate-driven landscape dynamics and in so doing, test the hypothesis that climate played a significant role in the diversification of* P. ursinus *in southern Africa. Individuals were sampled from geographic localities in South Africa, Namibia, Botswana and Zambia. An 896bp fragment of the mitochondrial Brown region was sequenced and 35 unique haplotypes were recovered from the sample, these were used in the final alignment which also included published sequences for* P. ursinus *(n=15) and additional* Papio *taxa (n=5). An intraspecific phylogeny was constructed using Maximum Parsimony and Bayesian methods and trees were rooted using* Theropithecus gelada *(n=1) and* Macaca *(n=2) sequences. Node ages were estimated for several clades using a number of fossil dates to constrain the time to most recent common ancestor (TMRCA) on the* Papio*-gelada split and the TMRCA for all* Papio*. Divergence estimates reveal that the chacma mitochondrial lineage diverged from ancestral* Papio *at ~1.80 Ma. Tree topologies were virtually identical for both methods of tree construction and reveal a major diversification event within chacma at ~1.60 Ma, splitting the emergent lineage into northern and southern populations. This diversification event is linked to climate driven aridification of central southern Africa and the subsequent expansion of the Kalahari Desert. Two distinct mitochondrial lineages emerge fully ~30ky later at ~1.35 Ma and ~1.27 Ma, respectively. The analysis of haplotye clustering suggests that there are three genetic clades within the two major lineages. There is also support for a geographic cluster of Namibian baboons nested within the southern lineage, emerging at ~1.00 Ma, coincident with climatically driven fluctuations in the level of the Orange River. Nested within the northern lineage is a well-supported north-eastern clade, which diverged at ~0.42 Ma, and a possible but poorly supported Zambian clade emerging at ~0.29 Ma. It is proposed that the northeastern baboons dispersed and diverged during an unusually long interglacial period. There is also some correlation between genetic and morphological groupings within chacma. Together these results support a significant role for climatically driven diversification within chacma baboons.*

University of Cape Town

Introduction

It has been suggested that diversification within *Papio* is linked to climatically driven habitat change (Jolly 2001). This model is not uniquely proposed for baboons, as large-scale climatic fluctuations have been shown to be a major determinant in the evolutionary history and population genetic structure of many medium to large bodied mammal species (Hewitt 2000, 2004; Flagstad et al. 2001; Muwanika 2003). Although molecular phylogenies of *Papio* from Newman et al. (2004) and Zinner et al. (2009) broadly support Jolly's (2001) hypothesis, the resolution of landscape sampling in these studies, is inadequate for a detailed analysis of the correlation between landscape changes and diversification events within the genus and especially within individual species. Here I test Jolly's (2001) hypothesis by assessing the strength of association between the timing and geography of landscape changes and genetic and morphological diversification within chacma baboons.

Fossil evidence reveals the first recognizable *Papio* in sub-Saharan Africa at least 2.5 Ma (El-Zaatari et al. 2005; Frost and Delson 2002; Jolly 2001) in response to the expansion of savannah habitats which occurred as a result of global aridification (de Menecol 1995; Lee Thorp et al. 2007; Vrba 1992). The modern genus is divided into two categories; the northern "maned" baboons, *P. papio, P. hamadryas, P. anubis* and the southern, "unmaned" baboons *P. cynocephalus* and *P. ursinus* (Jolly 1965, 2001). Each of the five species is defined by a suite of fixed morphological and behavioural traits that are shared between individuals. The divergence of these phenotypes is attributed to periods of genetic isolation between baboon populations during which time these diversified into separate species (Jolly 2001). Significant variation has been documented across the range of each of these species (Groves 2001; Hill 1970; Jolly 1993) and it is likely that within species diversification is also linked to landscape and habitat changes. Between three and eight sub-species of chacma baboon have been identified. The three generally accepted are *P.ursinus ursinus, P. u. ruacana* and *P. u. griseipes* (Groves 2001; Jolly 1993). A further five possible subspecies, *P. u. orientalis, P. u. occidentalis, P. u. ngamiensis, P. u. chobiensis* and *P. u. jubilaeus* are described (Hill 1970). The distribution of each phenotype is given in Hill (1970) and Appendix 3A and represented graphically in Fig. 3.3.

The evolutionary diversification of chacma baboons began in southern Africa ~2.09 Ma, since the lineage diverged from the ancestral form (Zinner et al. 2009). Proto chacma groups would most likely have emerged somewhere in northern South Africa (Zinner et al. 2009) and then

University of Cape Town

expanded south with the grasslands (Lee-Thorp et al. 2007), eventually giving rise to an independent baboon species at ~1.8 Ma. The mitochondria carried by chacma baboons have therefore been evolving independently for almost as long as modern *Papio* has been in existence. Given the evidence for climate and landscape driven structuring in many lineages in southern Africa it is expected that similar patterns of diversification will be indentified in baboons. Here I use chacma baboons as a case study to test the correlation between climate and landscape change in southern Africa, and genetic and morphological diversification within a baboon species. Phylogenies derived from molecular data provide an indirect record of diversification events (nodes) which can be dated. Together with geographical and environmental data, they can be used to test the role of geographic factors in the diversification of a taxa (Barraclough 1998; Berlocher 1998; Brooks and McLennan 1991).

The accuracy of phylogenetic reconstruction is greatly improved by high resolution taxon sampling. This is clearly illustrated by a comparison of three studies that have already attempted to reconstruct the evolutionary history of baboons. Using a small data set of 40 mitochondrial Brown region sequences Newman et al. (2004) produced a basic molecular phylogeny of *Papio*. For each of the five species sampled, individuals were all sourced from a single geographic point within the distribution of each species. The guinea baboon and chacma baboon samples are geographically isolated from the other three species and each other, and the olive, hamadryas and yellow baboons all come from the same geographic region in East Africa. Consequently, genetic variability within and between *Papio* species is largely under-represented and, as a result the analysis recovers three distinct, morphologically concordant, mitochondrial lineages within *Papio*, while the fourth grouping reveals an unresolved olive / yellow clade . The discovery of an unresolved clade suggested that the rest of the resulting tree construction may be an oversimplified representation of the evolutionary relationships between these species. Indeed this was confirmed in Zinner et al. (2009), who produced a phylogeny for *Papio* based on 67 mitochondrial Brown region sequences representing baboons from 53 localities across the distribution of the genus. This study again recovered four major clades, nonetheless, instead of distinct, morphologically concordant, mitochondrial lineages, the results reveal that three of the four taxonomic groups are paraphyletic. Such improved representation of variation within *Papio* indicates a significantly more complex evolutionary history for *Papio* than was revealed by Newman et al. (2004). As a result of improved taxon sampling, Zinner et al. (2009) were able to reconstruct a detailed picture of the events leading to the emergence of distinct *Papio* species and proposed a scenario of repeated introgression to explain the observed mitochondrial

University of Cape Town

paraphyly between them. However, a detailed analysis of the correlation between the timing of diversification events and landscape changes or of the processes driving diversification was not included in this study, nor did it focus on intraspecific diversification in great detail.

In contrast to the two studies discussed above (Newman et al. 2004; Zinner et al. 2009), Winney et al. (2004) focussed on variation within a single baboon taxon, producing an intraspecific phylogeny of hamadryas baboons (*P. hamadryas*) based on 26 haplotypes from 107 individuals. In this study, the authors were able to draw very specific conclusions about the role of landscape in driving differentiation of African and Arabian hamadryas baboons. In this case, the main divergence between the two taxa was driven by the periodic formation of landbridges across the Bab El Mandeeb in the southern Red Sea during the Last Glacial Maximum, a pattern observed independently by Wildman et al (2004). The level of detail achieved in this reconstruction of baboon population history is an example of what can be achieved through high resolution taxon sampling.

Phylogenetic studies often fail to adequately sample within taxon variation (e.g. Braun and Kimball 2002; Bremer et al. 1999; Chen et al. 2003; Freudenstein et al. 2003; Johnson 2001; Lin et al. 2002; Sorenson et al. 2003). This can lead to over simplistic phylogenies (Rydin 2002) such as that of Newman et al. (2004). Undersampling can also affect the estimation of node ages within a tree. This is particularly true for analyses that use more extreme rate smoothing (Linder et al. 2005). Inaccuracies in age estimation increase with the distance from calibration nodes such that the combined effect of undersampling and distance from the calibration node can result in up to a threefold difference in the age estimation of nodes from the same dataset with the same calibration point (Linder et al. 2005). Three additional factors that affect the accuracy of phylogenies are marker selection, sequence length, and choice of method for analysis (Rokas 2005; Swofford et al.1996). The effect of each factor depends on the nature of the study. In order to produce a robust phylogeny of chacma baboons and thereby generate a detailed picture of the evolutionary history the taxon, every attempt is made to counter these factors.

University of Cape Town

Materials and methods

Sampling

The DNA of a target species can be successfully extracted from faeces. This is possible as epithelial cells from the lining of the intestine are exfoliated and passed out with the pellet (Piggot and Taylor 2003). Faecal sampling is therefore a non invasive and plentiful source of DNA, however there are certain obstacles to downstream success inherent to this sample type (Deuter et al.1995). Firstly, most of the DNA in a faecal sample is actually microbial in origin. Faecal pellets therefore only have a relatively small amount of the target DNA (Piggot and Taylor 2004). It is imperative then, when using faecal DNA, that species specific PCR primers are used to amplify gene regions of interest. Faecal pellets are also subject to bacterial and enzymatic activity that may destroy DNA. It is therefore good practice to collect fresh samples and to stop all enzymatic and biological activity at the point of collection (Deuter et al.1995). In this way DNA degradation due to UV radiation is also minimized. Another possible challenge to amplification is PCR inhibition by plant tannins. Finally, when using faecal pellets as a source for DNA, the probability of sample duplication is high, and measures must be taken to minimize this.

To maximize the harvest of epithelial cells in each sample, pellets were collected soon after daybreak, which is when baboons generally have their first bowel movement of the day and presumably will exfoliate the most cells (Banks et al. 2002). In cases where sleeping sites were previously unknown, this was achieved by following a troop to their sleeping site in the late afternoon and returning before daybreak the next day to collect fresh samples. This approach also limits the deterioration of DNA by exposure to UV radiation. Samples were collected from wild, non-habituated troops, as a result individual sample identification was not possible. In order to minimize the duplication of samples, every effort was made to collect pellets of different sizes and consistency. Faecal samples from 261 free-living baboons were collected. Approximately 5g of wet faeces was collected per sample and placed in absolute ethanol to saturation. This stops all enzymatic and biological activity at the point of collection. In the field samples were kept at a constant ambient temperature in a thermally insulated container. Samples were then later stored at -20^{o}C in the laboratory. The removal of PCR inhibitors was achieved during the extraction protocol as outlined in the QIAamp DNA Stool Kit (Qiagen©) handbook.

Ideally samples would have been collected at equally spaced localities across the distribution of chacma baboons so as to remove any artefacts of sample choice. However due to a number of

University of Cape Town

constraints, sampling was limited to known localities of baboon troops and constrained by the accessibility of troops. The sampling strategy was therefore refined and aimed at collecting an ecologically representative sample. To capture a greater signature of geographic variability every effort was made to sample adult males in each troop. This strategy was employed based on the fact that baboons are male dispersing and the Brown region is a mitochondrial marker, factors which both predict that most of the females in the troop share the same mitochondrial haplotype. Large male baboons were observed and followed preferentially as the troop descended from its sleeping site in the mornings. Samples were collected within the ranges of each of the eight chacma subspecies (Hill 1970) and the geographic source locations represent a wide spectrum of ecological variability within the region. The distribution of samples is shown in Fig 4.1.

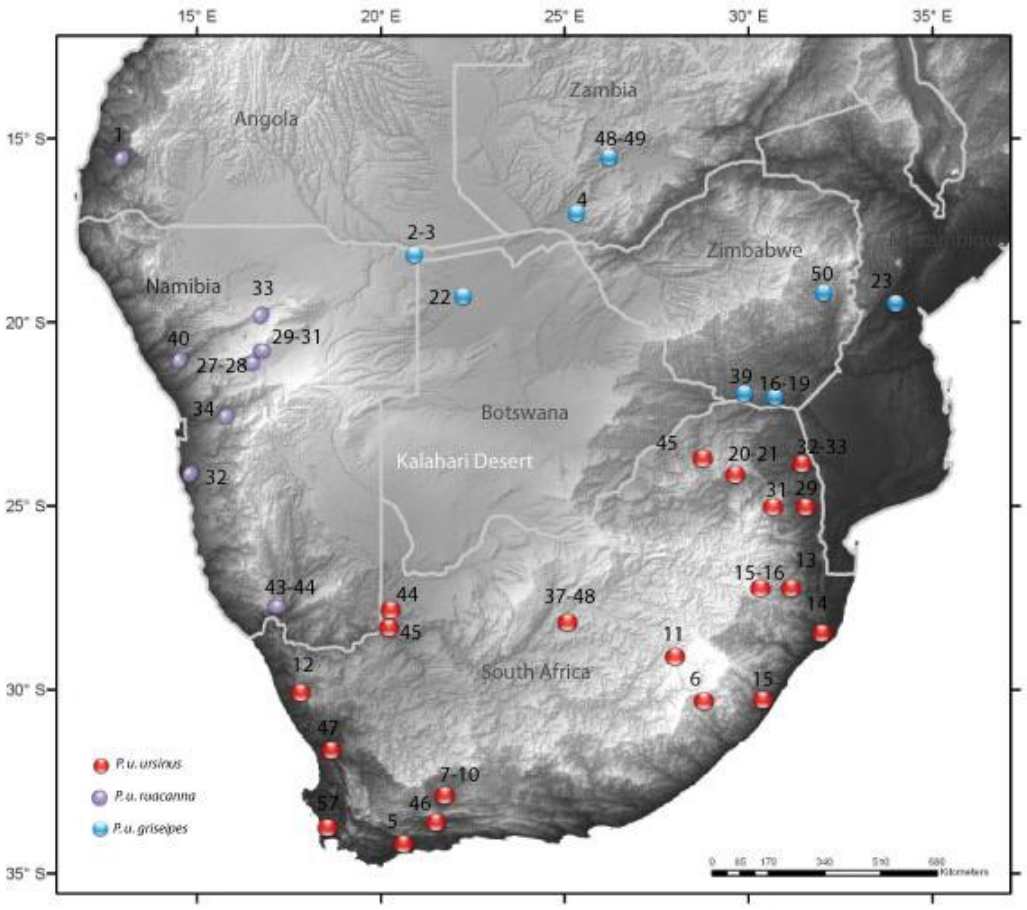

Figure 4.1- Map showing the approximate distributions* of each of the three chacma morphotypes as defined by Groves (2000) and Jolly (1993) and sampling localities of the* P. ursinus *sequences used in this study. See Map IDs in Appendix 4A for sample details.

Choosing a marker for intraspecific phylogenetic reconstruction

The structure and gene organization of mtDNA is generally highly conserved among mammals. The mitochondrial genome is a closed circular molecule, and codes for messenger RNA's, two ribosomal RNA's (rRNA) and twenty-two transfer RNA's (tRNA). The genome is approximately 16600 base pairs (bp) in length and contains no intervening sequences within transcribed genes, no spacer sequences between genes and no class of repetitive DNA (Giles 1980). In recent years mitochondrial DNA markers have been used extensively in phylogenetic and phylogeographic studies (e.g. Ascunce et al. 2007; Avise 2009; Flagstad et al. 2001; Herron et al. 2005; Newman et al. 2004; Zinner et al. 2009) and its popularity is based on five major factors.

The major factors contributing to the popularity of mitochondrial makers for phylogenetic reconstruction are outlined below along with certain caveats and corrections that have come to light in recent years and which must be considered in any analysis.

i. MtDNA exists as multiple copies in each cell

MtDNA is located in the mitochondrial matrix and is distinct from the nuclear genome (Taanman 1999). A typical cell has several hundred mitochondria and each organelle generally contains several identical copies of mtDNA (Michaels et al. 1982). The availability of multiple copies of the mitochondrial genome compared to a single copy of the nuclear genome in each cell, greatly improves the probability of a successful DNA isolation. This is particularly useful when using poor quality DNA sources such as faeces.

ii. MtDNA evolves faster than nuclear DNA (nDNA)

In order to capture significant genetic structure in a phylogeny of closely related organisms one requires a molecular marker that segregates rapidly between generations. The mitochondrial genome is constantly exposed to reactive oxygen metabolites (Avise 1991). In addition, and when compared to nDNA, there is a relative lack of mechanisms to repair the genome. This leads to a relatively high rate of mutations within the genome such that even recently diverged populations can be genetically differentiated from each other. Unfortunately because of the nature of these mutations, together with demographic change within species, the accumulation of mutations can be highly variable in space and time (Galtier 2009) and therefore difficult to predict and model mathematically. While this may decrease the reliability of branch length estimations, mtDNA polymorphisms are still informative as markers of intraspecific variation

University of Cape Town

(Ballard and Dean 2001) and shared substitutions along the length of the marker can be used to construct the evolutionary relationships between individual haplotypes.

iii. MtDNA is passed on through a simple pattern of known inheritance

Paternal mtDNA is generally eliminated before (Moses 1961), during (Ursprung and Schabtach 1965) or after (Sutovsky et al. 1999) fertilization (Xu et al. 2005). Though there have been reported cases of paternal mtDNA leakage (Kaneda et al. 1995; White et al. 2008) most studies continue to assume strict maternal inheritance of the mitochondrial genome (Hiraoka and Hirao, 1988; Taanman 1999; Ballard and Dean 2001) and no recombination between parental genomes. Furthermore the transmission of mtDNA is characterized by a strong bottleneck i.e. through a single egg cell (Shoubridge and Wai 2007). Each individual therefore has only one mitochondrial ancestor at each ancestral generation. This reduces within individual diversity and makes for a simple pattern of known inheritance (Harihara et al.1986; Hayasaka et al. 1986; Zhang and Shi 1993). MtDNA polymorphisms can therefore be used to create groups of related mtDNA haplotypes or haplogroups. Within a species these often exist in non-random spatial or temporal distribution (Avise 1991). The patterns within these distributions can then be used to identify factors driving diversification within a lineage.

The assumption of no recombination within the mitochondrial genome has been challenged. The original studies to suggest a problem with this assumption, showed that there was an unexpectedly high frequency of homoplasy (Eyre-Walker et al.1999) and a negative correlation between linkage disequilibrium and physical distance for pairs of sites within a mitochondrial marker (Awadalla et al.1999). This was evidence that mtDNA does in fact, recombine. Although the results of these original studies have been discredited, later work confirmed that mitochondria can in fact recombine and that significant departure from clonality does occur in several species, including primates (Piganeau et al. 2004; Tsaousis et al. 2005). Although the prevalence of mitochondrial recombination across animals has been difficult to assess, it does occur (Andolfatto et al. 2003; Ciborowski et al. 2007; Gantenbein et al. 2005; Hoarau et al. 2002; Ladoukakis and Zouros 2001; Ujvari et al. 2007) and this must be taken into account when building and interpreting intra-specific mtDNA genealogies (Hey 2000).

University of Cape Town

iv. The mitochondrial genome generally evolves under neutrality

Mitochondrial DNA contains many functionally important genes especially those related to respiration. This fundamental life process is extremely sensitive to changes in the genetic sequence, so it is assumed that deleterious mutations are quickly removed by purifying selection while adaptive mutations, which spread through positive selection, are very rare. The result is that that variation within mtDNA is assumed to be neutral and that this variation can be modelled according to neutrality theory. The geographic distributions of haplogroups can therefore be used to understand the external processes driving differentiation. A meta-analysis of >1600 animal species (Bazin et al. 2006) suggest that this assumption may not always hold. Bazi et al. (2006) found compelling evidence for recurrent selective sweeps within the mitochondrial genomes of many lineages and positive adaptive selection within mtDNA has also been identified within several groups of animals (Castoe et al. 2008, 2009) including primates (Grossman et al. 2004). MtDNA cannot therefore, be assumed to evolve neutrally and the assumption of neutrality should be explicitly justified in any study.

v. Mitochondria evolve at a known and constant rate

It is assumed that the evolutionary rate of the mitochondrial genome varies only 2-3 fold across highly divergent taxa (Martin 1995; Gissi et al. 2000; Bininda- Emonds 2007). This evolutionary rate is approximately 2% substitutions per million years for mammals (Avise 1986). However it has recently been shown that these studies (i.e. Martin 1995; Gissi et al. 2000; Bininda- Emonds 2007) are highly flawed due to undersampling of taxa and because they only compare highly divergent sequences (Nabholz et al. 2008, 2009). In a meta analysis of >1500 cytochrome b sequences, Nabholz et al. (2008, 2009) found that mtDNA substitution rate actually varies between 30-100 fold within phyla. Xu et al. (2006) also reported significant variation in substitution rate across species. The use of a constant molecular clock for calculating divergence estimates is therefore unjustified as far as mtDNA is concerned. Fortunately, reliable statistical methods have been developed to deal with these deviations from clock like evolution (Hasegawa et al. 2003; Huelsenbeck et al. 2000; Sanderson 1997, 2002; Thorne et al. 1998; Thorne and Kishino 2002).

University of Cape Town

Molecular methods

Due to deterioration either because of ethanol leakage or other storage issues, 224 samples collected were considered to be usable. Aliquots of between 0.5 and 1.0g of faecal material were separated and used for DNA extraction from 224 samples. Ethanol was removed by centrifugation followed by drying under a fume-hood. DNA extraction was carried out using a QIAamp DNA Stool Kit (Qiagen) following the manufacturer's protocol. DNA yield and purity were checked using a NanoDrop 2000 (Thermo Scientific) micro-volume spectrophotometer and DNA integrity was assessed on 1.0% agarose gels. All gels were stained with ethidium bromide and visualized under UV light.

The Brown region is bounded by two HindIII restriction sites and spans portions of the ND4 and ND5 genes as well as three tRNAs (Brown et al. 1982). In most baboons the Brown region comprises 453bp of the 3' end of NADH dehydrogenase subunit IV (ND4) gene, the tRNA genes for Histidien (His), Serine (Ser) and Leucine (Leu) and 239 bp of the 5' end of the NADH dehydrogenase subunit V (ND 5) gene (Wildman et al. 2004). The mitochondrial ND4/ND5 gene region has been shown to outperform other commonly utilized mtDNA genes such as CO1, cyt b, and 12S/16S rRNA genes in phylogenetic analyses at broad taxonomic scales because it is relatively long (ca. 3.4 kb) and contains more phylogenetically informative variation at first and second codon positions (Miya et al. 2006). The Brown region itself has been used effectively in several phylogenetic studies (Doosey et al. 2010; Newman et al. 2004; Zinner et al. 2009).

The presence of nuclear copies of mitochondrial genes is a major problem in phylogenetic studies which employ mitochondrial markers (Bensasson et al. 2001; Zhang and Hewitt 1996). These mtDNA-derived nuclear pseudogenes, also called ''*numts*" (numites), are generated when part of the mitochondrial genome is transferred to the nuclear genome during evolution. This is an ongoing process in most eukaryotes. *Numts* were first found by Du Buy and Riley (1967). When hybridizing purified mtDNA with nuclear DNA of mice they found many strong comibinations which they deduced were the result of mtDNA homologous sequences in the nuclear genome. *Numts* have since been reported in over 100 eukaryotic species (Bensasson et al. 2001) and are known to occur in primates (Collura and Stewart 1995; Collura et al. 1996; Mundy et al. 2000; Olson and Yoder 2002; van der Kuyl et al. 1995). Today it is clear that nearly all mitochondrial regions can be integrated into the nuclear genome, that the sizes of the *numts* vary in different species and genes (Davis and Parker 1998; DeWoody et al. 1999; Lopez et al. 1994; Pons and Vogler 2005; Richly and Leister 2004a, 2004b; Zischler, et al. 1998;) and that

University of Cape Town

numts can be derived from the entire mammalian mitochondrial genome.

When sequencing mtDNA the probability of inadvertently co-amplifying or even preferentially amplifying nuclear copies or *numts* is high (Lopez et al.1994). *Numts* are a potential source of phylogenetic error (Rauum et al. 2005) and have been indentified in higher primates such as gorillas and orangutans (Anthony et al. 2007; Chung and Steiper 2008). Symptoms of *numt* contamination include PCR ghost bands, extra bands in restriction profiles, sequence ambiguities, frameshift mutations, stop codons and unexpected phylogenetic placements. *Numt* contamination can be contained either by bench methods or by post amplification, quality control. In the lab, amplification of *numts* can be minimised using one of four methods; (a) purification and isolation of mtDNA (Anderson et al. 1981), (b) amplification of the target region from RNA derived, complementary DNA template (Collura et al. 1996; Jukes and Osawa 1993; Pamilo et al. 2007; Williams and Knowlton 2001), (c) semi-nested/nested PCR amplification on long-range PCR product or (d) PCR amplification on diluted DNA extract. Of these, PCR dilution is the cheapest and simplest method for reducing *numts* and it is also very robust (Song et al. 2008).

In this study DNA quality and concentration was often highly variable across samples. To overcome this challenge baboon specific primers weere developed by Newman et al. (2004) were used; the use of species specific primers has been suggested as being beneficial in eliminating *numts* (Song et al. 2008). The variability in DNA quality made it necessary to amplify DNA under a range of annealing temperatures between 52° C and 56° C. PCR primers amplified two overlapping fragments and conditions followed Newman et al. (2004) with minor modifications to primer B896HR (B896-new 5' GATAGACCAGGTAATGAATAGTGC 3'). All PCR reactions were performed in a total volume of 25ul containing 25-50ng of genomic DNA, 1x reaction buffer, 2.5mM $MgCl_2$, 0.2mM dNTPs, 0.5pmol/ul primers, and 1 unit *Taq* polymerase (GoTaq, Promega). PCRs were performed on a GeneAmp 9700 thermocycler (Applied Biosystems). PCR products (3ul) were visualised on a 1.0% agarose gel. Samples that failed to amplify were reprocessed a maximum of 3 times varying combinations of annealing temperature and $MgCl_2$ concentration. Of 224 individual samples, 162 amplified successfully.

Successful PCR products were electrophoresed in a 1.5% agarose gel and target bands of 500-600bp in length were excised. No ghost bands were identified but this step provides additional protection against *numt* contamination. These bands were then purified using the Promega

University of Cape Town

Wizard SV gel and PCR clean up system. Post clean-up PCR products were visualised on a 1.0% agarose gel and samples with a high likelihood of downstream success were identified. Of 162 samples that were successfully amplified 112 produced sufficient post clean up product for cycle-sequencing. Samples were further amplified following a standard cycle sequencing protocol in a 10-20ul reaction depending on the post clean up yield.

PCR reactions were sequenced in both directions using BigDye v1.1 chemistry on an ABI 3730 (Applied Biosystems) capillary sequencer. Sequences were edited in BIOEDIT v5.0.9 (Hall, 1999) and aligned using CLUSTAL X (Thompson et al. 1997). Of 112 samples sequenced, 104 produced sequence data and 51 were high quality sequences i.e. sequences with low frequency of ambiguous/missing sites. These were used in the phylogenetic reconstruction. Once all sequencing ambiguities were removed from the dataset a final alignment of 879bp mitochondrial Brown region (Brown 1982) was created. Sample details, collection site data and GenBank accession numbers are reported in Appendix 4A and the distribution of samples are shown in Figure 4.1.

Post amplification quality control methods were used to check for *numts* in the data set (Buhay 2009; Song et al. 2008). All sequences in the final alignment of 51 sequences were compared to a published *P. ursinus* (AY212105) sequence isolated from purified mitochondrial DNA. The target sequences aligned with no deletions or insertions, or stop codons. Sequences were also checked for frameshift mutations against Arnason's full *Papio* mtDNA genome (NC001992, Arnason et al. 1998) reference sequence (Bensasson et al. 2001; Buhay 2009; Song et al. 2008; Sorenson and Quinn 1998). A high rate of variability in the first and second codon positions can also indicate *numt* contamination. (Sorenson and Quinn 1998; Bensasson et al. 2001; Song et al. 2008; Buhay 2009) but this was not observed.

Phylogenetic tree construction

i. Choosing a method of tree construction

Phylogenetic trees represent the genealogical relationships of individuals or groups of organisms. Methods of phylogenetic tree construction fall into one of four primary categories. These are distance methods such as Neighbor-Joining (Saitou and Nei 1987), character based methods such a Maximum Parsimony (Fitch 1971; Hartigan 1973), Maximum Likelihood

University of Cape Town

(Felsenstein 1981) and Bayesian methods (Li et al. 2000; Mau and Newton 1997; Rannala and Yang 1996).

Distance methods start by calculating the number of base pair differences between two aligned sequences. This number is used as an estimate of genetic distance between those sequences. A matrix of pairwise distances is constructed for each combination of sequences and a clustering algorithm is then applied to generate a phylogenetic tree (Everit et al. 2001). If nucleotide substitution rate is assumed to be constant, then the genetic distance can be used as a proxy for time since the sequences diverged (Yang 2006). Neighbour joining (NJ) is such an algorithm (Saitou and Nei 1987). In this method the first node joins the taxa with the least evolutionary distance between them. The distance of each of the remaining taxa is calculated to this first node and the next closest individual is added. This process continues until the tree is complete. The NJ score is based on minimum evolution criterion which estimates the tree length (Avise 2004; Hall 2004; Nei and Kumar 2000). This is a fast method that permits the analysis of lineages with largely different branch lengths and allows a correction for multiple substitutions. However, sequence information is greatly reduced and the outcome is strongly dependent on the model of evolution that is chosen. Also only one possible tree is generated. While NJ may still be used to analyze large data sets due to its computational efficiency, other more accurate methods exist.

Maximum parsimony (MP) tree construction is based on the assumption that the most accurate tree is the one that represents the fewest number of evolutionary changes (Eck and Dayhoff 1966; Fitch 1971). Once sequences are aligned, each nucleotide base is treated as a character. Invariant characters are removed from the analysis as uninformative. This method presupposes that shared characters are more likely to be inherited from a common ancestor than as a result of homoplasy (convergence and parallelism). MP first constructs a tree for each informative site and then chooses the trees that minimize the number of evolutionary steps needed to explain the data (Nei and Kumar 2000). The tree score is calculated as the number of changes, minimized over ancestral states (Yang 2006). There are several variations on the simple parsimony algorithm. Weighted MP (Fitch 1971, Hartigan 1973) methods allow the user to give a greater weight to “more likely” types of character changes e.g. transitions may be considered more likely than transversions. Although the MP algorithm does not reduce sequence information to the extent of NJ it is computationally slower than NJ, but it is more accurate

University of Cape Town

because it involves reconstruction of ancestral sequences and evaluates many different trees. MP is an unreliable method for constructing phylogenies of highly divergent lineages which may contain multiple back mutations (Hall 2004). However MP is a quick and simple way to check your data and generate a basic tree topology which can then be compared to more sophisticated methods.

Maximum likelihood (ML) tries to find the evolutionary tree that maximizes the probability of observing the data, given an underlying model (Yang 2006).There are three main components of an ML analysis, the data, the model of evolution and the maximum likelihood criterion. The probability of observing the data under the assumed model will change depending on the parameter values of the model. Trees are given a log likelihood score which is optimized over branch lengths and model parameters. As ML methods use all the sequence information they are considered be less influenced by sampling error than MP and unlike NJ tend to be robust to many violations of the assumptions in the evolutionary model. However this is a computationally heavy algorithm (Huelsenbeck and Crandall 1997) and the result is still dependent on the model of evolution used. It has been shown that MP, NJ and simple ML which all use oversimplistic models of evolution are more likely to recover the wrong tree due to long branch attraction. (Felsenstein 1981). A significantly better method of tree construction uses Bayesian inference to generate phylogenies.

Bayesian analysis (BAYES) is similar to ML in that a model of evolution is imposed and the program searches for the best trees that are consistent with the model. However Bayesian methods are different from classical statistics in that, although parameters are unknown one can define a distribution which includes all possible values of the parameter, this is known as the prior distribution. This distribution is generated from prior knowledge of the life or evolutionary history of an organism. The theorem then generates a posterior distribution of the parameter given the data (Yang 2006). BAYES is a particularly attractive method as it allows the user to incorporate background information such as fossil data into the model (Beaumont and Rannala 2004).

Bayesian inference will also calculate the posterior probability which is used to evaluate the marginal probability of the data. This is a sum of all possible tree topologies integrated over all branch lengths in those trees and over all parameters in the substitution model. This computation is made possible through the application of Markov Chain Monte Carlo algorithms

University of Cape Town

(MCMC). The MCMC method searches the tree space along multiple independent paths that occasionally exchange information (Hall 2004). The computation starts with a tree of specified topology and branch length and a specific model of DNA substitution. This is the initial chain, a new state of the chain is proposed and the probability of this state, given the old state is calculated. A random number between 0 and 1 is drawn. If that number is less than the calculated probability of the new state, the new tree is accepted otherwise the old state remains, this is a single generation. BAYES does not distinguish between model parameters and data; instead, both are random values with a joint probability distribution (Beaumont and Rannala 2004). MCMC prevents BAYES from being trapped in a local optimum (Hall 2004). At some point the process will converge on a set of most likely trees. Linder et al. (2005) suggested that Bayesian methods are generally successful at finding optimal levels of smoothing to correct for rate heterogeneity, and are less sensitive to undersampling than other methods.

Alignment of the chacma haplotypes (n=50) generated for this study recovered 33 unique haplotypes which were used in the final analyses. An additional 17 chacma haplotypes were included from published articles (Newman et al. 2004; Zinner et al. 2009) and a Phd dissertation (Burrell 2008). Four sequences sourced from Genbank, representing each of the other *Papio* species are included in the alignment. Trees were rooted with *T. gelada* (n=1), *Macaca mulatta* (n=1) *and Macaca sylvanus* (n=1). An additional unpublished sequence representing *P. griesipes* (n=1) was made available by Andy Burrel . The resultant input file was a multiple alignment of 58 sequences. Sequence details are reported in Appendix 4A. Phylogenetic trees were constructed using MP and BAYES methods.

University of Cape Town

ii. Choosing a model of nucleotide substitution

Nucleotide substitutions are considered to be random events, a model of substitution can be specified *a priori* to provide an appropriate statistical description of this stochastic event.
The model of nucleotide substitution that best approximates reality for a dataset is chosen from one of the following. The Tamura Nei (TN93) model (Tamura and Nei 1993) allows for different substitution rates for transition and transversions as well as for purine transition and pyrimidine transitions. If the purine and pyramidine transition are assumed to be equal this model reduces to the HKY model (Hasegawa, Kishino and Yano 1985). If the base frequencies are equal then the HKY further reduces to Kimura 2 parameters (K2) model (Kimura, 1980). If the ratio of transition rate to transversion rate equals 1, then HKY becomes the F81 model (Felsenstein

1981) and the K80 model becomes JC69. The F84 model (Thorne et al. 1992; Felsensten 1993) is a special case of the HKY85. The Jukes Cantor (JC69) is the simplest model and assumes that each of the four nucleotides, AGCT occur in equal frequencies and have that each has an equal probability of replacing another (Jukes and Cantor 1969; Strimmer 2003). The HKY model was selected as the optimal model of nucleotide evolution for this data set according to the Aikaike Information Criterion (Aikaike 1974, 1980). This was determined using MODELTEST v3.6 (Posada and Crandall 1998).

iii. Constructing an MP tree

The MP tree was constructed on MEGA v4.0 and a majority consensus tree based on a 50% cut-off is shown here (Fig 4.2). The possible number of trees that can be generated by an algorithm increases exponentially with each sequence that is added to the calculation. Maximum parsimony searches for the optimal (minimal) tree. In this process more than one minimal tree may be found. In order to find the best possible tree an exhaustive evaluation of all possible tree topologies can be performed. An exhaustive search is one which calculates the score for every possible tree and identifies the tree with the best score. This is usually too computationally heavy for all but the smallest data sets. Alternative methods therefore been developed to speed up the search process. The branch and bound algorithm (Hendy and Penny 1982) does not attempt to construct every possible tree but it is still quite slow and can only efficiently be used on a very small data set. Other "heuristic" strategies start with a tree and rearrangements are sought to improve the tree score. A starting tree is generated either randomly or through a faster algorithm such as NJ. A set of similar trees (neighbours), are generated and compared to the starting tree. These neighbours are swapped by nearest neighbour interchange (NNI) which swaps the placement of subtrees. This can be done by subtree pruning and regrafting (SPR) where a subtree is pruned and reattached or by tree bisection and reconnection (TBR) where the tree is cut in two parts and then two branches are reconnected. The algorithm then uses optimality criteria to decide whether to move to one of these neighbour trees or not. This is repeated until there is no further improvement to the tree score (Yang 2006). Here the MP analysis was run with NNI search factor of 2 and a mini-heuristic search with level 100 was used to find the most parsimonious trees.

Bootstrapping is a statistical estimate of the reliability of each node in a tree. This method takes a sub sample of the characters in the alignment and creates trees based on those sub samples.

University of Cape Town

The proportion of times that a particular node is recovered is an estimate of the confidence that can be placed in that node (Hall 2004). This is the bootstrap value. Here I performed non-parametric bootstrap analysis with 1000 replicates and tree topology was optimized. The HKY model of nucleotide substitution was employed and empirical base frequencies were estimated directly.

iv. Constructing a Bayesian tree and estimating node ages

The software package BEAST v2.0.1 was used for Bayesian tree construction. This algorithm implements a 'relaxed Bayesian phylogenetic' method in which tree topology and branch lengths are estimated simultaneously from the data (Drummond and Rambaut 2007). By calibrating a number of nodes within the resulting tree, BEAST generates divergence estimates that allow for substitution rate changes across the tree and for variation around the age of fossil constraints (Drummond and Rambaut 2007; Ho et al. 2008). Parameters for phylogenetic tree construction are set up in BEAUTi which is distributed with the BEAST package. Here I selected the tree Yule tree as the tree prior as it has been shown to be the most suitable prior for trees describing relationships between individuals from different species (Drummond and Rambaut 2007). Nonclock-like evolution is common, and the departure from homogeneous rates can be very strong. Ho et al. (2005) reported a consistently faster mtDNA molecular clock at recent vs. ancient time scale. This could perhaps be explained by the existence of undetected, short-lived mutation hotspots within the genome (Galtier et al. 2006; Pulque´rio and Nichols 2007; Stoneking 2000)**.** In node age estimation one should therefore take care to adopt molecular dating methods which do not assume constant rate across the tree. Here I employed a relaxed molecular clock with rates for each branch drawn independently from an uncorrelated lognormal distribution. I used the HKY +G substitution model parameter as implemented in BEAST, and the gamma distribution was modeled with four categories.

To estimate node ages I employed three fossil dates for calibration. Following Steiper and Young (2006) the divergence between *Macacina/Papionina* was set at 7-8 mya while the split between *Papio/Theropithecus* was set at 3.5-4.0 mya (Jablonski 1993). Because rates of molecular change are generally greater within species than between, the inclusion of fossil calibrations for both ancient and more recent divergence events helps to prevent overestimating the age of terminal nodes (Ho et al. 2007). I therefore also used an internal node for the coalescence of the genus *Papio*, based on fossil evidence for the emergence of modern *Papio*

University of Cape Town

and conservatively calibrated to 1.0 - 2.7 mya (Heaton 2007). Although the taxonomy of African fossil Papionins is uncertain, a recent morphological assessment of South African fossil *Papio*nins places *P. augusticeps* and *P. robinsoni* within the modern radiation possibly emerging from ancestral *P. izodi* (Heaton 2007). The age estimations for these taxa are 1.0 -2.7 mya for *P. robinsoni* and 1.5 - 2.0 mya for *P. augusticeps* (El-Zaatari et al. 2005). There is some controversy surrounding the dates of *P. robinsoni* from the Sterkfontein fossil site in South Africa as this layer may be intrusive, I did however include this upper date (2.7 mya) here, to be conservative. In an assessment of calibration techniques, Ho et al. (2007) concluded that the lognormal distribution is the most appropriate for modeling palaeontological information, as the first appearance date of fossils most likely post-date the actual species divergence event. By employing a **distribution** of values around each calibration point the analysis accounts for uncertainty around each point. Calibration values were therefore input as the middle 95% of a lognormal distribution on each MRCA prior, except for the coalescence of *Papio*. Here, I employed a normal distribution to account for non-directional uncertainty accompanying the Sterkfontein date and used a mean of 1.8 Ma with a standard deviation of 0.2.

After several short optimising runs that examined the MCMC performance, two final BEAST runs of 10 000 000 iterations each were performed. Samples were taken every 1000 iterations and convergence was assessed in Tracer v1.4 (Drummond and Rambaut 2007). Model parameters and MRCA estimates were checked (ESS values >100) and tree and parameter estimates from the two runs were combined using Log Combiner v1.4.8. The first 10% of trees were discarded in the burn-in, before samples from the posterior distributions were summarized using the default settings in TreeAnnotator v1.4.8 (Drummond and Rambaut 2007) and the tree output was visualized in Figtree v1.1.2. Parameter estimate means and 95% higher posterior densities (HPD) were obtained from the combined outputs in Tracer v1.4. The 95% HPD corresponds to confidence intervals and represents the shortest interval that contains 95% of the sampled values from the posterior (Drummond and Rambaut 2007).

University of Cape Town

Results

A complete alignment of *Papio* sequences resulted in 181 polymorphic sites, of which 112 were parsimony informative and 69 represented singleton sites. Both tree reconstruction methods recover a major split within the sample which groups sequences into one of two mitochondrial lineages, Southern *ursinus* lineage (SL) or Northern *ursinus* lineage (NL) (Fig 4.2 and Fig 4.3) The fixation of a triple indel autapomorphy is a major feature distinguishing NL from SL. These monophyletic lineages are well supported by both bootstrap and BPP values (Table 4.1). Haplotypes can be further divided into one of 5 terminal clades or haplogroups. These are named according to their geographic provenance. Cape clade (Cc) and the Namibian clade (Nc) comprise the SL with Nc a monophyletic clade nested within Cc. The Limpopo clade (Lc), Zambian clade (Zc) and Eastern clade (Ec) comprise the NL. In relation to the distribution of Hills (1970) phenotypes Cc would coincide with the distribution of the ursinus phenotype, Nc with ruacanna, Lc with griseipes, Zc with jubilaeus (which we now know to be a yellow baboon) and Ec with orientalis. There is no evidence to support an occidentalis, ngamiensis or chobiensis clade. These phenotype localities as described by (Hill 1970) are instead represented by Lc haplotypes. Bootstrap and BPP values for all haplogroups are reported in Table 4.1.

Tree topologies from both analyses were virtually identical except for the allocation of a single individual; *Caprivi1* which is allocated to the SL by the MP analysis and NL by the BAYES reconstruction (Fig. 4.2 and Fig 4.3). A further incongruence between the two trees is that Nc has a low bootstrap value in the MP reconstruction (29) but is well supported by a BPP=0.99 in the BAYES reconstruction. Due to the weaknesses inherent in the MP algorithm, further discussion is based on the topology of the BAYES reconstruction.

Bayesian dating estimates indicate that the modern chacma separated from the rest of *Papio* ~1.80 Ma (node D on Fig. 4.3) and diversified into two strongly supported mitochondrial lineages at ~1.6 Ma (node E). These lineages achieve reciprocal monophyly by ~1.27 Ma (node F) and ~1.35 Ma (node G), respectively. SL essentially comprises the Cc in which the Nc is nested. Bayesian dating estimates indicate that the Nc haplogroup differentiated from Cc ~1.00 Ma (node H, BPP 0.99). The northern lineage includes an Ec which differentiated at 0.42 Ma (node I, BPP 1.00) and a possible but weakly supported Zc which differentiated at 0.29 Ma (node J, BPP 0.18).

University of Cape Town

Figure 4.4 shows the geographic distribution of each of the clades (Cc, Lc and Ec, Nc and Zc) identified in the phylogenetic reconstructions. The distributions described here are based on Bayesian clade allocations. Although there is some evidence for movement of individuals across clade boundaries the distributions of each clade are geographically quite distinct. Cc (pink) extends north from the southwest tip of South Africa to the Orange River, inland up until the edge of the Kalahari Desert and up the east coast to Kwa-Zulu Natal. The Nc (purple) haplogroup could be considered to be a Namibian clade starting in the Richtersveld, just north of the Orange River and going up to central Namibia. Two Cc haplotypes are distributed within the range of Nc, one in central Namibia (*Namibia9*) and a second one unexpectedly far north in Angola (*Angola1*). Nc is therefore genetically and geographically nested within Cc, bounded by Cc haplotypes to the north and south.

Lc (blue) is the most broadly distributed clade within NL. The distribution of Lc extends across a large part of central sub-Saharan Africa and meets Cc in central Namibia to the west, Zc to the north at the Zambezi River and Oc to the south in Mpumalanga. The southern border extends to the edge of the Kalahari Desert. The Ec (green) is a localized subset of Lc found in the Mpumalanga region. This clade shares a border with Cc in the south. A single Cc haplotype was found within the range of Ec in the mountainous Waterberg region of South Africa. Zc (white) is distributed north of the Zambezi River. Three regions have been identified as potential mixing zones between NL and SL, and include, Kwa-Zulu Natal, Central Namibia and Loskop.

University of Cape Town

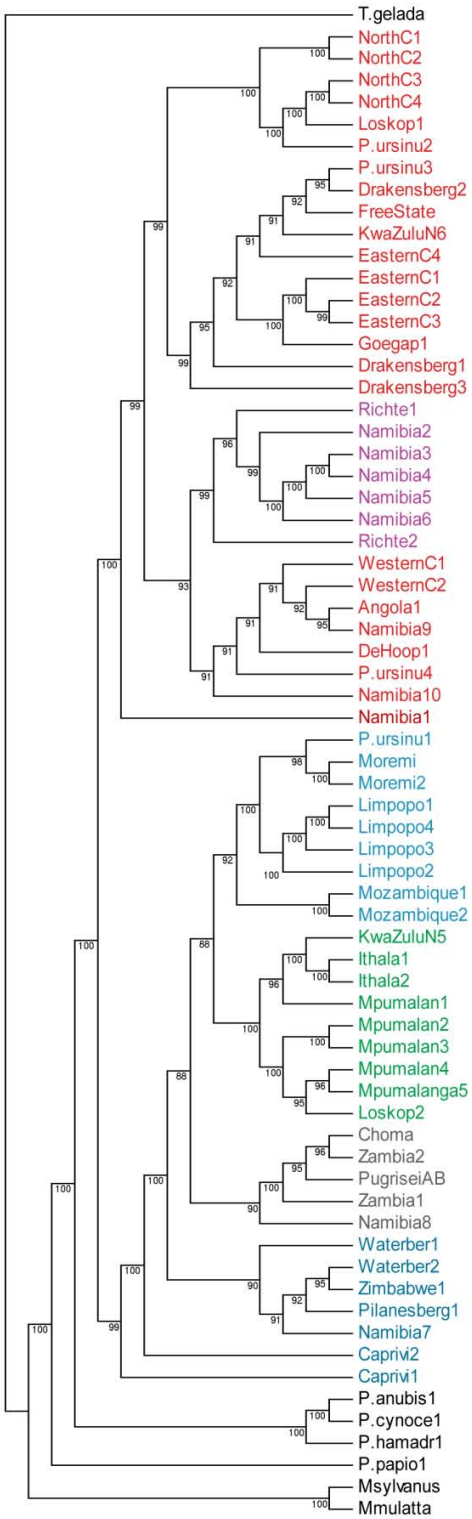

Figure 4.2- Maximum Parsimony tree with bootstrap branch support values based on 1000 bootstrap replicates. Bootstrap support values are reported in Table 4.1, clade allocations are reported in Appendix 4A.

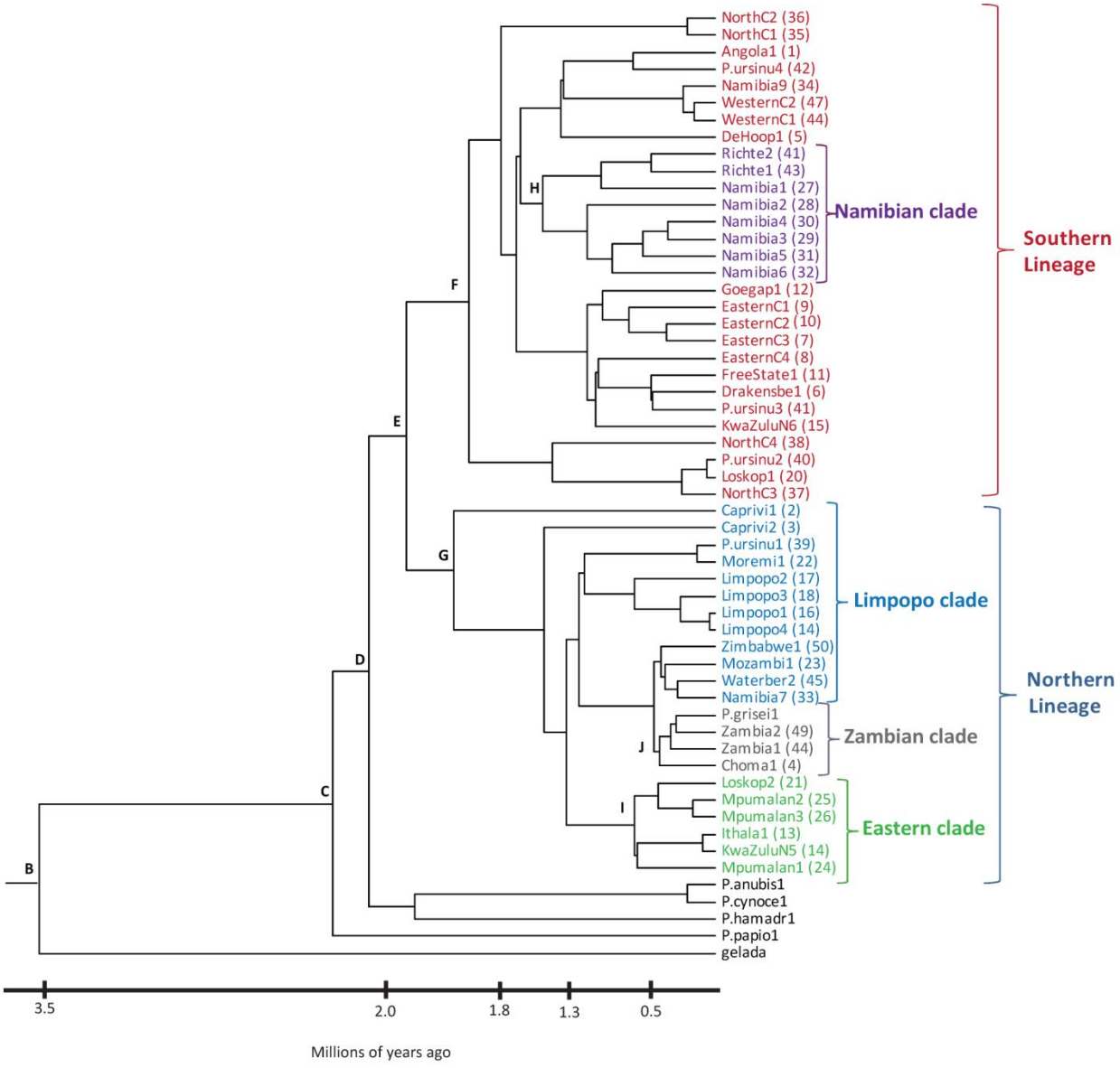

Figure 4.3- Maximum clade credibility tree estimated using Bayesian phylogenetic analysis in BEAST. Posterior probability values for the main nodes are reported in Table 4.1.

Node (clade)	Clade	Node age (Mya)	95% HPD BEAST	BPP BEAST	bootstrap support % for clades MP
A*	*Macaque*-African *Papio*nins	6.95	6.69 - 7.82	1	100
B*	*Theropithecus* - *Papio*	3.50	3.38 - 3.94	1	100
C*	*Papio* root	1.98	1.73 - 2.45	1	100
D	*P. ursinus* root	1.80	1.73 - 2.45	1	99
E	*NL /SL split*	1.60	1.31 - 2.18	1	86
F	Cc+ *Nc*	1.27	0.85 - 1.77	1	91
G	*Lc*+ *Ec*+ *Zc*	1.35	0.90 - 1.86	1	100
H	*Nc*	1.00	0.53 - 1.30	0.99	29
I	*Ec*	0.42	0.19 - 0.86	1	91
J	*Zc*	0.29	0.04 - 0.69	0.18	8

Table 4.1- Bayesian divergence estimates in millions of years before present (mya). Nodes are labelled as in Figure 4.3. *Nodes A-C were used in calibration; node A is not shown in Figure 4.3. Values represent the mean node age (mya), 95% highest posterior distribution (HPD) and Bayesian posterior probability (BPP). Bootstrap values for each of the tree constructions are also shown.

University of Cape Town

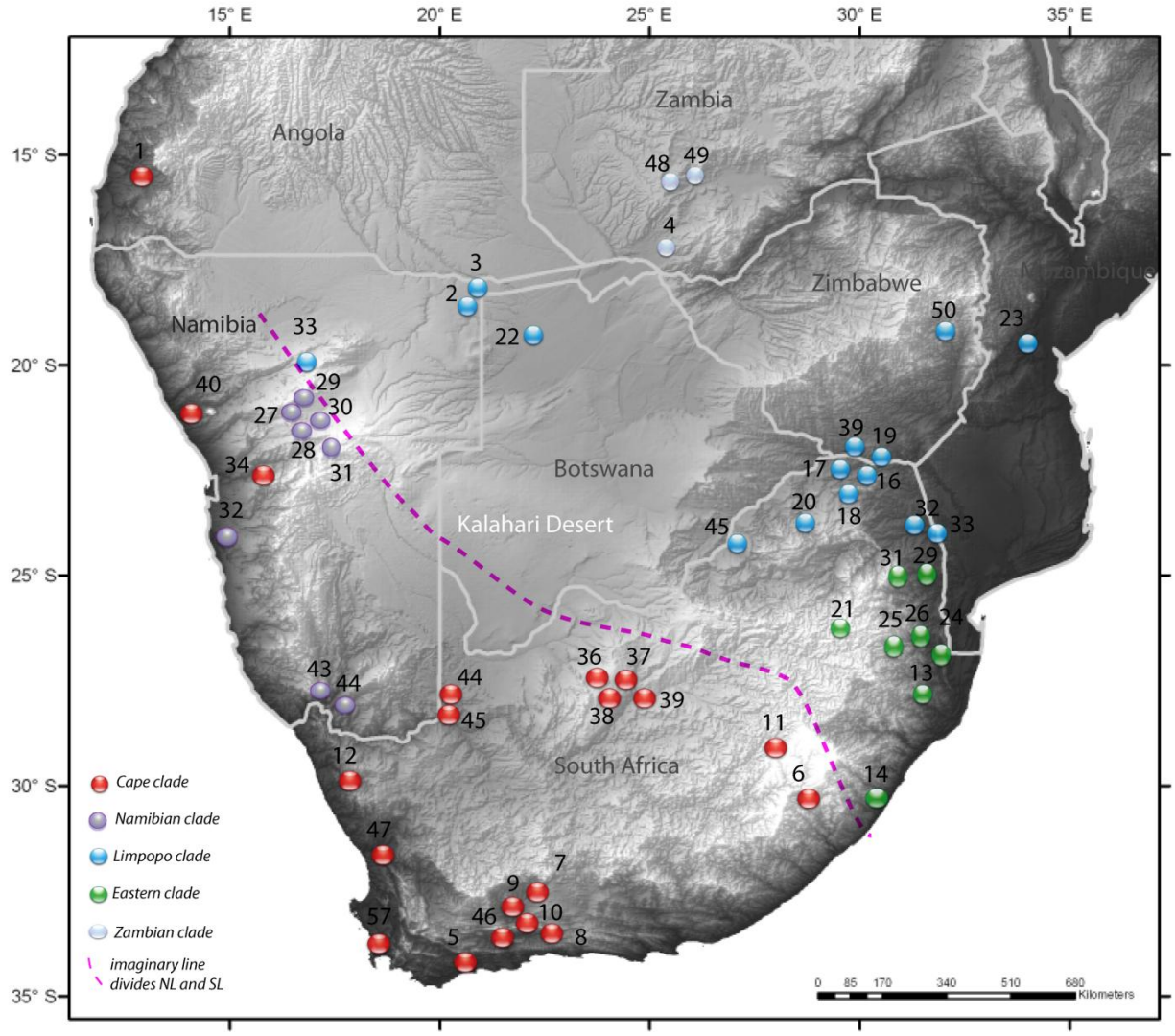

Figure 4.4- Map showing the approximate distributions of each of the five clades identified in the Bayesian reconstruction and sampling localities of the sequences used in this study. A line shows the division between NL and SL. See Appendix 4A for details.

Discussion

Results from this study suggest that a precursor to chacma baboons diverged from the ancestral *Papio* as an independent lineage during the Pleistocene ~1.80 Ma. This estimate is congruent with published ages (Newman et al. 2004, Sithaldeen et al. 2009; Zinner et al. 2009) for the emergence of chacma baboons. Chacma baboons then diverge at ~1.6 Ma and fully emerge as independent lineages ~300kyr later. Zinner et al. (2009) recover a slightly earlier date for this

fragmentation event at ~1.8 Ma [1.28-2.36]. To be conservative, this major diversification event within the chacma mitochondrial lineage therefore can be estimated to have occurred between 1.6 and 1.8 Ma and is contemporaneous with the diversification of another large mammal lineage, the southeastern bushbuck haplogroup, *Tragelaphus scriptus,* at ~1.7 Ma (Moodley and Bruford, 2007). Moodley and Bruford (2007) attribute this event to a reduction in gene flow between populations as a result of the expansion of grasslands in southern Africa (Cerling 1992). Given the ubiquitous distribution of the chacma baboons today and their general success in present day grasslands, it does not seem likely that the expansion of grasslands would drive population fragmentation within chacma baboons. Instead, this fragmentation may be the result of the extreme aridification that led to the expansion of grasslands and the subsequent desertification of central southern Africa (Jürgens 1997; Meadows 2001; van Zinderen Bakker 1978).

The Kalahari Desert, which extends from 22.1°S to the Orange River at 29.1°S (Shaw and Thomas 1996), is the only substantial, natural, modern day barrier to the distribution of chacma baboons in southern Africa. Today most of the Kalahari is stabilized, with active dunes only in the southwest (Stokes 1998). During the last 1.8 myr large scale glaciations drove phases of aridification in southern Africa (Jürgens 1997; Meadows 2001; van Zinderen Bakker 1978) resulting in the large-scale deposition of dunes, including those of the Kalahari Desert (Meadows 2001). Globally it is reported that, during these Pleistocene glacial phases, active sand dunes covered much larger areas than they do today (Goudie et al.1973; Grove 1958; Lancaster 1981; Mabbutt 1971). In central southern Africa, four periods of dune activation in the northwest of the Kalahari have been documented for the late Pleistocene giving rise to Mega Kalahari Sands (MKS) (Stokes 1998). The timing of these MKS periods is coincident with periods of global cooling. While there is no direct data for a Kalahari expansion for the period 1.6-1.8 Ma, by extrapolation, the large scale cooling that began at 1.8 Ma is likely to have resulted in a general expansion of the Kalahari Desert and a period of MKS.

Recent work on the Cape rock elephant shrew (*Elephantulus edwardii*) recovers a separation of Namaqua and central Fynbos clades i.e. northwest and southeast of the Kalahari Desert, dated to 1.73 Ma (Smit 2007). In addition, Matthee and Flemming (2002) recover a phylogeographic discontinuity east and west of the Kalahari between clades of the rock lizard, *Agama atra* (2.2–4.4 Ma). Studies of species on a more recent time scale indicate that Kalahari sandflows do act as a physical barrier to gene flow between populations (Deacon and Lancaster 1988; Haacke

University of Cape Town

1989; Lancaster 1989) and can drive diversification e.g. *Elephantulus edwardii* (Smit et al. 2007) and *Micaelamys namaquensis* (Russo et al. 2010), suggesting a significant role for the Kalahari Desert as an agent of vicariance in mammals. The expansion of the Kalahari region may have contributed to part of the hypothesized 'arid corridor' that is proposed to have stretched, during drier periods, from the Horn of Africa to the Cape of Good Hope. This corridor provided a route by which xeric flora and fauna dispersed from central to southern Africa (Knoch and Schulze 1956; van Zinderen Bakker 1969, 1978) and at the same time would have acted as a barrier to more mesic adapted species. An extension of this arid corridor from the Kalahari region to the north east towards the Horn of Africa is supported by the discovery of Pleistocene age discontinuities in bushbuck *Tragelaphus* spp. (Moodley and Bruford 2007) and giraffe (*Giraffa* spp.) (Brown et al. 2007) that partition populations from southern and East Africa from those in west Africa.

The association of diversification in chacma baboons with aridification is unexpected, as chacma baboons can be extremely drought tolerant, for example the well-studied Tsoabis Leopard Park (Davies and Cowlishaw 1997) and Kuiseb Canyon troops in Namibia (Anderson 1982) who are able to cope with extreme water stress . Muwanika et al. (2003) however noted that even the hardy warthog (*Phacochoerus africanus*), which is capable of surviving in extremely arid conditions, experienced periods of genetic isolation during the Pleistocene. Indeed paleoclimatic evidence suggests that for much of this period, and earlier still, the African continent has been very dry, and today's climates are probably closer to the moist, warm end of the scale (Hamilton and Taylor 1992). While chacma baboons have been known to adapt behaviourally to extremely dry living conditions, there does appear to be a threshold to the levels of water stress that these baboons can withstand.

The distribution of the northern and southern mitochondrial lineage haplotypes suggests some degree of correlation between these genetic subdivisions and morphological subdivisions within the taxon, with *P. u. ursinus* in the south and *P. u. griseipes* in the north (Groves 2001; Jolly 1993). These phenotypes are distinguishable in size, with *P. u. ursinus* being the larger of the two*. P.u. ursinus* is also darker*,* tending towards very dark grey while *P. u. griseipes* is fawn coloured and has grey hands and feet the same colour as their limbs, and a longer tail (Hill 1970; Jolly 1993). They are also more lanky in build (Barret et al. 2003). Individuals that carry NL and SL haplotypes mix readily in no less than three sampling localities. Given that NL and SL

University of Cape Town

fully diverged ~300kyr after the initial fragmentation of chacma at 1.6 Ma, we can conclude that the populations were separated for at least this long. If we take the generation time for chacma to be 5 years, based on average age of first time of reproduction in chacma females (Ascunce 2006; Zunino 1996), then NL and SL were separated for at least ~60000 generations. In all this time, although post zygotic fertility has not been assessed, it appears that no reproductive discontinuity has emerged between the two lineages. This supports the hypothesis that, given time, genetically isolated baboon populations can accumulate a significant amount of phenotypic difference without the formation of obvious reproductive barriers. Of course, the degree of genetic isolation between lineages during this period remains to be tested. Furthermore, at areas of mixing, there is no obvious phenotypic difference within troops (pers. obs.) which reinforces the importance of nuclear swamping in generating phylogenetic complexity within *Papio* (Zinner et al. 2009).

This analysis also reveals three further geographic groupings of related haplotypes (shown in Fig 4.4). These are the Namibian clade (possibly representing *P.u.ruacana*), Ec and Zambian clade. The results of Keller et al. (2010), based on variation in the mitochondrial *cytochrome-b* marker in chamca baboons, do not support the delineation of an independent Namibian clade. This may be an artifact of sampling as Keller et al. (2010) do not include samples from the very south of the *P.u. ruacana* range in Nambia, or it may be that the difference in our findings are influenced by the relative history represented by the two markers. It is possible that the Brown region, which evolves more slowly, retains the signature of an earlier differentiation event. Further sampling together with the inclusion of nuclear markers should provide valuable data with which to further investigate incongruence. Keller et al. (2010) do recover a robustly supported north eastern clade within their sample which is represented here by the Eastern clade.

Hill (1970) suggested that *P. u. ruacana* is a Namibian subspecies distributed north of the Orange River and into Angola. In this analysis a cluster of haplotypes (Namibian clade) was recovered within SL from just north of the Orange River to central Namibia that date to ~1.0 Ma (Fig.4.3). Geographically, this clade is distributed between Cape clade haplotypes to the south and a single Cape clade haplotype from Angola in the north. A period of intensified cooling is thought to have occurred sometime between 900 and 700 kya and as a result led to a second glacial maximum with the dominant periodicity of glacial-interglacial cycles shifting from 41kyr to100 kyr (e.g. Hooghiemstra et al. 1993; de Menocal 1995). The node age of the emergence of

University of Cape Town

a Namibian clade coincides with this shift to prolonged periods of cooling and aridification in Africa. It is therefore possible that diversification of the *P.u. ruacana* lineage resulted from the formation of a genetic barrier between baboons north and south of the Orange River (the largest river crossing the arid interior of southern Africa). Rivers have been shown to be important determinants of gene flow in a number of large primates e.g. gorillas (*Gorilla gorilla*) (Anthony et al. 2007) and chimpanzees (*Pan troglodytes*) (Gagneux et al. 2001) and in the case of orangutans (*Pongo pygmaeus*) are particularly disruptive to the movement of females i.e. the dispersal of mitochondrial DNA (Arora et al. 2010). It is possible that the dynamic history of the Orange River that has included periods of reduced flow rate and catastrophic flooding (Dollar 1998; Zawada 1995) could have contributed to the diversification of a Namibian clade. However apes cannot swim while monkeys can, so an alternative explanation may be that it is the flat, seasonally-flooded grassland without trees, and thus a lack of night refuges, that act as a boundary to to distribution of baboons near the Orange River, much like in south west Zambia today, which is apparently another baboon free area (Cliff Jolly, pers comm). The single haplotype *Angola1*, which belongs to the Cape clade that is distributed north of the Namibian clade may represent a later, unique event of natural or anthropologically facilitated dispersal. Alternatively these haplotypes could represent a once relict group that became isolated from the Cape clade and, as a result of founding effects, diversified rapidly. Later, probably due the expansion of suitable habitats, this group expanded to meet Cape clade haplotypes in the north and south. The most northerly group within this data set is the Zambian clade. While the geographic clustering of genenetically related haplotypes may be due to a diversification event, there is very little statistical support for a monophyletic Zambian clade, distinct from the rest of NL.

Finally a well supported Eastern clade (possible orientalis phenotype) differentiated from NL ~420 kya. This clade is also reported in Keller et al (2010). The node age of this clade is coincident with Marine Isotope Stage 11 which spans from 420 to 360 kya. This period of earth history was remarkable as an extraordinarily long interglacial period and it is possible that the related change in habitat contributed to the divergence of of this clade. During warm wetter periods, forests expanded (Lawes 1990); these expanded tracts of forest may have acted as significant biogeographical barriers to gene flow in baboons (Zinner et al. 2009) and this may have been the case along the great escarpment of southern Africa, resulting in the isolation of the lineage that gave rise to orientalis from griseipes. Chacma baboons however, do show a tolerance for forest living today, as observed in Tsitikamma, St Lucia and Mpumalanga in South

Africa (pers. obs.); forests or other barriers to gene flow therefore may not necessarily have therefore resulted in complete genetic isolation (particularly in the absence of competition with tree dwelling species) between populations as is assumed in the allopatric speciation model presented by Zinner et al. (2009). Incomplete isolation between groups may explain why, after 400 kyr since diversification, Eastern clade is nested within NL and not sister to it.

Conclusions

The strong concordance between diversification within the chacma baboon lineage and climate driven landscape change is evidence for the important role played by climate change, habitat shifts and local geography in structuring diversity within the *Papio* lineage as a whole. In addition to forests as barriers to gene flow, these results also introduce a variety of important landscape features that have also contributed to vicariance within the genus, and suggests a more complex history of events in structuring genetic and phenotypic variation within baboons than has previously been appreciated.

The suggestion of a relationship between morphological and molecular groups (albeit this pattern is shown for a maternal molecular marker only) demonstrates the power of increased taxon sampling in phylogenetic inference and provides further support for the relationship between genetic and phenotypic divergence within *Papio*. Additionally, as in the case of *P.u. ursinus* and *P.u. griseipes,* these results reveal how significant phenotypic diversification can be generated within baboons with little or no decrease in the reproductive compatibility of lineages, even after 300 kyr.

Recommendations for future phylogenetic research on *Papio* include an increased focus on sampling resolution of individual taxa; while significant sampling has been carried out within hamadryas baboons (Wildman et al., 2004) and to a lesser extent other baboon allotaxa (Zinner et al. 2009), many gaps remain. Within the the chacma baboon, further sampling from Zimbabwe and Mozambique, as well as at the boundaries between the five species, is important both to fill in the geographic gaps and to capture a more representative sample of the total variation available within the species. It would also be useful to increase resolution at areas of potential gene flow between mitochondrial lineages. It is likely that detailed sampling of nuclear markers at the contact point between *P. u. ruacana* and the Cape chacma may provide additional insight into the ecological drivers of lineage diversification in *Papio* by providing new perspectives on

University of Cape Town

baboon population history Ideally, broader molecular sampling would be combined with detailed morphological and behavioural data, in order to better understand the extent to which these data correspond. With in creased fine scale sampling of individual *Papio* taxa and the addition of nuclear markers, we will better understand the complex evolutionary history of this genus.

Appendix 4A- Collection details for samples used in this chapter and GenBank accession numbers of the mitochondrial Brown region sequences used in chapter 4. Haplotype designations follow Newman et al. (2004).

Tree ID	Taxon	Map ID	Haplotype	Sample locality	GPS co-ordinates	Genbank Accession no.
Angola1	*P. ursinus*	1	U53	Serra Leba, AO	15°08'30 "S, 13°14'30"E	EU885826
Caprivi1	*P. ursinus*	2	U37	Rundu, NA	18° 09'09.16"S, 21°36'15.60"E	FJ531529
Caprivi2	*P. ursinus*	3	U38	Rundu, NA	18° 09'09.16"S, 21°36'15.60"E	FJ531530
Choma1	*P. ursinus*	4	U40	Choma, ZW	17°51'12.71"S, 25°51'18.97"E	FJ531532
DeHoop1	*P. ursinus*	5	U54	De Hoop NR, ZA	34°27'22 "S, 20°24'23"E	EU885833
Drakensbe1	*P. ursinus*	6	U56	Giants Castle, ZA	29°20'00 "S, 29°29'00"E	EU885831
EasternC1	*P. ursinus*	7	U20	Leeu Gamka, ZA	33°20'38.16"S, 21°42'56.43"E	FJ531512
EasternC2	*P. ursinus*	8	U21	Leeu Gamka, ZA	33°20'38.16"S, 21°42'56.43"E	FJ531513
EasternC3	*P. ursinus*	9	U22	Leeu Gamka, ZA	33°20'38.16"S, 21°42'56.43"E	FJ531514
EasternC4	*P. ursinus*	10	U23	Nieu Bethesda, ZA	31°51'59.04"S, 24°33'17.28"E	FJ531515
FreeState1	*P. ursinus*	11	U26	Golden Gate NP, ZA	28°33'49.42"S, 28°35'47.35"E	FJ531518
Goegap1	*P. ursinus*	12	U55	Goegap NR, ZA	29°41'58 "S, 18°01'56"E	EU885832
Ithala1	*P. ursinus*	13	U63	Ithala NR, ZA	27°32'00 "S, 31°16'00"E	EU885822
KwaZuluN5	*P. ursinus*	14	U24	Hluhluwe NR, ZA	28° 00'31.16"S, 32°16'09.76"E	FJ531516
KwaZuluN6	*P. ursinus*	15	U25	Oribi Gorge NP, ZA	29°49'40.50"S, 31°01'17.07"E	FJ531517
Limpopo1	*P. ursinus*	16	U27	Messina, ZA	22°20'56.85"S, 30° 02'23.51"E	FJ531519
Limpopo2	*P. ursinus*	17	U28	Messina, ZA	22°20'56.85"S, 30° 02'23.51"E	FJ531520
Limpopo3	*P. ursinus*	18	U29	Messina, ZA	22°20’58.22”S, 30° 2'23.11"E	FJ531521
Limpopo4	*P. ursinus*	19	U30	Messina, ZA	22°20'58.22"S, 30° 2'23.11"E	FJ531522
Loskop1	*P. ursinus*	20	U61	Loskop Dam, ZA	25°08'15 "S, 29°22'43"E	EU885825
Loskop2	*P. ursinus*	21	U62	Loskop Dam, ZA	25°08'15 "S, 29°22'43"E	EU885824
Moremi1	*P. ursinus*	22	U39	Moremi NP, BW	19°27'53.18"S, 23° 06'12.49"E	FJ531531
Mozambi1	*P. ursinus*	23	U49	Gorongosa NP, MZ	18°58'42"S, 34°21'40"E	EU885814

University of Cape Town

Appendix 4A- Collection details for samples used in this chapter and GenBank accession numbers of the mitochondrial Brown region sequences used in chapter 4. Haplotype designations follow Newman et al. (2004).

Tree ID	Taxon	ID	Haplotype	Sample locality	GPS co-ordinates	Genbank Accession no.
Mpumalan1	*P. ursinus*	24	U70	Blyde River , ZA	24°51'40.36"S, 30°50'47.46"E	FJ531523
Mpumalan2	*P. ursinus*	25	U32	Blyde River , ZA	24°51'40.36"S, 30°50'47.46"E	FJ531524
Mpumalan3	*P. ursinus*	26	U34	Blyde River , ZA	24°45'37.10"S, 30°46'21.75"E	FJ531526
Namibia1	*P. ursinus*	27	U71	Okahanja, NA	21°56'50"S, 16° 50'55"E	FJ531504
Namibia2	*P. ursinus*	28	U13	Okahanja, NA	21°56'50"S, 16° 50'55"E	FJ531505
Namibia3	*P. ursinus*	29	U14	Okapuka, NA	22°11'40"S, 17° 04'31"E	FJ531506
Namibia4	*P. ursinus*	30	U15	Okapuka, NA	22°11'40"S, 17° 04'31"E	FJ531507
Namibia5	*P. ursinus*	31	U16	Okapuka, NA	22°11'40"S, 17° 04'31"E	FJ531508
Namibia6	*P. ursinus*	32	U71	Swakopmund, NA	22°40'24.94"S, 14°32'04.01"E	FJ531509
Namibia7	*P. ursinus*	33	U51	Waterberg Plateau, NA	20°30'16"S, 17°14'32"E	EU885819
Namibia9	*P. ursinus*	34	U59	Spreetshoogte NR, NA	23°38'51 "S, 16°12'20"E	EU885828
NorthC1	*P. ursinus*	35	U6	Augrabies NP, ZA	28°36'02.46"S, 20°19'32.11"E	FJ531498
NorthC2	*P. ursinus*	36	U7	Augrabies NP, ZA	28°36'02.46"S, 20°19'32.11"E	FJ531499
NorthC3	*P. ursinus*	37	U8	Barkely West, ZA	28°43'45.15"S, 24°45'13.01"E	FJ531500
NorthC4	*P. ursinus*	38	U9	Barkely West, ZA	28°43'45.15"S, 24°45'13.01"E	FJ531501
P.ursinu1	*P. ursinus*	39	U1	unknown	22°12"S, 29°23"E	AY212105
P.ursinu2	*P. ursinus*	40	U2	unknown	unknown	AY212057
P.ursinu3	*P. ursinus*		U3	unknown	unknown	AY212058
P.ursinu4	*P. ursinus*		U5	unknown	unknown	AY212059
Richte1	*P. ursinus*	43	U10	Potjiespram, ZA	28°04'16.80"S, 16°57'59.00"E	FJ531502
Richte2	*P. ursinus*	44	U11	Potjiespram, ZA	28°04'16.80"S, 16°57'59.00"E	FJ531503
Waterber2	*P. ursinus*	45	U36	Lapalala, ZA	24°17'53.88"S, 28° 06'36.00"E	FJ531528
WesternC1	*P. ursinus*	46	U18	Cape Peninsula, ZA	34° 9'27.23"S, 18°24'25.58"E	FJ531510
WesternC2	*P. ursinus*	47	U19	Elandsbaai, ZA	32°33'09.84"S, 18°20'35.97"E	FJ531511

University of Cape Town

Appendix 4A- Collection details for samples used in this chapter and GenBank accession numbers of the mitochondrial Brown region sequences used in chapter 4. Haplotype designations follow Newman et al. (2004).

Tree ID	Taxon	ID	Haplotype	Sample locality	GPS co-ordinates	Genbank Accession no.
Zambia1	*P. ursinus*	48	U68	Kafue, ZM	15°44'16 "S, 25°59'53"E	EU885812
Zambia2	*P. ursinus*	49	U69	EU885811	15°46'25 "S, 26°00'39"E	
Zimbabwe1	*P. ursinus*	50	U48	EU885813	19°03'28"S, 32°48'46"E	
P.grisei1	*P.griseipes*	49	U41		unavailable	
P.anubis1	*P. anubis*	50	U42		unknown	AY212039
P.cynoce1	*P. cynocephalus*	51	U43		unknown	AY212043
P.hamadr1	*P.hamadryas*	52	U44		unknown	AY212034
P.papio1	*P.papio*	53	U45		unknown	AY212049
mulatta	*M.mulatta*	54	U47		unknown	AY612638
sylvanus	*M.sylvanus*	55				
gelada	*T.gelada*	56	U46		unknown	AY212130

Table Legend

TreeID- Identifier used on the phylogenetic tree constructions (Fig. 4.2 and Fig 4.3).

Taxon- Species name.

ID- Identifier used to show sample locality on Fig 4.1.

Sample locality- Sample provenance, place name and country. Countries are identified by two letter codes; AO= Angola, BW= Botswana MZ= Mozambique, NA= Namibia, ZA= South Africa, ZW= Zimbabwe, ZM= Zambia.

University of Cape Town

CHAPTER 5

PHYLOGEOGRAPHY OF THE CHACMA BABOON

Abstract

Contemporary patterns of genetic variation and structure within a species are the product of genetic drift, selection and both historical and contemporary gene flow. Gene flow is governed both by geographical factors and life history traits such as dispersal behaviour and ecological adaptation. Here I aim to explore how palaeoenvironmental change affects population responses and subsequently evolutionary history of a baboon species thereby driving structure within it. This is achieved through the analysis of a neutrally evolving mitochondrial genetic marker. I investigate the distribution of neutral genetic diversity across chacma baboon populations in southern Africa and examine patterns of genetic structure in order to reconstruct past population histories within the species. DNA sequence data from the mitochondrial D-loop are analysed, representing 132 individuals from South Africa, Namibia, Botswana and Zambia. I describe the distribution of diversity within and between chacma baboon populations and investigate spatial structure within the dataset. I also analyse the phylogenetic relationships between individuals, ascertain the number of genetically distinct chacma populations represented within the sample and reconstruct population demographic histories for each of these. These analyses recovered at least three distinct chacma clades (SoC, NwC and NeC), nested within two mitochondrial lineages. Reconstruction of population history of these cladess indicates that both SoC and NwC have experienced population expansion events, and evidence points to a population bottleneck and glacial refuge use for baboons in the southwest. NeC may represent a later expansion of NwC that later became isolated from the parent population due to the development of biogeographic barriers to gene flow between them. Further haplogroup subdivisions are compared with phylogeographic patterns of other southern African species. Results confirm that population contractions and expansions in response to climate driven landscape change contribute significantly to driving diversity within the species.

University of Cape Town

Introduction

The model of diversification for *Papio* (Jolly 2001) presented in chapter 1 proposes that the current genetic structure observed within the genus is the result of several associated geographic and biological processes. The first is climate and habitat change. Together, due to their combined effect on gene flow between populations, these two factors have influenced population genetic structure within many mammalian lineages in Africa and Europe (Arctander et al. 1999; Arora et al. 2010; Brouat et al. 2009; DeMenocal 2004; Ducroz et al. 1998; Eggert et al. 2002; Flagstad et al. 2001; Herron et al. 2005a; Hewitt 2000; Hewitt 2004; Lorenzen et al. 2009; Matthee and Robinson 1997; Muwanika et al. 2003; Russo et al. 2010). In chapter 4, I explored the correlation between climate driven landscape change and diversification within chacma baboons through a phylogenetic analysis of mitochondrial Brown region data. The results identify temporal overlaps between diversification events within chacma baboons and established climate driven landscape changes in southern Africa. This suggest that habitat changes are likely to explain a significant portion of the structure and diversity observed within chacma baboons today. Nonetheless, it is not clear from the phylogenetic analysis of chacma individuals to what degree demographic changes in local populations of chacma baboons have contributed to contemporary genetic structure in the region.

Phylogeographic studies in southern Africa are rich with examples of species that have diversified as a result of population fragmentation and subsequent expansion from glacial refugia. A sample of these studies is presented in Appendix 5A which illustrates the diversity of lineages that owe their current distribution and genetic diversity to Plio-pleistocene climate driven landscape change. The harsh conditions of the Pleistocene are likely to have driven population responses to habitat change in many other species as well, including *Papio ursinus*.

In Jolly's (1993, 2001) model of climate driven diversification for baboons he provides a mechanism by which climate change and subsequent habitat fragmentation led to range fragmentations and contractions in baboons. Jolly (2001) hypothesised that these genetic bottleneck events are directly related to the level of phenotypic diversity that is observed across *Papio* today and that subsequent range expansions account for the modern distribution of species. The past events that have shaped the genetic structure observed in baboons today cannot be directly observed in ancestral populations; however insight into these processes in a suitable modern day proxy would provide valuable insight into the role of landscape change and population responses in shaping genetic structure in a large

University of Cape Town

bodied, Pleistocene primate. An investigation of this nature requires the use of phylogeographic methods. Phylogeographic techniques can be used to investigate myriad processes related to biological evolution such as historical patterns of gene flow and divergence (Hickerson et al. 2010) the timing and biogeography of speciation (Hewitt 2001; Zinner et al. 2010), how and why populations differentiate (Heaney et al. 2005; Newton et al. 1999; Russo et al. 2006; Russo et al. 2010; Wilfert et al. 2006; Zhang and Jiang 2006) and how hybrid zones form (Hewitt 2001; Karanth et al. 2008; Keller et al. 2010; Newman 1997; Stone 2000; Tung et al. 2008; Vianna et al. 2006; Zinner et al. 2010). Phylogeographic analyses also provide a greater understanding of how paleoenvironmental change affects population and subsequently evolutionary history (Arora et al. 2010; Brown et al. 2010; Hernández and Vrba 2006; Hewitt 2000; Hewitt 1996b; Hewitt 2001; Hewitt 2004; Hickerson et al. 2010; Marshall et al. 2009; Russo et al. 2010; Smit et al. 2010).

Phylogeographic techniques can be used to investigate myriad processes related to biological evolution such as historical patterns of gene flow and divergence (Hickerson et al. 2010), the timing and biogeography of speciation (Hewitt 2001; Zinner et al. 2010), how and why populations differentiate (Heaney et al. 2005; Newton et al. 1999; Russo et al. 2006; Russo et al. 2010; Wilfert et al. 2006; Zhang and Jiang 2006) and how hybrid zones form and are maintained (Hewitt 2001; Karanth et al. 2008; Keller et al. 2010; Newman 1997; Stone 2000; Tung et al. 2008; Vianna et al. 2006; Zinner et al. 2010). Phylogeographic analyses also provide a greater understanding of how paleoenvironmental changes have shaped population and subsequently evolutionary history (Arora et al. 2010; Brown et al. 2010; Hernández and Vrba 2006; Hewitt 1996, 2000, 2001, 2004; Hickerson et al. 2010; Marshall et al. 2009; Russo et al. 2010; Smit et al. 2010).

The phylogeographic analysis of a sample of chacma baboon populations provides the opportunity to investigate genetic structure within this primate species and allow for the exploration of both abiotic and biotic factors in the diversification of the lineage. Phylogeographic methods are used here to quantify genetic variation within current chacma baboon populations and its distribution across the geographic landscape of southern Africa. Further analyses reconstruct the evolutionary relationships of individuals within the sample and infer the distribution and demographic history of past populations. These results are interpreted within the geological and ecological context of the southern African region.

University of Cape Town

Methods

Sampling methods

Powerful phylogeographic inference can be achieved through high resolution taxon sampling. Faecal samples for DNA extraction were collected from 261 free-living baboons from 29 localities in South Africa, Namibia, Botswana and Zambia. The distribution of sampling localities is plotted in Figure 5.1. Samples were collected, stored and processed as outlined in chapter 4 and sample information is tabled in Appendix 5B.

Molecular methods

Papio-specific PCR primers (H12652- 5'AATGTTTGGGTCTGAGTTTATATATCA 3' from Winney et al. 2004 and *P.ursinusL10*, L11574 5'CTATCCCTATGAGGGATAATTATAAC 3' designed in this study) amplified a 473 base pair (bp) fragment of the mtDNA genome commonly known as the 'D-loop' or mitochondrial hypervariable region. This section occurs in the main non-coding region of the mitochondrial DNA molecule. All PCR reactions were performed in a total volume of 25ul containing 25-50ng of genomic DNA, 1x reaction buffer, 2.5mM $MgCl_2$, 0.2mM dNTPs, 0.5pmol/ul primers, and 1 unit *Taq* polymerase (GoTaq, Promega). PCRs were performed on a GeneAmp 9700 thermocycler (Applied Biosystems) at an annealing temperature of 54° C. PCR reactions were sequenced in both directions using BigDye v1.1 chemistry on an ABI 3730 (Applied Biosystems) capillary sequencer. Sequences were edited in BIOEDIT v5.0.9 (Hall 1999) and aligned using CLUSTAL X.

A total of 132 individuals were chosen from the collected samples and sequenced. Nuclear transposition of mtDNA sequences (*numts*) was checked by comparing all sequences generated in this study to a *P. ursinus* sequence (AY212105) isolated from purified mitochondrial DNA, and to a *Theropithecus gelada* sequence (FJ785426). No evidence for *numt* contamination was observed.

Analytical methods

i. Quantifying mitochondrial sequence variation

The results of Sithaldeen et al. (2009) and Keller et al. (2010) were used to identify areas of possible mixing between northern and southern *P. ursinus* mitochondrial lineages. The localities identified as possible mixing zones are Drakensberg, Hluhluwe, Kosi Bay, Loskop, Namibia Central and Waterberg and are highlighted in yellow in Fig 5.1. Individuals from possible mixing zones were excluded when estimating genetic variation and the remaining samples were grouped into one of two *geographic* populations, Northern geographic

University of Cape Town

population (NgP) and Southern geographic population (SgP). The sampling localities of individuals belonging to NgP and SgP are shown in Fig 5.1, in blue and red respectively.

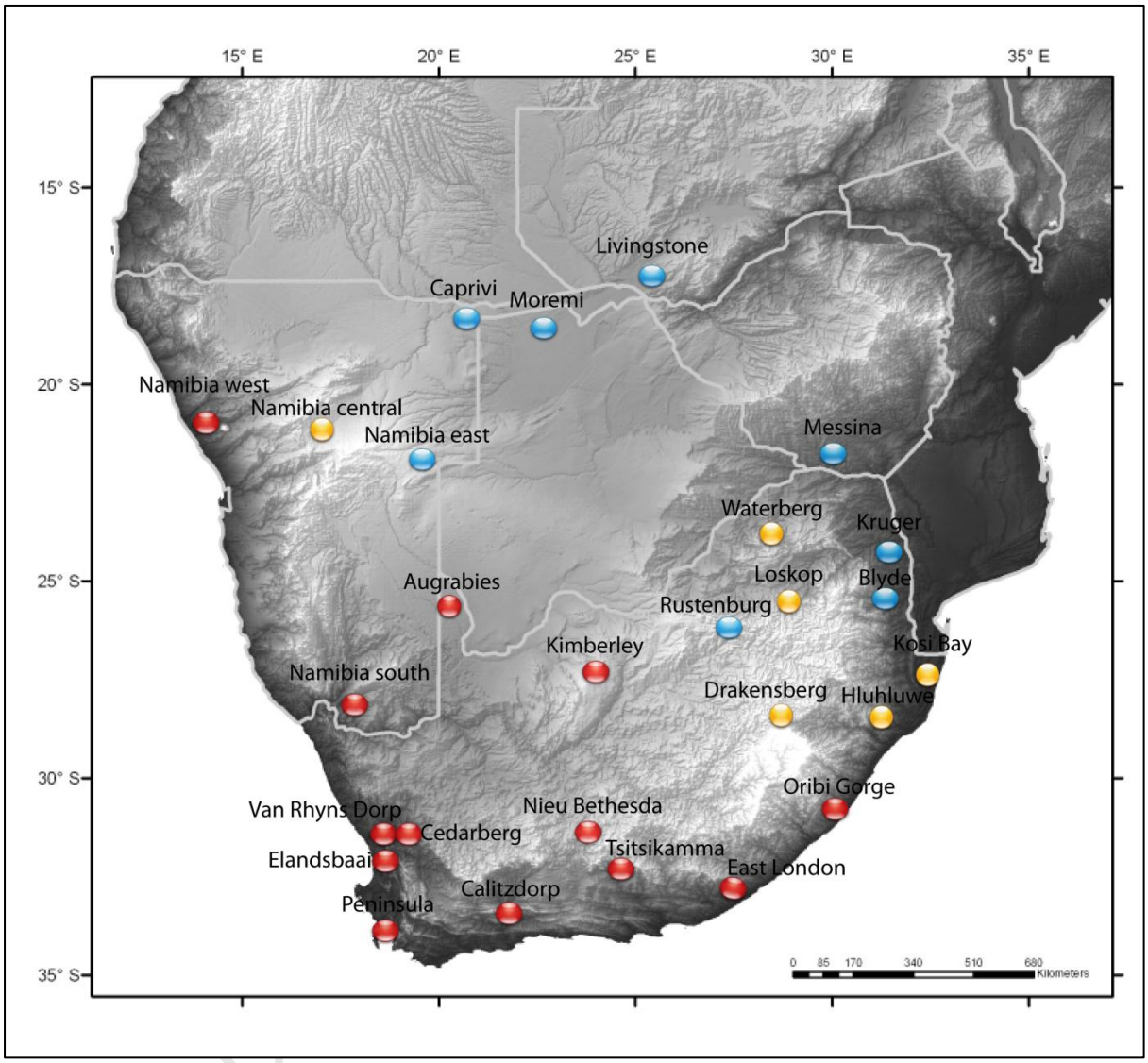

Figure 5.1- Map showing the 29 unique sampling localities from which 261 samples were collected for this study. Appendix 5B provides all sampling information. Zones of admixture are highlighted in yellow and distributions of NgP (blue) and SgP (red) are shown.

Diversity indices were then calculated for each geographic sample using the program DNASP v 5.0 (Librado and Rozas 2009). These indices include haplotype diversity (h) (Nei 1987), nucleotide diversity (π) (Nei 1987) and the number of segregating sites (S). Haplotype diversity measures the probability that two alleles drawn at random from the population will be different from each other (Nei 1987). In DNASP v5.0 this is calculated using equations 8.4 and 8.12 from Nei (1987) but replacing 2n by n for mitochondrial data and is a measure of the uniqueness of an allele averaged for the sample. Nucleotide

diversity is the average number of nucleotide differences between pairs of sequences randomly chosen from the population. This value is affected by the length of the sequences therefore a correction for sequence length is applied (Nei and Li 1979). Pi (π) is the probability that two randomly chosen but homologous nucleotide sites are different (Nei 1987). The number of segregating sites is the number of sites that show variation among the sequences in the sample. If these mutations occur in only a single sequence, they are referred to as "singleton" sites. An excess of singleton sites is generally indicative of a population that has experienced a rapid expansion in the recent past. These indices were calculated for the entire sample (N=132) and each of the two geographic lineages, SgP (N=48) and NgP (N=50). The results of these analyses are shown in Table 5.1.

ii. Spatial distribution of variation

Different hierarchical Analyses of Molecular Variance (AMOVA) are used to evaluate the partitioning of genetic variance within the sample. Here the user defines nested groups of samples based on pre-determined criteria and the AMOVA analysis as implemented in Arlequin v 3.1 (Excoffier et al. 2005) partitions the total variance into covariance components due to intra-individual differences, inter-individual differences, and/or inter-population difference (Excoffier et al. 2005). These covariance components are then used to calculate *Fst* values. The *Fst* statistic can be seen as the correlation between homologous genes taken from a given subdivision level compared to a higher subdivision level (Wright 1943,1951, 1965) and estimates the correlation between genes within a subdivision relative to the genes of the total population; in so doing the method determines if the two groups being compared are significantly different from each other. The calculation of the *Fst* statistics assumes a random mating population and the absence of inbreeding (Excoffier et al. 2005). Here Renold's distance as employed in Arlequin v 5.1 is used to correct for the haploid nature of the mitochondrial data set.

In this AMOVA analysis individual sequences are grouped according to geographic provenance. Unique sample localities were then grouped into geographic areas. These geographic areas and the samples allocated to them are shown in Appendix 5B; each geographic area contains a minimum of 10 individuals. The one exception is *Loskop* which comprises only 9 unique haplotypes. Kimura two-parameter distances, which account for variation in the transition/transversion ratio and for rate variation among sites in DNA sequences, were used to calculate the F-statistics; these pairwise estimates are tabled in Appendix 5C.

University of Cape Town

To complement the AMOVA, a spatial analysis of molecular variance (SAMOVA) was performed in SAMOVA v1.0 (Dupanloup et al. 2002) and was used to identify possible geographic barriers to gene flow across the southern African landscape. This analysis identifies groups of populations that are maximally different to each other without requiring user defined groups as in AMOVA (Excoffier et al. 2005). In this case 'populations' (N=26) were defined as all the haplotypes represented at a unique latitude, longitude combination. The programme was run for 10000 iterations for each of 100 random initial conditions, testing all grouping options ranging from K=2 to K=26. K refers to the number of groups into which haplotypes are divided.

The use of landscape shape interpolation can be used to generate a graphical representation of the distribution of genetic diversity in species, in physical space (Miller 2005; Smit et al. 2007; Tolley et al. 2004). The programme Alleles in Space (AIS) (Miller 2005) produces a 3D surface plot where the Z-axis (height) represents the amount of genetic diversity at a point on the landscape. The analysis begins by first constructing a connectivity network among all of the sampling locations in the data set. Genetic distances between pairs of geographic localities are then calculated and plotted at the midpoint between each location. A 2D (x,y) co-ordinate grid is generated to include all points on the network. Surface heights (z values) are then calculated for each of the locations. The version of AIS used here generates an "all pairwise locations" connectivity network based on Delaunay triangulations (Brouns et al. 2003; Watson 1992). Here the calculation of the surface is based on the midpoints of edges derived from the Delaunay triangulation and surface heights log transformed distance-corrected genetic distances (Miller 2005).

iii. Intraspecific gene genealogy

Network construction methods were used to construct an intraspecific gene genealogy for the final dataset. More commonly, evolutionary relationships between species are estimated using phylogenetic tree inference methods; intra-specific gene genealogies, however, are not always accurately described using methods based on bifurcating trees. This is because ancestral alleles can persist and therefore co-exist with younger descendent alleles which, in a bifurcating tree, would result in close to zero branch lengths. Tree construction also implicitly models the evolution of a single gene (Pleines et al. 2009). The more complex the model, the more likely the analysis will recover incompatible branching events that cannot be incorporated onto a single tree.

The splits network is a more general type of phylogeny that can represent any combination of branch splits, incompatible or not. Networks do not assume a tree like ancestor-

University of Cape Town

descendent relationship and therefore do not force the data onto a tree (Huson and Bryant 2006). This type of genealogical reconstruction is significantly more appropriate to represent intra-species relationships using a relatively fast evolving marker such as the hypervariable D-loop region and are generally more 'sensitive' (Excoffier et al. 1992) than tree-based criteria to trace finer-scale population structure through space and time (Bermingham and Moritz 1998; Posada and Crandall 2001).

A network of relationships between individual sequences was constructed in SPLITSTREE v 4.8 (Huson and Bryant 2006). SPLITSTREE creates splits-graphs which are constructed according to the rules of split-decomposition theory (Bandelt and Dress 1992a, b, 1994). A split occurs when a single branch of a tree is removed, dividing the tree into two sets. If the intersection of two sets includes at least one empty subset, then the split is considered compatible. If all intersections are non-empty then the split is incompatible. It has been shown that tree searching algorithms are essentially searching for compatible collections of terminal nodes only (Buneman 1971). The splits decomposition allows for the inclusion of incompatible splits into the graph by relaxing the rules of compatibility. This generates a set of weakly compatible splits, that is, for every three splits there is one empty intersection. These splits can then be represented by a weighted splits graph or network (Holland and Moulton 2003).

UncorrectedP, NeighborNet and EqualAngle options were explored. The UncorrectedP method computes the distance between sequences as the proportion of positions at which two sequences differ. Neighbor-Net computes a set of incompatible splits from the data in the form of a given distance matrix to produce a set of splits that is circular (Huson and Bryant 2006). Two gap and all parsimoniously uninformative sites were removed from the analysis. The two gap sites at positions 152 and 317 were removed from the analysis as it was unclear if they were real gaps or an artifact of sequencing error. Clade allocation was tested by 1000 bootstrap replicates and 0.7 was used as a minimum cut off for confidence.

The spatial distribution of clades was plotted onto a map and patterns of geographic clustering were studied and are reported on. Network analysis allowed individuals to be grouped into populations defined by genetic rather than geographic relatedness.

iv. Population demographic history

The neutral theory of molecular evolution was proposed by Kimura (1968) and King and Jukes (1969) and further developed by Kimura and Takahata (1983). The basic premise of neutrality theory as applied to molecular evolution is that the majority of mutations that

University of Cape Town

achieve fixation are selectively neutral and as populations are finite, population differentiation is predominantly driven by random genetic drift. Within a relatively simple mathematical framework, neutral theory provides a powerful tool with which to investigate aspects of population demographic history (Nachman et al. 2003). Essentially, if the marker under investigation is known to be neutral then any deviations from neutrality are attributable to population size changes. This is explained in detail below.

Fundamental to employing neutrality tests, is that the molecular marker being used must be neutrally evolving. Tests for neutral sequence evolution fall into three categories, all of which assume that samples are taken from a single randomly mating population. The first and second category of tests attempt to distinguish demography from evolution but require multiple independent markers or congruent data from multiple species and are therefore not used here.

The third category of tests use the ratio of silent and synonymous substitutions within a marker and which can be calculated as Tajima's *D* and Fu's *Fs* statistics. Nonsynonymous mutations result in a change in phenotype which in turn affects how an organism will interact with its environment and is thus subject to selection pressure. Synonymous substitutions however are neutral, becoming fixed only by random genetic drift. It is therefore possible to compare the effect of positive and neutral selection by comparing the rates of synonymous and non synonymous substitution (Vandamme 2003).

Neutrality tests are based on one of three mutational models including the infinite sites model (ISM). The ISM is typically used to model DNA sequence evolution. The ISM assumes first that every new mutation generates a new allele and second, that each mutation happens at a unique locus (Nachman et al. 2003).There are two important predictions of the neutral model: (i) that the amount of variation within a population is a balance between gain from new mutations and loss or fixation due to genetic drift and (ii) that the distribution of alleles at equilibrium constitutes the neutral distribution. The neutral distribution is described as many low frequency alleles and a decreasing number of high frequency alleles (Nachman et al. 2003). Statistical tests of neutrality in essence, test for deviations from this distribution.

Statistical tests as implemented in Arlequin v3.1 (Excoffier et al. 2005) were used to test hypotheses of selective neutrality for the mitochondrial D-loop and to detect past population growth. Tajima's *D* (Tajima 1989), and Fu's *F*s (Fu 1997b) were calculated. Tajima's (1989) *D* statistic tests the null hypothesis that two estimates of the neutral mutation parameter, one

University of Cape Town

derived from the average number of pairwise nucleotide differences and the other based on the number of segregating sites in the sample, are equal. The statistic compares the observed number of polymorphic sites with the observed nucleotide heterozygosity in a sample and is negative for an excess of low frequency polymorphisms. Fu's (1997) *F*s statistic tests the probability of having no fewer than the number of observed alleles in the sample and tends to be negative when there is an excess of recent mutations (or rare alleles), and significant negative *F*s statistics suggest either that the marker is under selection or that the population under investigation has experienced demographic changes in the past. The null hypothesis (Ho) for this test is that the marker evolves neutrally. Rejection of Ho suggest either that the marker is not evolving neutrally or that the population under investigation has experienced demographic changes in the past. The significance of tests was determined using 10000 coalescent simulations without recombination, conditional on the pairwise number of differences.

Whilst recent changes in population size leave detectable patterns in the distribution of genetic differences within neutral markers (Harpending 1994; Rogers and Harpending 1992; Slatkin and Hudson 1991) a limitation of statistical estimates like Tajima's *D* and Fu's *F*s tests is that they cannot distinguish the effects of selection from departure from stable population size. Another set of tests must therefore be used for a more accurate interpretation of population history.

There are three main types of statistical approaches used for testing deviation from population equilibrium. Tests that use the frequency of segregating sites, tests of haplotype distribution and tests that use information from the mismatch distribution. Generally an expansion event leads to an increase in the number of haplotypes. This is due to an excess of singleton mutations. The majority of tests are highly influenced by this increase in the number of segregation sites (Rogers and Harpending 1992; Harpending 1994).

An assessment of the robustness of tests (Ramos-Onsins and Rozas 2002) revealed that the most powerful of these tests, particularly for non-recombining DNA, are Fu's *F*s which is based on haplotype distribution, R^2 based on the frequency of segregating sites and Harpending's raggedness index (*rg*) which uses the mismatch distribution. Fu's *F*s is better for larger samples (>50) while R^2 is more powerful with smaller samples or when the number of segregation sites is low (Ramos-Osins and Rozas 2002). R^2 tests for significant recent population growth, as indicated by low R^2 values.

University of Cape Town

The mismatch distribution of pairwise nucleotide differences among sequences (Rozas et al. 2003) is an estimate for the number of differences between each pair of sequences. The pairwise nucleotide differences of haplotypes drawn from a population should be unimodal if a population expansion has occurred otherwise it is multimodal, implying a stochastic population (Rogers and Harpending 1992). The smoothness of the mismatch distribution is measured as Harpending's raggedness index (*rg*) (Harpending 1994).This value is expected to be > 0.05 for multimodal or ragged distributions typical of populations at equilibrium, and < 0.05 for unimodal or smoother distributions typical of expanding populations. These tests are most powerful in identifying recent population size changes however it may be possible to identify older events by using only unique haplotypes in the analysis (Harpending et al. 1993).

Based on the clustering of haplotypes in the network reconstruction, the sample was divided into three clades which were treated as separate populations. These were labelled, Southern Clade (SoC), North eastern clade (NeC) and a North western Clade (NwC). Here I test the hypothesis of recent population growth from low-diversity founder populations within the different clades using two methods, Ramos-Onsins and Rozas (2002) R^2 statistic and pairwise mismatch distributions. Goodness of fit, i.e. the sum of square deviations (*SSD*) between the observed and the expected mismatch distribution (Schneider and Excoffier, 1999), and Harpending's raggedness index (*rg*) (Harpending 1994) were calculated to examine the statistical support of the expansion event. Significance of these tests were evaluated through 1000 coalescent simulations assuming a neutral infinite-sites model and constant population size. The simulations estimate the probability of obtaining values of the statistics (*rg*, R^2) equal to or lower than the observed using empirical sample sizes and theta (*theta = 2Ne mu*), which is estimated from the data.

University of Cape Town

Results

i. Quantifying diversity

After sequencing ambiguities were removed from the dataset the final sequence alignment comprised 473bp of the mitochondrial D loop. A total of 61 unique mitochondrial haplotypes were identified from the sample of 132 *P. ursinus* individuals. Alignment of all sequences resulted in 160 polymorphic sites, of which 104 were parsimony informative. Haplotype designations are reported in Appendix 5B.

The D-loop showed a high degree of sequence variation with 160 (~34%) variable sites, of which 123 represent transitions and 37 represent transversions. Of the 61 unique

haplotypes, 36 were scored only once. Descriptive statistics for each geographic lineage are tabled below (Table 5.1). The haplotype diversity for SgP (h= 0.957), is only slightly higher than for NgP (h =0.923). Nucleotide diversity per site was also slightly higher for SgP (0.047) than for NgP (0.038). SgP has a higher number of segregating (S) and singleton (s) sites than NgP although the ratio of S/s is equivalent for both geographic lineages.

	Full dataset	**Full dataset excluding individuals from identified mixing zones**	**SgP**	**NgP**
No. sequences	132	98	56	42
No haplotypes	61	43	28	15
No. segregating sites (S)	160	128	122	56
No. singleton sites (s)	56	27	24	12
ratio S/s	2.9	4.7	5.1	4.7
Nucleotide div (Pi)	0.08	0.08	0.05	0.04
Haplotype div (h)	0.98	0.97	0.96	0.90

Table 5.1- Diversity indices calculated for each of four subsamples. Estimates are based on the full dataset of 132 sequences, the dataset excluding individuals from mixing zones and datasets of the two geographic populations.

ii. Spatial diversity

The spatial interpolations produced a graphic representation of the distribution of genetic diversity (Fig 5.2) and reveals clustering of diversity at two hotspots. The first is across the distribution of NgP and the second is in a small part of the south-western Cape of South Africa. The peaks labelled in blue represent localities of NgP while red labels represent localities of the SgP. The area of highest diversity is seen in the Waterberg region of South Africa and further peaks are seen in Moremi, Mpumalanga, south east Namibia, the West coast and the Karoo.

The AMOVA analysis shows that the major component of variation corresponds to intra-population variance (46%), while differences among populations within groups accounted for 30% of the variation and only 24% among groups. These results indicate that isolation by distance accounts for very little of the genetic structure observed across the sample.

The SAMOVA analysis differs from AMOVA in that it should allow the user to capture geographic differentiation on a much smaller spatial scale and in doing so identify possible localised geographic impediments to gene flow between 'populations'. In this case the

University of Cape Town

SAMOVA analysis was uninformative as each increase of K removed one 'population' at a time from the whole sample, thus not producing informative clusters.

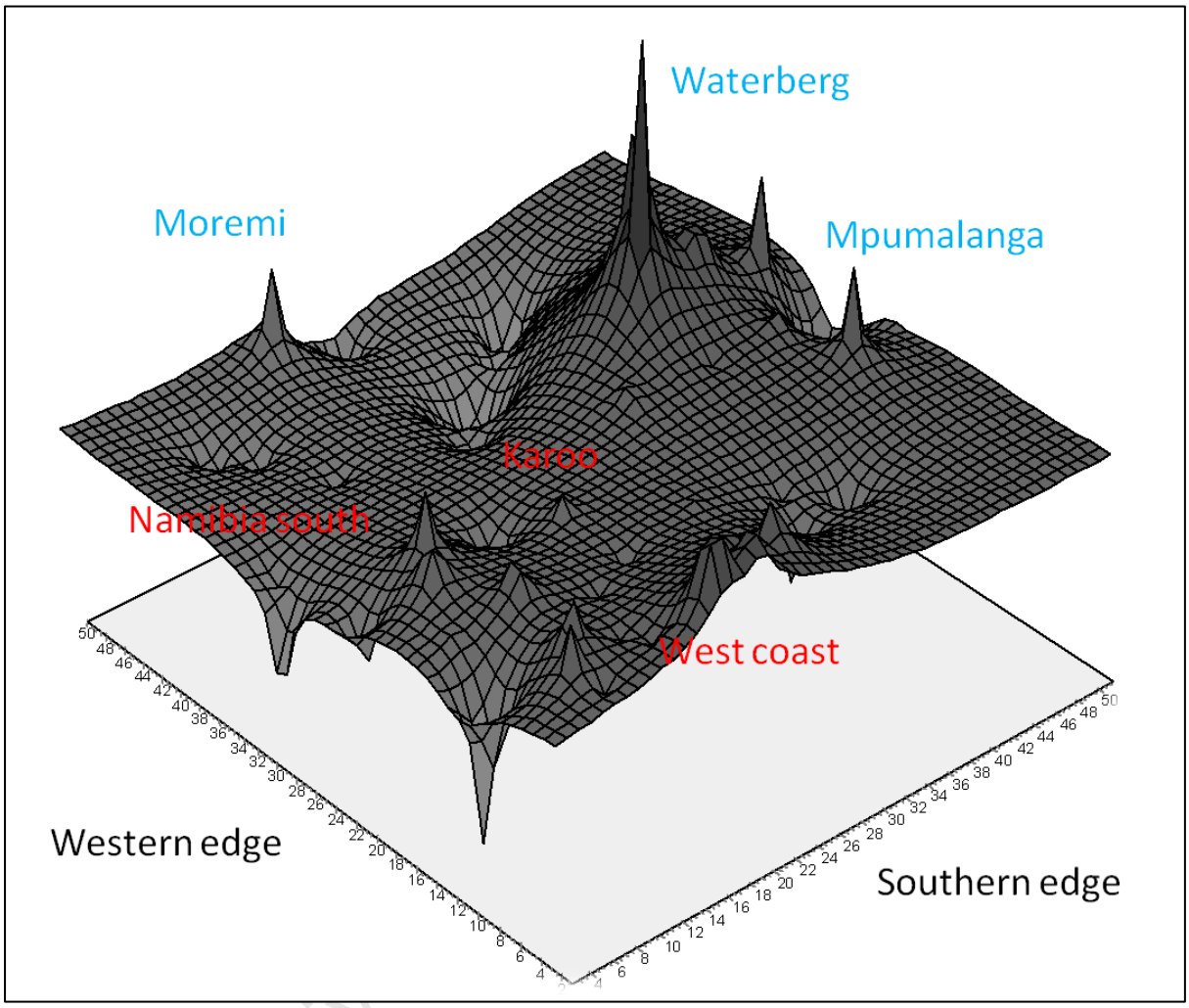

Figure 5.2- Spatial interpolation graph of chacma mitochondrial diversity. Blue labels indicate localities of NgP while labels in red indicate SgP localities. The Waterberg region has the highest peak indicating the greatest degree of genetic diversity per unit of distance.

iii. Intraspecific gene genealogy

These genetic clustering of haplotypes recovered in the network analysis (Fig 5.3) clearly reflects the geographic distribution of haplotypes as seen in Fig 5.4. The SPLITSTREE network reveals a major division within the sample which groups individual sequences into one of two major genetic lineages, northern D-loop (ND) and southern D-loop (SD) with further subdivisioning apparent (Fig 5.3). SD may be divided into 3 populations SoC 1, SoC2 and SoC3. However due to small sample numbers for SoC2 and SoC3, SoC1 is the only southern population considered for further analysis. It is therefore referred to as SoC.

ND is clearly divided into two populations: a North-western Clade (NwC) and a North-eastern clade (NeC).

The geographic distribution of clades and subclades is shown in Fig 5.4 below and the allocation of haplotypes is tabled in Appendix 5B. Boundaries have been drawn to show the general distribution area of each haplogroup. Individuals (outliers) that do not fit the general picture are shown on the map but are ignored when drawing the boundary lines, for example one NeC2 individual lies within the border of SoC2. These outliers most likely represent recent or rare dispersal events and are not included in the general pattern. NwC is undifferentiated and individuals are distributed in the north of the study area. Further sampling may show that this boundary is further north than estimated here. NeC is composed of three subclades in the north, west and east, and lies to the south east of NwC, concentrated on the east coast of South Africa. SoC1 hugs the South African west and south coasts and extends inland into the Drakensberg Mountains. SoC2 is distributed on the very arid edge of the Kalahari Desert along the Orange River. SoC3 extends the distribution of SD further north to meet NwC and NeC.

iv. Population demographic history

Results of the neutrality tests are tabled below together with the results of the tests for population expansion (Table 5.2). The shape of mismatch distribution for SoC is unimodal, and *rg* and *SSD* values are small and non-significant at the 0.05 level. R^2 is low and significant for this population. Together these results indicate that a sudden expansion model for SoC cannot be rejected. One would therefore expect to see highly negative values for Tajima's *D* and Fu's *Fs* as indicative of population size changes in a neutrally evolving marker (Fu 1997; Fu and Li 1993; Tajima 1989), however for the same population both are only slightly negative and not significantly so (Table 5.2). A similar pattern is identified for NwC which, despite having a unimodal mismatch distribution, non-significant *rg* and *SSD* values and a low and significant R^2 value, also support low Tajima's *D* and Fu's *Fs* values. These results may result from the statistical properties of neutrality tests identified in the study by Ramos-Onsins and Rozas (2002); the authors demonstrated that Tajima's *D* regularly performs poorly compared to Fu's *Fs* and R^2, and that for smaller samples i.e. < 50, R^2 outperforms Fu's *Fs*. Together these factors may contribute to the poor power of Tajima's *D* and Fu's *Fs* as applied to data set in this study. As applied to the NeC, tests for deviations from neutrality are all non-significant and a multimodal mismatch distribution suggests a population demographic history characterised by stability rather than expansion.

University of Cape Town

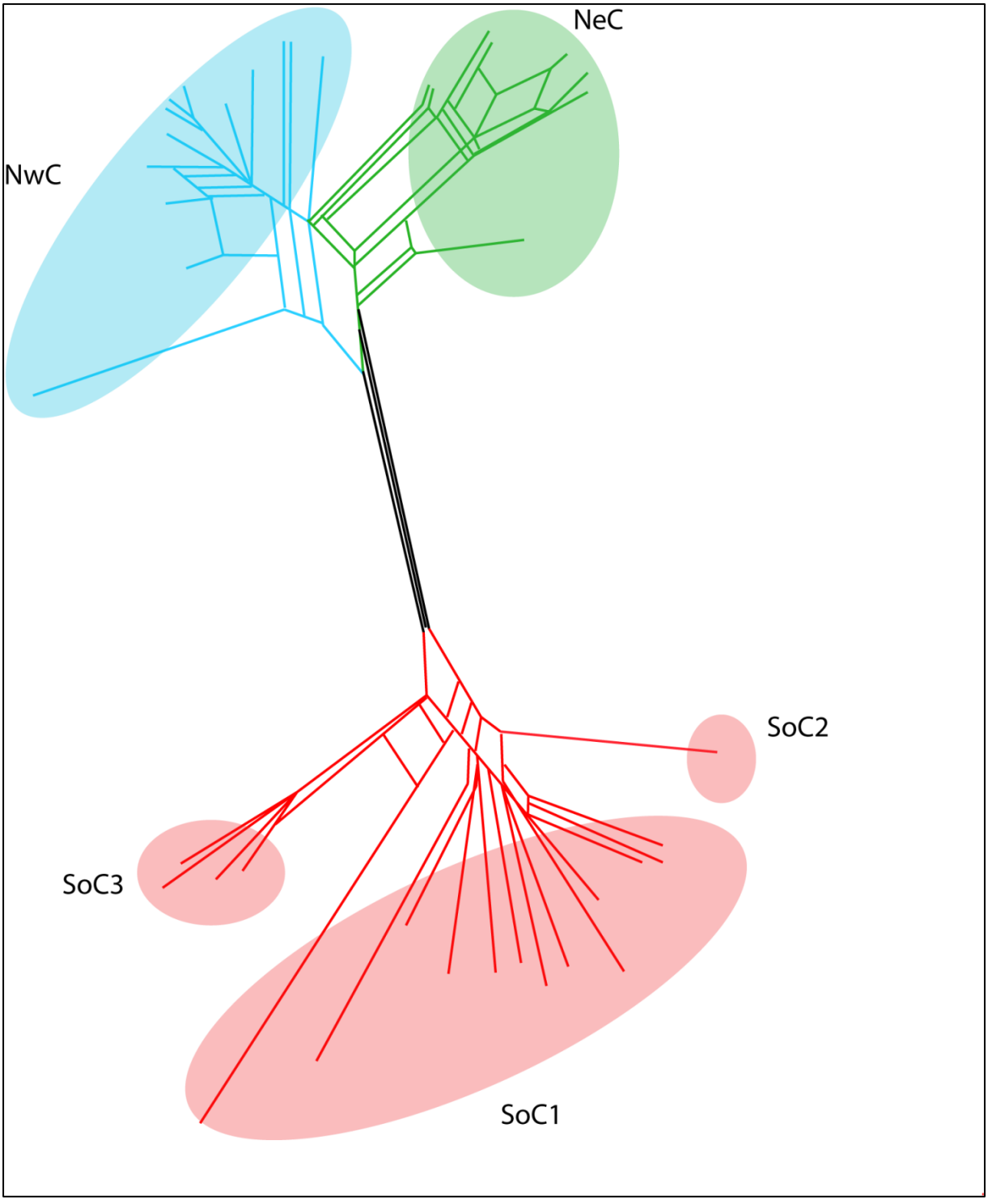

Figure 5.3- A splits-decomposition network of 132 D-loop sequences reveals a deep divergence in chacma baboons. Individuals are divided into one of two mitochondrial lineages, ND and SD.ND separates into two defined clades, NwC and NeC while SD is divided into 3 clades SoC1 which is a star cluster, SoC2 and SoC3

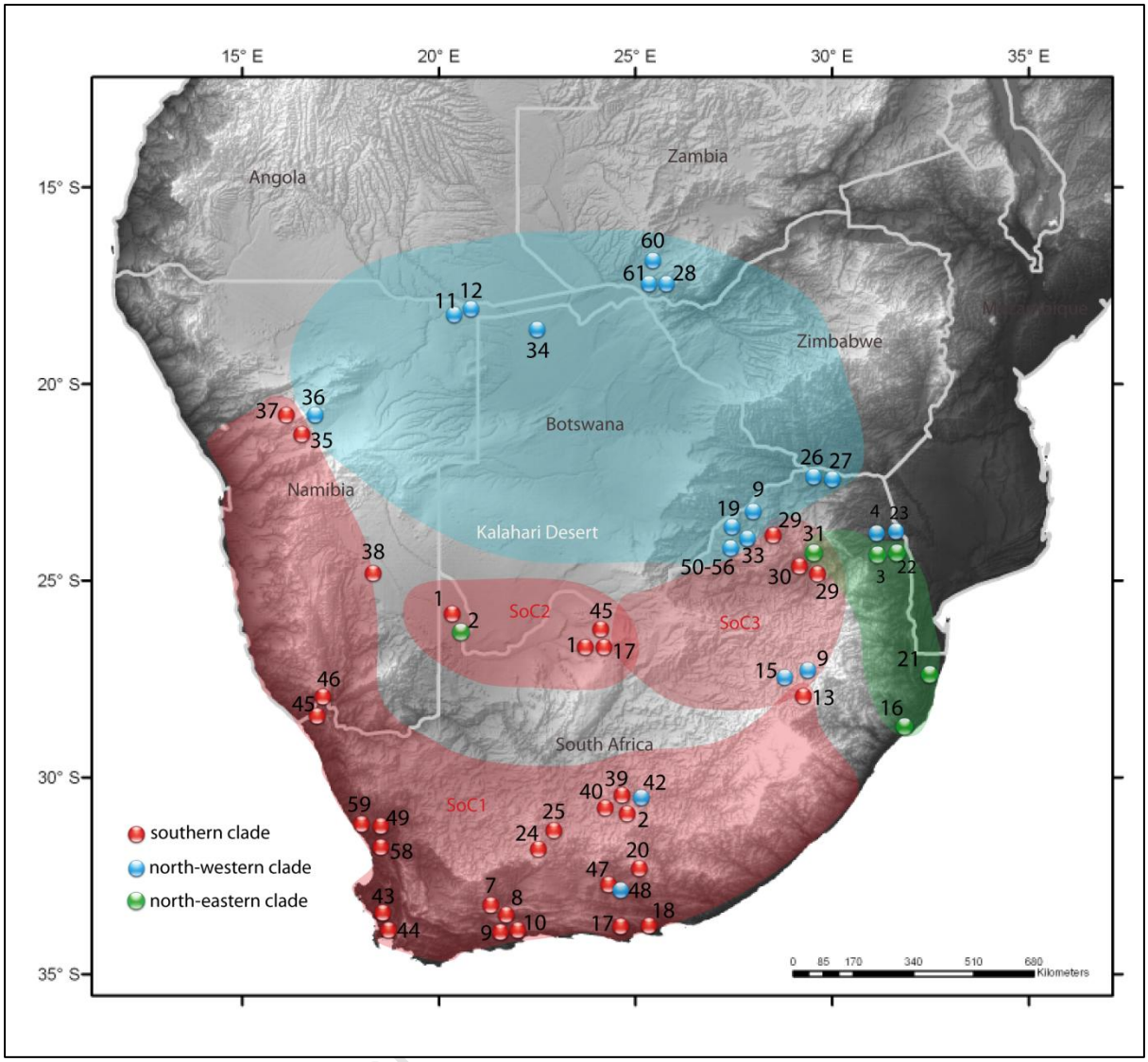

Figure 5.4- Map showing the distribution of samples representing unique D-loop haplotypes. Dots are coloured according to clade assignments from the network analysis which are tabled in Appendix 5B. Distributions of each of the clades and each of the subclades of SoC are also presented.

	SoC	NwC	NeC
Tajima's *D*	-0.67, *p*=0.28	-1.4, *p*=0.05	0.39, *p*=0.70
Fu's *Fs*	-0.28, *p*=0.52	-3.92, *p*=0.05	3.42, *p*=0.90
R^2	0.08, *p*=0.05	0.07, *p*=0.05	0.14, *p*=0.7
Raggedness index (*rg*)	0.01, *p*=0.9	0.01, *p*=0.7	0.04, *p*=0.2
SSD	0.01, *p*=0.34	0.005, *p*=0.80	0.03, *p*=0.16

Table 5.2- Diversity indices and neutrality estimates for SoC, NwC and NeC. P values are reported and significant values are highlighted in red. Significance is determined at the 0.05 level.

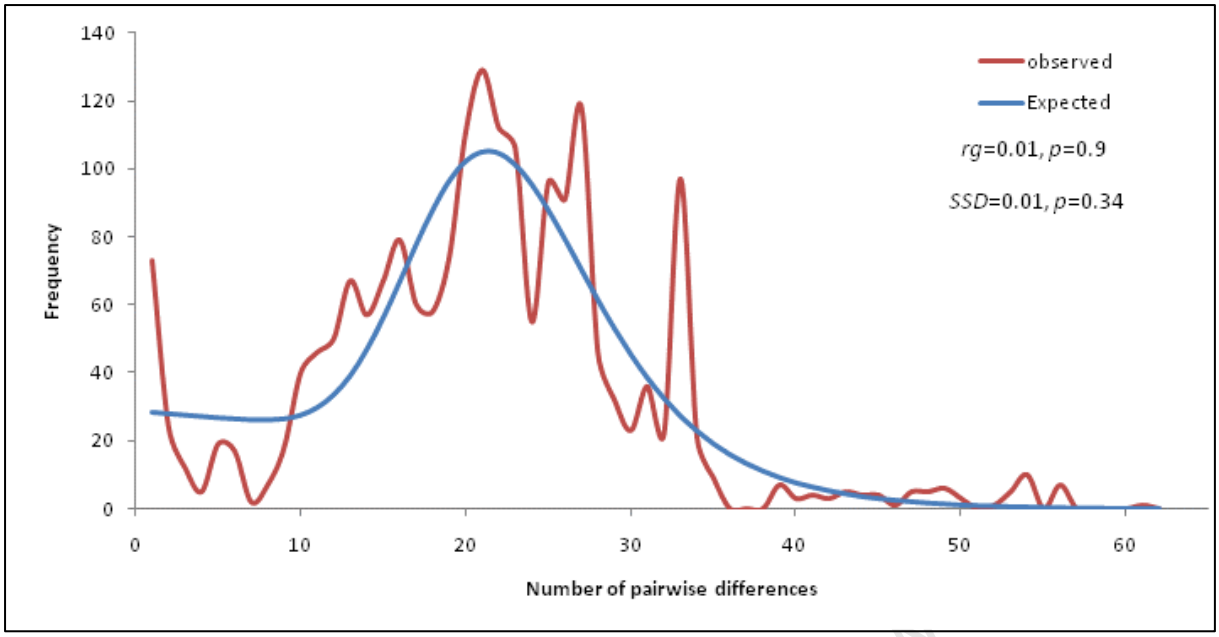

Figure 5.5a- The pairwise mismatch distribution of SoC plotted together with the curve of model expectations for a population under expansion.

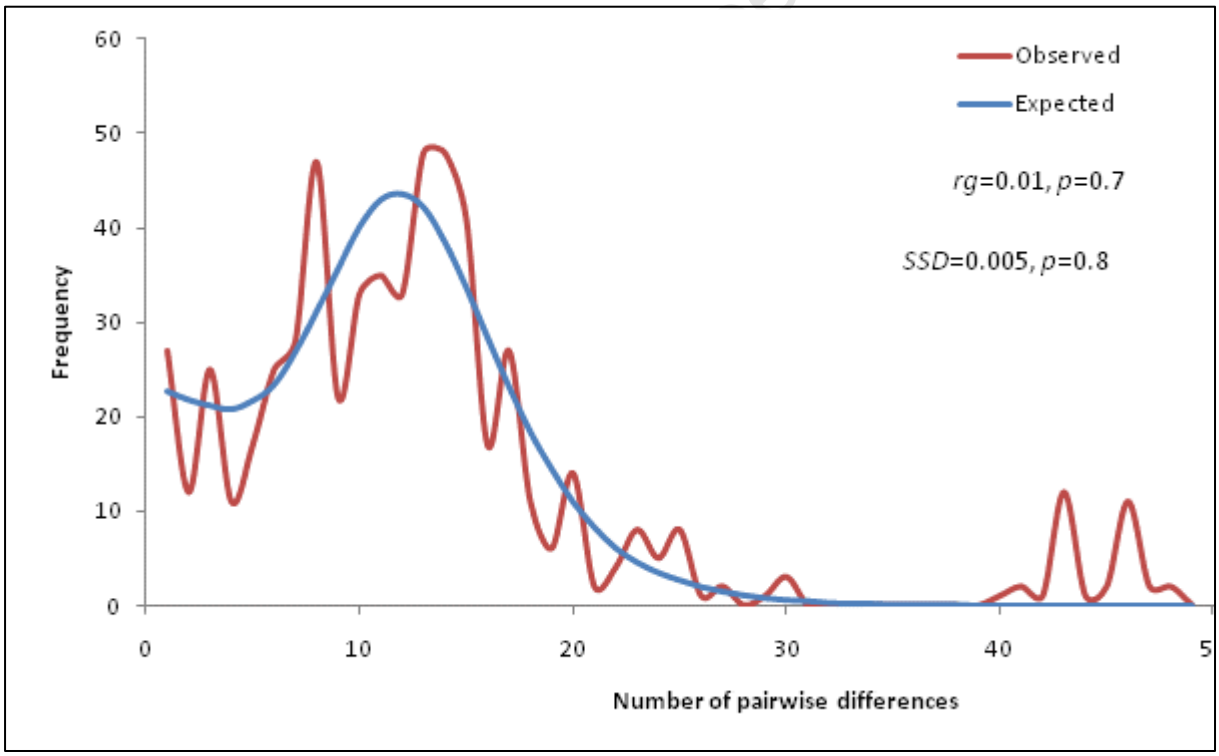

Figure 5.5b- The pairwise mismatch distribution of NwC plotted together with the curve of model expectations for a population under expansion.

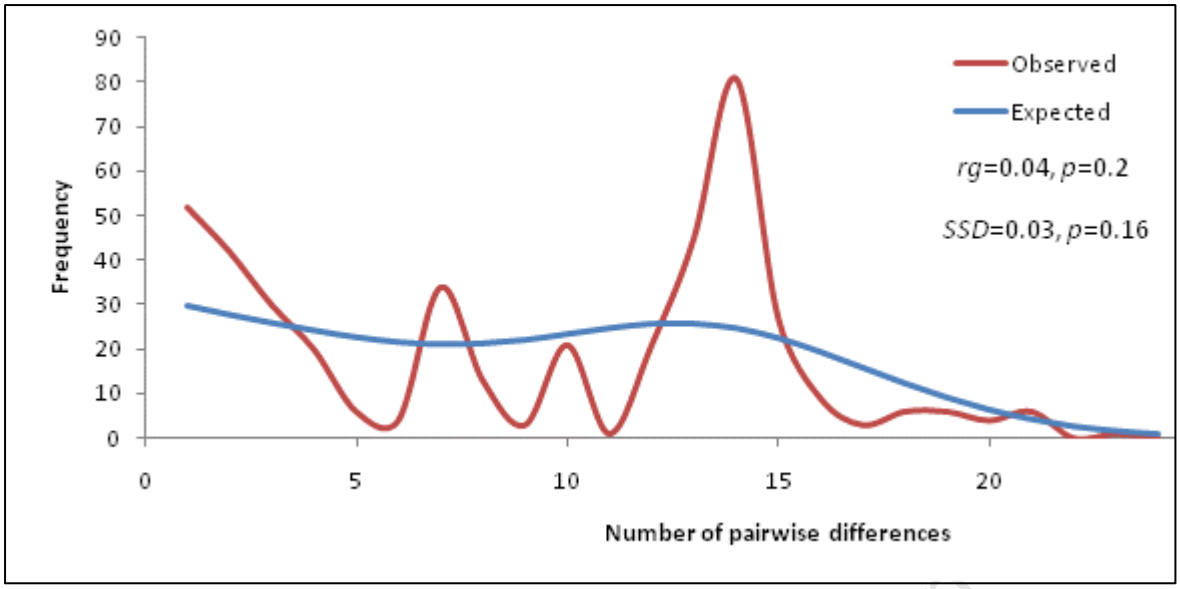

Figure 5.5c- The pairwise mismatch distribution of NeC plotted together with the curve of model expectations for a population under expansion.

University of Cape Town

Discussion

The results presented in this chapter allow for the fine-scale re-construction of chacma population history contextualised within the climate and landscape history of southern Africa. This is made possible through a phylogeographic analysis of the mitochondrial D-loop. The D-loop is a relatively fast evolving marker (Excoffier and Yang 1999; Pesole et al. 1992; Vigilant 2009; Vigilant 1986), substitutions are therefore less likely to be fixed for long periods of time and substitution patterns related to particular demographic events are increasingly lost as more time passes since the event occurred. This marker then, is used as a window into the recent evolutionary past (Vigilant 1986; Excoffier 1999; Pesole et al. 1992; Belay and Mori 2006).

A network analysis of the full sample (n=132) divides haplotypes into one of two genetic clusters, SD and ND. This division further reflects recent published findings by Sithaldeen et al. (2009) and Zinner et al. (2009). A north-south division within chacma baboons is also recovered in analyses of the mitochondrial *cytochrome-b* gene by Keller et al. (2010). These studies, together with findings reported in this chapter, provide unequivocal support for a major divergence in chacma baboons between 1.13 - 2.36 Ma (Sithaldeen et al. 2009; Zinner et al. 2009), most likely as a result of the Kalahari Desert expansion (Sithaldeen et al. 2009). That evidence for this vicariance event is retained in the evolutionary history of the chacma D-loop, suggests that northern and southern populations were isolated until the

recent past; further support for this finding is the limited sharing of D-loop haplotypes between previously isolated northern and southern haplotype localities.

The continuous distribution of chacma baboons today is possibly due to population expansions that are suggested, but not confirmed by neutrality tests for population growth. Both the excess of singleton sites and mismatch distributions for SoC and NwC suggest that these populations have recently experienced some degree of demographic growth. Furthermore, in between the northern and southern diversity hotspots shown in the landscape interpolation plot (Fig 5.2), the diversity landscape is relatively flat, suggesting that some of these areas may represent a later expansion fringe (Bernatchez and Wilson 1998; Hewitt 2000; Hewitt 1996; Ibrahim et al. 1996). Hewitt (2000) proposed that the leading edge of an expanding population is usually achieved through long range dispersals; consequently a series of founding events at the leading edge of expansion might leave a signature of reduced allelic and haplotype diversity. The model predicts low genetic diversity at the rapidly colonized leading edge of an expanding population (Hewitt 2000) and this can be seen between the genetic diversity hotspots identified in this study.

The spatial interpolation analysis also reveals a greater degree of diversity concentrated in the northern (ND) regions of the chacma distribution. Both Zinner et al. (2009) and Keller et al. (2010) show a closer mitochondrial relationship between northern chacma populations and yellow baboons from Zambia than between northern and southern chacma populations. It is therefore likely that past gene flow has contributed to contemporary levels of genetic diversity in the ND gene pool. The highest concentration of genetic diversity is seen in the Waterberg region of South Africa, a pattern that is likely the result of lineage contact and some degree of gene flow between all three chacma clades (SoC, NeC, NwC).

A second diversity hotspot was identified in the south-western Cape. Unlike the majority of southern Africa, the south-western Cape region of Africa is a winter rainfall zone dominated by *fynbos* vegetation. The fynbos biome is characterised by extremely high levels of diversity and endemism (Cowling et al. 1992; Galley and Linder 2006 ; Goldblatt and Manning 2002). It has been proposed that this region represented a glacial refuge characterised by relative stability for many species during the Pleistocene (Barraclough 2006; Dynesius and Jansson 2000; Meadows and Baxter 1999). For example, although the winter rainfall region was cooler than much of the rest of South Africa during the last glacial maximum, it was also a lot wetter (Meadows and Baxter 1999). As such this region could have served as a refuge for species that may have been cold tolerant but unable to withstand the intense aridity

University of Cape Town

expanding in central and southern Africa at the time. Additionally the south-western Cape is home to a diverse plant community which includes a number of nutritious, carbohydrate rich geophytes (see Cowling et al. 1992) which baboons are known to exploit (Devore and Hall 1965). These resources are also thought to have provided a rich source of calories for hominid species that took refuge in this region during glacial periods (Rector and Reed 2010). It is therefore possible that the expansion of the Kalahari Desert, which fragmented chacma baboons at ~1.6 Ma, coincided with a period of population contraction into the south-western Cape which provided a glacial refuge for the southern population. This event also likely resulted in reduced dispersal and gene flow to the north, driven both by environmental barriers and declining population size, as baboons responded to habitat stress brought on by the harsh glacial conditions of the Pleistocene.

It is clear that periods of reduced gene flow driven by population contraction contributed to the divergence of chacma populations both genetically and morphologically. This proposal is supported by the overlap of genetic and morphological groupings as reported in chapter 4. SD in this analysis would be analogous to Southern Lineage in chapter 4 and so represent *P.u. ursinus* haplotypes. ND in this analysis would be analogous to Northern Lineage in chapter 4 and so represent *P.u griseipes* haplotypes. In fact, the geographic distribution of further haplogroups representing ruacana and orientalis phenotypes in chapter 4 suggest that these groups may also represent a history of repeated population contractions. I propose that, as has been observed for several diverse lineages in southern Africa (Matthee and Flemming 2002; Matthee and Robinson 1996; Russo et al. 2010; Smit et al. 2010; Swart et al. 2009; Tolley et al. 2006), throughout the Pleistocene, climate driven landscape changes overlaid on the topographical features of southern Africa repeatedly forced chacma baboons to seek more suitable habitats. As the result of localized extinctions and population retractions into refugia, gene flow across southern Africa would have been episodic. Refugial groups would therefore have differentiated both genetically and phenotypically, most likely under a regime of genetic drift (Ackermann and Bishop 2010), until a return to conditions that favoured population expansions. These data provide support for Jolly's (2001) hypothesis that population contractions and expansions can lead to a significant degree of phenotypic diversity in baboons, even over fairly short evolutionary time periods.

Chacma baboons are ecological generalists capable of adapting to extreme habitat stress. For example, Namibian baboons are able to cope with intense aridity (18 - 85mm per annum) (Cowlishaw and Davies 1997) and have also been documented as fairly cold tolerant e.g. at Suikerbosrand which is a high altitude mountain range which can experience frost for more than 150 days per year (Anderson 1982) with temperatures routinely falling

University of Cape Town

below 5^{O}C in the winter (Segal 2008). A pattern of glacial refuge use in this species is therefore unexpected. Studies of baboon ecology, however, confirm that troops can be restricted by certain ecological parameters related to an individual's need to perform three important tasks on any given day: foraging, resting and socializing (Dunbar 1992). In an ideal situation an individual baboon will be able to budget all three in a way that still optimizes their reproductive success; however, in times of habitat stress, both the individual and the troop must choose to sacrifice resting or socializing time in order to have sufficient daylight hours (Hill et al. 2003) to gather their required daily calories (Alberts et al. 2005). If habitat stress is taken to an extreme and the troop reaches a critical threshold whereby they do not have sufficient resting or socializing time, they will be forced to fission or move (Altmann et al. 2002) It is therefore possible that given the harsh conditions of the Pleistocene in Africa (Thunell and Williams 1983; Shackleton et al.1984) chacma baboons may have experienced repeated population shifts, which have, in turn, shaped genetic structure within the species.

At a finer geographic scale, the network analysis revealed three further sub-clades (SoC, NeC and NwC) in the data set. These results further support the findings of Sithaldeen et al. (2009), Keller et al. (2010) in which 3 chacma clades -- southern, northern and eastern -- were identified. Based on the geographic proximity and genetic similarity of NeC and NwC, and the relatively broad distribution of NwC when compared to NeC, I propose that NeC represents a clade that once fragmented from NwC. Hewitt's (2000) leading edge hypothesis provides a model for expansion and subsequent diversification within lineages that may apply here. Hewitt proposed that once populations on the leading expansion edge fill up new habitats, it becomes more difficult for further dispersal out of the refugial groups; this slows down the rate and extent of the range expansion, thereby maintaining a spatial distance between the expanded edge and the interior populations, and in time leading to population genetic differentiation via drift (Hewitt 2000).

The date of divergence for NeC is estimated to ~400kya (chapter 4), concurrent with, and possibly correlated to, a period of extremely warm and moist climates, known as a hyperthermal, which lasted from 420-360 kya (Howard 1997; Muller and MacDonald 1997; Olson and Hearty 2009 ; Raynaud et al. 2005). The warming trend would have stimulated the expansion of suitable habitats and allowed baboons to extend their range high up into the Great Escarpment to the east and taking advantage of newly available niches in the highlands of Mpumalanga. Hyperthermal periods are also correlated to periods of forest expansion in South Africa (Deacon and Lancaster 1988) such that forest cover within southern Africa occupied much greater areas during those periods preceding and following

University of Cape Town

glacial maxima (van Zinderen Bakker 1978), with forest distributed from the Eastern Cape into Mpumalanga and extensive forest cover extending from the Eastern Cape into Gauteng (Cooke 1962). These forest expansions may have acted as semi-permeable genetic barriers between baboon populations that had dispersed to the eastern highlands and their parent populations, eventually differentiating and giving rise to NeC.

Despite gene flow between lineages, between populations and across clades, the network analysis suggests limited sharing of mitochondrial haplotypes between geographic localities. The distributions of these clades are shown in Fig 5.5. Judging by eye NwC appears to be realtively undifferentiated when compared to NeC and SoC. The apparent lack of subdivisioning in NwC could suggest either a late expansion event for this population or alternatively a long uninterrupted history of continuous distribution and gene flow within the population. NeC lies to the south east of NwC concentrated on the east coast of South Africa within and is differentiated into 3 subclades to the north, west and east. SoC is composed of three subclades which are discussed below.

Southern African temperature fluctuations and associated wet/dry cycles have played a major role in shaping genetic structure in many species (Arctander et al. 1999; Brain 1981; Brouat et al. 2009; DeMenocal 2004; Ducroz et al. 1998; Eggert et al. 2002; Herron et al. 2005b; Lorenzen et al. 2009; Matthee and Flemming 2002; Matthee and Robinson 1996; Matthee and Robinson 1997; Nyakaana et al. 2002; Russo et al. 2010; Swart et al. 2009; Tolley et al. 2010; Tolley et al. 2006; Van Hooft et al. 2002) and may have also contributed to genetic structuring among chacma baboon populations. The first SoC subclade, SoC1, extends from central Namibia south to the Cape Peninsula and north to the Drakensberg-Maluti mountain range. In a study of Smith's red rock rabbit (*Pronolagus rupestris)*, Matthee and Robinson (1996) identified a clade with a similar distribution to SoC1 and which appears bound to the mountain ranges and abutting regions of the Great Escarpment of South Africa. Matthee and Robinson (1996) describe these clades as 'high alpine' communities which represent a relict population that had taken refuge in the mountains and from which the species later expanded. Here I propose that SoC1, which is the dominant group in SD, shared a similar history. This group likely represents the dispersal of the relict population that had taken refuge in the south-western Cape and later expanded out of the galcial refuge along the coastal Cape mountains towards the north and east. A similar pattern has been observed in the rock hyrax (*Procavia capensis)* with the distribution of the south-eastern clade again linked to the Great Escarpment mountain range (Prinsloo and Robinson 1992)).

University of Cape Town

In their studies Matthee and Robinson (1996) and Prinsloo and Robinson (1992)recover additional clades associated with the Kuruman Hills near Augrabies in South Africa. The distribution of these clades closely approximates the distribution of SoC2. Matthee and Robinson (1996) propose that a high-lying region, the Kuruman Hills, could have served as a refugium from which subsequent recolonization within the marginal habitat of the northen Cape plains took place. SoC2 may represent a colonizing group which dispersed away from the expanding SoC1 during favourable condition and then during a later cold cycle took refuge in these mountains. The distribution of these haplogroups within SoC, and their relationship to the landscape, further support the hypothesis of glacial refuge use in baboons.

Conclusions

Together with the results presented in Chapter 4, the findings discussed here suggest that current genetic structure within chacma baboons is the result of complex interactions between regional climatic and geographic factors together with behavioural factors central to the success of *Papio* in Africa. Environmental change, brought on by climate fluctuations in the Plio-Pleistocene, together with species behavioural responses, have resulted in clear mitochondrial signatures of lineage diversification within chacma baboons across southern Africa. An initial diversification into two mitochondrial lineages was most likely driven by population contractions away from a region of intense aridification in central southern Africa, forcing one lineage to seek refuge in the south-western Cape. Climatic amelioration allowed for the expansion of both lineages and the colonisation of new habitats, possibly driving early structure in the expanding lineages. The return to cold and dry conditions in the next cold cycle drove these newly dispersed groups into new environmental refugia. Localised extinctions led to localised genetic bottlenecks and refugial populations acted as sources for later expansions thereby generating a highly differentiated but geographically continuous baboon lineage.

University of Cape Town

Appendix 5A- A sample of phylogeographic studies illustrates the diversity of lineages that owe their current distribution and genetic diversity to Plio-pleistocene age climate driven landscape change.

Reference	Species	Study conclusions
Mathee and Robinson 1997	springhare (*Pedetes capensis*)	Today woodlands form a genetic barrier between springhare ranges however these outcrops are too recent to account for the accumulated genetic differences that distinguish species within the lineage. Instead, it seems that the current pattern of structuring within the genus is due to relatively recent range expansions from smaller source populations that were previously isolated in Pleistocene refugia across southern Africa .
Ducroz et al. 1998	grass rat (genus *Arvicanthis*)	The genus experiences several cladogenic events during the late Pliocene, when the extension of open habitats, may have driven speciation in this savannah-dwelling genus.
Arctander et al. 1999	hartebeest (*Alcelaphus buselaphus*), topi (*Damaliscus lunatus*), and the blue wildebeest (*Connochaetes taurinus*)	The phylogenetic patterns found in these three bovid lineages suggest a scenario of climate driven localised extinctions as species were forced into refugia through habitat change.
Eggert et al. 2002	African elephant (*Loxodonta africana*)	The complex phylogeographic patterns detected in African elephants are thought to be the result of repeated continental-scale landscape changes as climates fluctuated over the last 5-6 myr.
Matthee and Flemming 2002	rock lizard (*Agama atra*)	Three distinct geographical assemblages within *Agama atra* result from population isolations and dispersals in response to natural climatic changes during the past three million years.
Nyakaana et al. 2002	African elephant s (*Loxodonta africana*)	The distribution of genetic variation within and between African savannah elephants are attributed to the use of Pleistocene refugia followed by population admixture due to later population expansions.
Van Hooft et al. 2002	African buffalo (*Syncerus caffer*)	Results indicate a history of late middle to late Pleistocene population expansions in Cape buffalo which may be related to climate driven vegetation changes.
de Menocal 2004	fossil African mammals	An analysis of fossil databases shows speciation events within African mammals roughly corresponding with Pliocene-Pleistocene climate change and vegetation shifts at ~2.9-2.4 Ma and after 1.8 Ma. These intervals also correspond to major events in early hominid evolution, including the emergence of the genus *Homo*.

University of Cape Town

Appendix 5A- A sample of phylogeographic studies illustrates the diversity of lineages that owe their current distribution and genetic diversity to Plio-pleistocene age climate driven landscape change.

Reference	Species	Findings
Brouat et al. 2009	Guinea multimammate mouse (*Mastomys erythroleucus*)	Vicariance events within this species are related to the fragmentation of savanna habitats during the Pleistocene and recent demographic expansions probably occurred during arid phases of the Holocene with the southward expansion of savannas.
Herron et al. 2005	African ground squirrel (*Xerus inauris*)	The distributions of African ground squirrel populations are concordant with divergences within and disjunctions between a number of other mammalian taxa, which have been interpreted as results of Plio-Pleistocene climate cycles.
Lorenzen et al. 2009	common eland (*Taurotragus oryx*)	The data support the hypothesis of Pleistocene refugia for eland occurring in East and southern Africa and is in agreement with palynological, palaeovegetation and fossil studies.
Swart et al. 2009	southern rock lizard (*Agama atra*)	Southern rock lizards exhibit at least four distinct genetic provinces within the Cape Floristic region, and the dates of separation among the clades coincide well with the documented Pleistocene climate fluctuations.
Russo et al. 2010	Namaqua rock mouse (*Micaelamys namaquensis*)	The radiation of the Namaqua rock mouse occurred during the Pliocene and Pleistocene coinciding with major periods of aridification and the expansion of savannah habitats. The major diversification within lineages occurred during the Pleistocene.
Tolley et al. 2010	clicking stream frog (*Strongylopus grayii*)	Phylogeographic analyses show climatic transitions as generating effective barriers to gene flow resulting in vicariant speciation within this genus.

University of Cape Town

Appendix 5B- A table cataloguing sample information and genetic and geographic groupings of the phylogeographic analysis. These haplotype numbers are for mitochondrial D-loop sequences.

Sequence no.	Locality	Haplotype/ MapID	longitude	latitude	Network Clade	Geographic Population	Geographic area
1	Augrabies, ZA	1	20.320125	-28.547017	SoC2	SgP	Orange River
2	Augrabies, ZA	1	20.320125	-28.547017	SoC2	SgP	Orange River
3	Augrabies, ZA	1	20.320125	-28.547017	SoC2	SgP	Orange River
4	Augrabies, ZA	2	20.320125	-28.547017	NeC2	SgP	Orange River
5	Augrabies, ZA	2	20.320125	-28.547017	NeC2	SgP	Orange River
6	Augrabies, ZA	1	20.320125	-28.547017	SoC2	SgP	Orange River
7	Augrabies, ZA	1	20.320125	-28.547017	SoC2	SgP	Orange River
8	BlydeRiver, ZA	3	30.846517	-24.861211	NeC3	NgP	Mpumalanga
9	BlydeRiver, ZA	3	30.846517	-24.861211	NeC3	NgP	Mpumalanga
10	BlydeRiver, ZA	3	30.846517	-24.861211	NeC3	NgP	Mpumalanga
11	BlydeRiver, ZA	3	30.846517	-24.861211	NeC3	NgP	Mpumalanga
12	BlydeRiver, ZA	3	30.846517	-24.861211	NeC3	NgP	Mpumalanga
13	BlydeRiver, ZA	4	30.846514	-24.861211	NeC3	NgP	Mpumalanga
14	BlydeRiver, ZA	4	30.846514	-24.861211	NeC3	NgP	Mpumalanga
15	BlydeRiver, ZA	4	30.846514	-24.861211	NeC3	NgP	Mpumalanga
16	BlydeRiver, ZA	4	30.846514	-24.861211	NeC3	NgP	Mpumalanga
17	BlydeRiver, ZA	3	30.846517	-24.861211	NeC3	NgP	Mpumalanga
18	BlydeRiver, ZA	5	30.846514	-24.861211	NeC3	NgP	Mpumalanga
19	Calitzdorp, ZA	6	21.470617	-34.024325	SoC1	SgP	Cape
20	Calitzdorp, ZA	7	21.470617	-34.024325	SoC1	SgP	Cape
21	Calitzdorp, ZA	8	21.470617	-34.024325	SoC1	SgP	Cape
22	Calitzdorp, ZA	9	21.470617	-34.024325	NwC	SgP	Cape
23	Calitzdorp, ZA	6	21.470617	-34.024325	SoC1	SgP	Cape
24	Calitzdorp, ZA	6	21.470617	-34.024325	SoC1	SgP	Cape

University of Cape Town

Appendix 5B- A table cataloguing sample information and genetic and geographic groupings of the phylogeographic analysis. These haplotype numbers are for mitochondrial D-loop sequences.

Sequence no.	Locality	Haplotype/ MapID	longitude	latitude	Network Clade	Geographic Population	Geographic area
25	Calitzdorp, ZA	10	21.470617	-34.024325	SoC1	SgP	Cape
26	Caprivi, NA	11	21.604333	-18.152544	NwC	NgP	Okavango
27	Caprivi, NA	11	21.604333	-18.152544	NwC	NgP	Okavango
28	Caprivi, NA	11	21.604333	-18.152544	NwC	NgP	Okavango
29	Caprivi, NA	12	21.604333	-18.152544	NwC	NgP	Okavango
30	Caprivi, NA	11	21.604333	-18.152544	NwC	NgP	Okavango
31	Caprivi, NA	11	21.604333	-18.152544	NwC	NgP	Okavango
32	Caprivi, NA	11	21.604333	-18.152544	NwC	NgP	Okavango
33	Drakensberg, ZA	13	28.961906	-28.538458	SoC1	Mx	East coast
34	Drakensberg, ZA	13	28.961906	-28.538458	SoC1	Mx	East coast
35	Drakensberg, ZA	14	28.961906	-28.538458	SoC1	Mx	East coast
36	Drakensberg, ZA	9	28.961906	-28.538458	NwC	Mx	East coast
37	Drakensberg, ZA	15	28.961906	-28.538458	NL	Mx	East coast
38	Hluhluwe NP, ZA	16	32.269378	-28.008656	NeC2	Mx	East coast
39	Hluhluwe NP, ZA	16	32.269378	-28.008656	NeC2	Mx	East coast
40	Hluhluwe NP, ZA	16	32.269378	-28.008656	NeC2	Mx	East coast
41	Hluhluwe NP, ZA	16	32.269378	-28.008656	NeC2	Mx	East coast
42	Hluhluwe NP, ZA	16	32.269378	-28.008656	NeC2	Mx	East coast
43	Hluhluwe NP, ZA	16	32.269378	-28.008656	NeC2	Mx	East coast
44	Kimberley, ZA	17	24.528669	-27.972458	SoC1	SgP	Kimberley
45	Kimberley, ZA	45	24.528669	-27.972458	SoC1	SgP	Kimberley
46	Kimberley, ZA	19	24.528669	-27.972458	NwC	SgP	Kimberley
47	Kimberley, ZA	19	24.528669	-27.972458	NwC	SgP	Kimberley
48	Kimberley, ZA	20	24.528669	-27.972458	SoC3	SgP	Kimberley
49	Kimberley, ZA	20	24.528669	-27.972458	SoC3	SgP	Kimberley

University of Cape Town

Appendix 5B- A table cataloguing sample information and genetic and geographic groupings of the phylogeographic analysis. These haplotype numbers are for mitochondrial D-loop sequences.

Sequence no.	Locality	Haplotype/ MapID	longitude	latitude	Network Clade	Geographic Population	Geographic area
50	Kimberley, ZA	20	24.528669	-27.972458	SoC3	SgP	Kimberley
51	Kimberley, ZA	20	24.528669	-27.972458	SoC3	SgP	Kimberley
52	Kimberley, ZA	1	24.035161	-28.186036	SoC2	SgP	Kimberley
53	Kimberley, ZA	1	24.035161	-28.186036	SoC2	SgP	Kimberley
54	Kimberley, ZA	20	24.528669	-27.972458	SoC3	SgP	Kimberley
55	KosiBay, ZA	21	32.867289	-26.906211	NeC2	Mx	East coast
56	Kruger NP, ZA	22	31.638253	-24.486086	NeC3	NgP	Kruger
57	Kruger NP, ZA	22	31.638253	-24.486086	NeC3	NgP	Kruger
58	Kruger NP, ZA	22	31.638253	-24.486086	NeC3	NgP	Kruger
59	Kruger NP, ZA	22	31.638253	-24.486086	NeC3	NgP	Kruger
60	Kruger NP, ZA	23	31.638253	-24.486086	NL	NgP	Kruger
61	Leeu Gamka, ZA	24	21.715675	-33.343933	SoC1	SgP	Karoo
62	Leeu Gamka, ZA	25	21.715675	-33.343933	SoC1	SgP	Karoo
63	Leeu Gamka, ZA	24	21.715675	-33.343933	SoC1	SgP	Karoo
64	Limpopo, ZA	26	30.039753	-22.349506	NwC	NgP	Limpopo
65	Limpopo, ZA	27	30.039753	-22.349506	NwC	NgP	Limpopo
66	Livingstone, ZM	28	25.855269	-17.853531	NwC	NgP	Okavango
67	Livingstone, ZM	28	25.855269	-17.853531	NwC	NgP	Okavango
68	Livingstone, ZM	28	25.855269	-17.853531	NwC	NgP	Okavango
69	Loskop Dam, ZA	29	28.210792	-24.435117	SoC3	Mx	Loskop
70	Loskop Dam, ZA	30	28.210792	-24.435117	SoC3	Mx	Loskop
71	Loskop Dam, ZA	30	28.210792	-24.435117	SoC3	Mx	Loskop
72	Loskop Dam, ZA	31	28.210792	-24.435117	NeC1	Mx	Loskop
73	Loskop Dam, ZA	31	28.210792	-24.435117	NeC1	Mx	Loskop
74	Loskop Dam, ZA	31	28.210792	-24.435117	NeC1	Mx	Loskop

University of Cape Town

Appendix 5B- A table cataloguing sample information and genetic and geographic groupings of the phylogeographic analysis. These haplotype numbers are for mitochondrial D-loop sequences.

Sequence no.	Locality	Haplotype/ MapID	longitude	latitude	Network Clade	Geographic Population	Geographic area
75	Loskop Dam, ZA	30	28.210792	-24.435117	SoC3	Mx	Loskop
76	Magaliesberg, ZA	32	27.317314	-24.947225	NwC	Mx	Loskop
77	Magaliesberg, ZA	33	27.317314	-24.947225	NwC	Mx	Loskop
78	Moremi NP, BW	34	23.103469	-19.464772	NwC	NgP	Moremi
79	Moremi NP, BW	34	23.103469	-19.464772	NwC	NgP	Moremi
80	Moremi NP, BW	34	23.103469	-19.464772	NwC	NgP	Moremi
81	Moremi NP, BW	34	23.103469	-19.464772	NwC	NgP	Moremi
82	Moremi NP, BW	34	23.103469	-19.464772	NwC	NgP	Moremi
83	Moremi NP, BW	34	23.103469	-19.464772	NwC	NgP	Moremi
84	Moremi NP, BW	34	23.103469	-19.464772	NwC	NgP	Moremi
85	Moremi NP, BW	34	23.103469	-19.464772	NwC	NgP	Moremi
86	Moremi NP, BW	34	23.103469	-19.464772	NwC	NgP	Moremi
87	Moremi NP, BW	34	23.103469	-19.464772	NwC	NgP	Moremi
88	Windhoek, NA	35	17.080672	-22.575553	SoC1	Mx	Namibia
89	Okapuka, NA	36	17.080672	-22.575553	NwC	Mx	Namibia
90	Otjiwarango, NA	37	17.080672	-22.575553	SoC1	Mx	Namibia
91	Windhoek, NA	37	17.080672	-22.575553	SoC1	Mx	Namibia
92	Kheetmanshoop, NA	38	18.1466	-26.5762	SoC1	SgP	Orange River
93	Nieu Bethesda, ZA	39	24.5548	-31.8664	SoC1	SgP	Karoo
94	Nieu Bethesda, ZA	39	24.5548	-31.8664	SoC1	SgP	Karoo
95	Nieu Bethesda, ZA	40	24.5548	-31.8664	SoC1	SgP	Karoo
96	Nieu Bethesda, ZA	40	24.5548	-31.8664	SoC1	SgP	Karoo
97	Nieu Bethesda, ZA	41	24.5548	-31.8664	SoC3	SgP	Karoo
98	Nieu Bethesda, ZA	39	24.5548	-31.8664	SoC1	SgP	Karoo
99	Nieu Bethesda, ZA	39	24.5548	-31.8664	SoC1	SgP	Karoo

University of Cape Town

Appendix 5B- A table cataloguing sample information and genetic and geographic groupings of the phylogeographic analysis. These haplotype numbers are for mitochondrial D-loop sequences.

Sequence no.	Locality	Haplotype/ MapID	longitude	latitude	Network Clade	Geographic Population	Geographic area
100	Nieu Bethesda, ZA	42	24.5548	-31.8664	NwC	SgP	Karoo
101	Nieu Bethesda, ZA	2	24.5548	-31.8664	NeC2	SgP	Karoo
102	Cape Peninsula, ZA	43	18.407106	-34.157564	SoC1	SgP	Cape
103	Cape Peninsula, ZA	44	18.407106	-34.157564	SoC1	SgP	Cape
104	Richtersveld, ZA	45	16.996694	-28.441139	SoC1	SgP	Orange River
105	Richtersveld, ZA	46	16.996694	-28.441139	SoC1	SgP	Orange River
106	Tsitsikamma, ZA	47	23.6196	-33.9493	SoC1	SgP	South Coast
107	Tsitsikamma, ZA	47	23.6196	-33.9493	SoC1	SgP	South Coast
108	Tsitsikamma, ZA	17	23.6196	-33.9493	SoC1	SgP	South Coast
109	Tsitsikamma, ZA	18	23.6196	-33.9493	SoC1	SgP	South Coast
110	Tsitsikamma, ZA	18	23.6196	-33.9493	SoC1	SgP	South Coast
111	Tsitsikamma, ZA	48	23.6196	-33.9493	NwC	SgP	South Coast
112	Tsitsikamma, ZA	20	23.6196	-33.9493	SoC3	SgP	South Coast
113	Tsitsikamma, ZA	17	23.6196	-33.9493	SoC1	SgP	South Coast
114	Tsitsikamma, ZA	18	23.6196	-33.9493	SoC1	SgP	South Coast
115	Tsitsikamma, ZA	20	23.6196	-33.9493	SoC3	SgP	South Coast
116	Van Rhyns Dorp, ZA	49	18.7541	-31.6004	SoC1	SgP	Cape
117	Waterberg, ZA	50	28.259406	-23.842417	NwC	Mx	Waterberg
118	Waterberg, ZA	51	28.259406	-23.842417	NwC	Mx	Waterberg
119	Waterberg, ZA	9	28.259406	-23.842417	NwC	Mx	Waterberg
120	Waterberg, ZA	52	28.259406	-23.842417	NL	Mx	Waterberg
121	Waterberg, ZA	9	28.259406	-23.842417	NwC	Mx	Waterberg
122	Waterberg, ZA	19	28.259406	-23.842417	NwC	Mx	Waterberg
123	Waterberg, ZA	53	28.259406	-23.842417	NwC	Mx	Waterberg
124	Waterberg, ZA	54	28.259406	-23.842417	NwC	Mx	Waterberg

University of Cape Town

Appendix 5B- A table cataloguing sample information and genetic and geographic groupings of the phylogeographic analysis. These haplotype numbers are for mitochondrial D-loop sequences.

Sequence no.	Locality	Haplotype/ MapID	longitude	latitude	Network Clade	Geographic Population	Geographic area
125	Waterberg, ZA	55	28.259406	-23.842417	NwC	Mx	Waterberg
126	Waterberg, ZA	56	28.259406	-23.842417	NwC	Mx	Waterberg
127	Waterberg, ZA	29	28.259406	-23.842417	SoC3	Mx	Waterberg
128	West Coast, ZA	57	28.259406	-23.842417	SoC1	SgP	Cape
129	West Coast, , ZA	58	19.450925	-32.800533	SoC2	SgP	Cape
130	WestCoast, , ZA	59	19.450925	-32.800533	SoC3	SgP	Cape
131	Kafue, ZA	60	25.855269	-17.853531	NL	NgP	Okavango
132	Kafue, ZA	61	25.855269	-17.853531	NL	NgP	Okavango

Table Legend:

Sequence no.- The number allocated to each individual that was sequenced for this analysis

Locality- Provides descriptive data for sample provenance as place name and country. Countries are identified by two letter codes; AO= Angola, , BW= Botswana, MZ= Mozambique, NA= Namibia,
ZA= South Africa, ZW= Zimbabwe, ZM= Zambia.

MapID- The unique haplotype number assigned to that sequence and also serves as an identifier in the map of sample distribution (Fig 5.1).

Longitude/Latitude- GPS co-ordinates for sample provenance.

Network clade – The clade to which that individual has been assigned by the network analysis.

Geographic population- The population to which an individual has been assigned based on its geographic location. NB: "Mx" refers to zones of lineage mixing.

Geographic area - Group to which an individual has been assigned for the ANOVA analysis.

University of Cape Town

Appendix 5C- Results of the AMOVA analysis. F tests are based on Reynold's distances. Statistical significance is calculated using 110 random permutations and reported at the 0.05 level.

F statistics

	Cape	East Coast	Karoo	Kimberley	Loskop	Moremi	Mpumalanga	Okavango	Orange River	Tsitsikamma	Waterberg
Cape	0.000										
East Coast	0.433	0.000									
Karoo	0.099	0.411	0.000								
Kimberley	0.130	0.306	0.180	0.000							
Loskop	0.404	0.097	0.383	0.162	0.000						
Moremi	0.799	0.447	0.786	0.669	0.493	0.000					
Mpumalanga	0.796	0.366	0.780	0.682	0.509	0.979	0.000				
Okavango	0.716	0.291	0.703	0.567	0.354	0.517	0.803	0.000			
Orange River	0.155	0.292	0.180	0.090	0.282	0.713	0.696	0.618	0.000		
Tsitsikamma	0.034	0.361	0.099	0.021	0.294	0.764	0.760	0.670	0.121	0.000	
Waterberg	0.496	0.132	0.495	0.329	0.119	0.295	0.570	0.065	0.385	0.429	0.000

P values

	Cape	East Coast	Karoo	Kimberley	Loskop	Moremi	Mpumalanga	Okavango	Orange River	Tsitsikamma	Waterberg
Cape	0.000										
East Coast	0.000	0.000									
Karoo	0.018	0.009	0.000								
Kimberley	0.027	0.000	0.009	0.000							
Loskop	0.000	0.009	0.000	0.018	0.000						
Moremi	0.000	0.000	0.000	0.000	0.000	0.000					
Mpumalanga	0.000	0.000	0.000	0.000	0.000	0.000	0.000				
Okavango	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000			
Orange River	0.018	0.009	0.000	0.008	0.009	0.000	0.000	0.000	0.000		
Tsitsikamma	0.144	0.018	0.005	0.279	0.000	0.000	0.000	0.000	0.072	0.000	
Waterberg	0.009	0.045	0.000	0.009	0.027	0.000	0.000	0.108	0.000	0.000	0.000

University of Cape Town

CHAPTER 6

TESTING MODELS OF DIVERSIFICATION IN *PAPIO URSINUS* USING STATISTICAL PHYLOGEOGRAPHIC METHODS

Introduction

This thesis has focussed on the role that palaeo-environmental change has played in driving diversification within *Papio ursinus.* Results identify three landscape related mechanisms as important drivers of diversification in chacma baboons. The first is allopatric diversification via the emergence of geographic barriers to gene flow while the second deals with population responses to climate-driven habitat change during the Pleistocene. Although the inference is indirect, the results also suggest a third possible mechanism, that of long range dispersal events, contributing to divergence between colonizing and ancestral populations. Finally, a fourth mechanism, ecologically limited dispersal, is likely to contribute to ongoing population differentiation during periods of environmental stability and is investigated in this chapter.

From the results of chapters 4 and 5, a timeline of diversification in chacma baboons can be constructed for the last 1.8 myr. Age estimates reveal that chacma baboons diverged from ancestral *Papio* at ~1.80 Ma, at ~1.60 Ma, the species subdivided into two further genetically distinct lineages, to the areas north and south of the Kalahari Desert. This diversification event is linked to climate driven aridification of central southern Africa and the subsequent expansion of the Kalahari Desert (Sithaldeen et al. 2009; Stokes 1998) and suggests that climate driven landscape change has contributed significantly to diversification within baboons. There is also further evidence for two later, climatically linked diversifications within chacma. First, the emergence of a clade of Namibian baboons at ~1.00 Ma is coincident with climatically driven fluctuations in the level of the Orange River (Dollar 1998; Zawada 1995) and second, the emergence of a north-eastern chacma clade at ~420kya which may be linked to an unusually long interglacial period (Howard 1997; Muller and MacDonald 1997; Olson and Hearty 2009; Raynaud et al. 2005). It is assumed, particularly for the estimation of node ages in the Bayesian analysis, that major divergence events within the chacma baboon results from genetic isolation among populations. This assumption underpins vicariance as a mode of diversification between northern and southern mitochondrial lineages and is tested here using coalescent modelling techniques.

University of Cape Town

While landscape change may cause the initial disruptions to dispersal and gene flow ultimately leading to significant lineage divergence, population responses to these changes also play a role in determining the degree to which two groups diverge. The second mechanism tested here is the role of population size changes in shaping genetic diversity within baboons. A reconstruction of population demographic histories, using proxy measures for changes in population size, indicates that both the northern and southern populations experienced past demographic expansions. Together with landscape diversity measures, results suggests that the southern population first experienced a contraction into a glacial refuge in the south-western Cape from which it later expanded. Given that ~20 cycles of glacial-interglacial conditions have occurred in the past 1.8 myr, it is highly likely that repeated cycles of range contraction and expansion have been a major factor in shaping diversity within baboons (Jolly 2001). Here I use Bayesian skyline plots to test for past changes in *theta Θ*, the effective population size scaled by mutation rate, over time in both the northern and southern chacma lineages.

The spatial fragmentation of a species can also result when small groups expand their range into new areas; as they diverge, dispersals and subsequent gene flow with the parent population can become reduced (Hewitt 2000). The diversification of a north-eastern clade at a time of climatic warming is interpreted here as the result of a long range dispersal event during periods of climatic amelioration ~360-420kya. In this scenario a subset of the northern population is proposed to have dispersed into new habitat, accompanied by limited dispersal and gene flow back into the parent population. This process alone could drive population differentiation at the fringe of a species' distribution (Hewitt 2000) and is tested here using calescent modelling approaches.

Finally, an understanding of how contemporary populations respond to periods of environmental stability may contribute significantly to our understanding of how variation accumulates across the distribution of a species; in so doing, studies testing these ideas can also inform our understanding, and allow for predictions, of how populations might respond to strong environmental selection during periods of instability. Here I test two models of population genetic differentiation in chacma baboons, these are (i) isolation by distance and (ii) ecologically limited dispersal.

Each of these scenarios described above is tested using mitochondrial DNA sequence data and analysed using statistical phylogeographic methods. Below is a basic introduction to statistical phylogeographic methods, followed by the analysis section where the four main scenarios are presented. For each scenario methods of testing are outlined, analyses are

University of Cape Town

detailed and results are presented and interpreted. The chapter ends with a discussion of these findings contextualized within a proposed model of baboon diversification.

Methods

Nested clade phylogeographic analysis (NCPA) is a traditional phylogeographic method that focuses on the association of genotype and phenotype, and sample provenance is treated as a phenotypic character. Here the association between genotype and phenotype are non-independent (Templeton 1998). The NCPA user begins by estimating a gene tree or a network of haplotypes for the sample. Haplotype clusters are then used to group samples into nested geographic clades. Statistical support for these clades is then computed using an explicit inference key from which biologically meaningful interpretations are constructed (Templeton 1998).

The NCPA protocol involves several methodological steps that are prone to unquantifiable error; (i) when constructing the tree, assumptions must be made to accommodate homoplasy, (ii) the nesting of clades is the user's choice and therefore subjective, and (iii) there can also be significant ambiguity in the interpretation of the inference key (Nielsen 2009). Additionally one cannot estimate the error associated with the analysis or test if alternative hypotheses could fit the data equally well (Knowles and Madisson 2002). NCPA can also produce a high rate of false positives (Knowles and Madisson 2002; Panchal and Beaumont 2007; Petit and Grivet 2002). Although this particular issue has been addressed and the inference key adjusted accordingly, there has been an associated reduction in the power of the method (Templeton 2004).

The evolutionary process is inherently stochastic. This is due to the process of lineage sorting which produces disparities in population and gene histories and thereby introduces a degree of randomness to the reconstruction of the evolutionary past. It is therefore generally recognized that one cannot assume a direct positive correlation between demographic history and evolutionary processes with genealogical pattern alone as this ignores the stochasticity (Arbogast et al. 2002; Edwards and Beerli 2000; Irwin 2002). Many have therefore abandoned NCPA in favour of methods that provide greater statistical confidence and integrate stochasticity into the framework (Hickerson 2010) such as simulation based phylogeography which models probable population histories and is based on a coalescent framework.

University of Cape Town

Coalescent theory allows the user to sample a population, model demographic history and estimate phylogeographic parameters all the while taking into account the stochastic nature of evolution (Wakeley 2008). This stochasticity is integrated into the coalescent framework by using the genealogy only as a transition parameter to obtain other biologically important parameters (Hey and Manchado 2003). There are three significant advantages of coalescent theory, for one it is sample based and seeks to describe the sample rather than the whole population, secondly it is highly suitable for DNA sequence data and thirdly is highly efficient due to the development of algorithms to estimate model parameters (Fu and Li 1999). The coalescent also allows the use of the full set of sequence data rather than a subset of segregating sites as in summary statistics. This allows the investigator to address a wider range of interesting biological questions related to the processes that have influenced taxon evolution, the historical relationships amongst different subpopulations and their respective ancestry (Stephens 2001).

Coalescent theory hinges on the probabilities associated with the coalescent event. Coalescence occurs when genes from the present merge into their common ancestor at some point in the past (Knowles 2009). These coalescent events only end when the process converges on the MRCA of the entire sample (Nordberg 2001). The coalescent is defined in terms of effective population size (Ne) and time (t) and the mutation rate μ and (vi) Theta (the scaled mutation parameter) = 2Nu (Stephens 2001). These mathematical properties allow for the genealogy to be modelled backwards in time and the mutations superimposed on the genealogy, forwards in time (Nordberg 2001). The use of coalescent theory in phylogeography therefore allows us to model the most likely history of populations (Fu and Li 1999).

The coalescent approach is modelled on the Wright -Fisher population model which assumes that the sequences present in a population today are a random sample (with replacement), from those in the previous generation, that population size remains constant from generation to generation and that all mutations are selectively neutral (Fu 1997). Neutral mutations do not affect reproductive success; these mutations can therefore be separated from genealogy. This ideal model is biologically unrealistic but the coalescent is easily adapted to more general, real populations, provided that time is scaled appropriately. The coalescent can accommodate a number of violations of the Wright-Fisher model including variable population size, population structure in time, geographical structure, segregation, recombination and selection (Nordberg 2001).

University of Cape Town

Mathematical modelling of the coalescent provides a statistical framework for estimating many of the hypothetical model parameters such as effective population size, migration rates and divergences times (Knowles and Madisson 2002). This information can be incorporated into a model which includes other known external information about the species and can ultimately be used to generate a set of testable hypotheses about the process of differentiation (Knowles 2004). Competing hypotheses or null hypotheses can then be evaluated using different theoretical methods to assess which best fits the data. This is done using likelihood based inference, either Maximum Likelihood or Bayesian approaches (Beerli and Felsenstein 2001; Hey and Nielsen 2004, 2007; Kuhner 2006; Kuhner et al. 1998). The coalescent modelling of population histories has become the standard statistical and theoretical framework to approach the testing of phylogeographic hypotheses (Avise 2000).

Analysis

Scenario 1: Allopatric diversification during the Pleistocene

The phylogenetic reconstruction presented in chapter 4 provides clear evidence for two distinct mitochondrial lineages within chacma baboons: a 'northern lineage' (NL) and a 'southern' (SL) lineage, a pattern which is also recovered in published studies (Keller et al. 2010; Sithaldeen et al. 2009; Zinner et al. 2009). This event is dated to ~1.60 Ma and was likely the result of allopatric fragmentation driven by glacial aridification and expansion of the Kalahari Desert, a proposal which is highly compatible with the data from other terrestrial records which show a trend towards drying in South Africa between 1.80 Ma and 1.60 Ma (Lee-Thorp et al. 2007).

Most theoretical models of population subdivision assume that divergent populations have either been exchanging low levels of migrants at a constant rate for an infinitely long time or, that populations which are descended from a common ancestral population diverged without further gene flow (Nielsen and Wakeley 2001). This model of 'Isolation with Migration' is implemented in the programme IM (Hey 2007) and estimates six demographic parameters from DNA sequence data in a pairwise test; the effective population sizes of each of the two extant populations (N1 and N2), the ancestral population size prior to population divergence (NA), immigration rates (m1= the proportion of the population of N2 replaced by individuals from N1 each generation and m2= the proportion of the population of N1 replaced by individuals from N2 each generation), and the time (t) since divergence for the two populations.

University of Cape Town

The program (IM) uses coalescent theory (Kingman 1982, 2000) and Bayesian methods to differentiate divergence through isolation with no migration from divergence with some degree of continued gene flow (Nielsen and Wakeley 2001; Hey and Nielsen 2004). IM uses Markov Chain Monte Carlo (MCMC) simulations to simultaneously estimate the six parameters that affect the mean genetic divergence between the populations being tested and is dependent on user defined priors for each parameter. Priors are input as scalars of each parameter and may be quite narrow if informed by external data or must be wide enough to accommodate any biological reality.

The assumptions of the IM model are that, the sampled populations are more closely related to each other than to any other extant population, the marker is selectively neutral, there is no recombination within the locus and that mutation has followed the model that is being applied to the data (Hey 2007). A well resolved parameter will generate a normal distribution curve where both tails are flat and fall to zero (Hey 2007). IM can also estimate the time to most recent common ancestor (TMRCA) for all sampled haplotypes. If an accurate estimation of mutation rate, and data on average reproductive age is available, these scalars can be further converted to demographic parameters which are biologically informative (Hey and Nielsen 2004).

Here I propose that the subdivision of chacma baboons into two mitochondrial lineages was driven by a period of genetic isolation between two populations, north and south of the expanded Kalahari Desert, with minimal, if any continued gene flow.

Using simulated data sets, Nielsen and Wakeley (2001) showed that even a single non-recombining locus can provide substantial power for differentiating between divergence in complete isolation and divergence with migration (also see Griswold & Baker 2002). Here I use IM to simultaneously estimate model parameters for the two main mitochondrial lineages, NL and SL, identified in chapter 4. The input file was constructed with NL assigned as population 1 and SL as population 2. An HKY mutation model was employed and an inheritance scalar of 0.25 used, as is appropriate for mitochondrial DNA (Hey 2007). Here I test the fit of two models to the dataset, (1) a model of divergence with isolation between lineages and (2) a model of divergence with low levels of gene flow between lineages. Given a mutation rate and generation time for the sample IM automatically converts parameters into demographic units. While not ideal as explained earlier in this thesis, due to limitations of the software this conversion requires the implementation of a strict molecular clock. The generation time was input as 5 years (-u 5) based on average age of first time of reproduction in chacma females (Ascunce et al. 2007; Zunino 1996). The substitution rate of

University of Cape Town

the Brown region is estimated at 2% per million years (Brown 1982). Using this approximation and a sequence length of 879bp, I calculated a mutation rate $0.177x10^{-4}$ per year for this marker.

A series of short simulations of 1 million chains (-l 1000000) with a 10% burn in (-b 100000) were first run in order to explore the data and assess the shape of posterior parameter distributions under the assumptions of each of the two models. Demographic parameter priors were input as scalars which are defined in Table 6.1. For a model of no migration, m1 and m2 were set to zero and q1, q2, qa and t were estimated. To generate a probability distribution for each of these parameters, priors were set to encompass all biological possibilities (-q1 20, -t 40). I then tested a model of migration between populations. Here again I began with wide priors (m1=m2=10, -q1 10, -t 10). Distribution curves are presented in Figure 6.1.Given the set of priors, distribution curves for a model of no migration (Fig. 6.2) suggest that this is a better fit for the dataset. Posterior parameter distributions were therefore estimated for a model of no migration using longer sampling chains.

Demographic Parameter	Scalar	Description
N1	q1	Effective number of female chacma in present day NL
N2	q2	Effective number of female chacma in present day SL
NA	qa	Effective number of females in the ancestral population at the time of population divergence
m1	*m1*	Probability of migration from SL to NL, per gene copy per generation
m2	*m2*	Probability of migration from NL to SL, per gene copy per generation
T	t	Time since split between NL and SL

Table 6.1- Definition of the parameters estimated by IM in the analysis of isolation with migration between NL and SL.

Metropolis coupling was implemented using five chains with five chain swap attempts per step and a two-step heating increment. I used a burn-in period of 500000 and allowed the program to run for 5000000 steps so that the lowest effective sample sizes (ESS) for each parameter were at least 500 (Hey and Nielsen 2004). To test that the priors were sampling the full data space, I executed IM under identical conditions, but with different random number seeds, for a total of four independent runs. If the data space is being fully explored in each run, independent runs will converge on highly similar parameter estimates. Because all four runs gave very similar parameter estimates and 95% highest posterior distributions (HPDs), I report results from one run only. Distribution curves of the final run are presented

University of Cape Town

in Fig. 6.3. For each parameter, I report the mean of the posterior distribution and the 95% HPDs in Table 6.2. These values were converted into demographically meaningful units and are presented as distribution curves in Fig. 6.2 and are reported in Table 6.3.

Results and interpretations

i. Results under a model of migration

The posterior distributions of parameters constructed under a model of migration between NL and SL are presented in Fig. 6.1. Both q1 and q2 are well resolved i.e. the distribution approaches a normal curve and both tails are flat and fall to zero (Hey 2007). However, qa and t are not well resolved as the distributions do not approximate a normal curve (Hey 2007). Multiple attempts were made to explore the data with wider priors but no improvement on these distributions was observed. For m1 and m2 the shape of the curve indicates a well resolved parameter which approaches zero. In this case, due to the fact that the parameter cannot be negative, this distribution cannot approach normality. This model performs poorly when compared to the resolution achieved with a model of no migration. This model was therefore abandoned and demographic estimations were based on a model of no migration.

i. Results under a model of no migration

With one exception, parameter estimates were all well resolved under a model of no migration between NL and SL. Both q1 and q2 were well resolved with parameter distributions approaching a normal curve and both tails becoming flat and falling to zero (Hey 2007). For qa the posterior distribution increases steadily as parameter value increases but does not reveal a peak. This suggests that there is not enough information in the data for an accurate estimation of this parameter given the set of priors. Time since divergence (t) is well resolved. These results support a history of genetic isolation between northern and southern chacma populations maintained for a period of time.

The outputs for the 5 000 000 chain run for a model of no migration are summarised in Table 6.2. All ESS values are >500. This is followed by a presentation of the posterior distribution curves for each of the well resolved parameters estimated under a model of no migration and presented in demographic units (Fig. 6.3). The mean values of each of the demographic parameters estimated under a model of no migration are presented in biologically meaningful units with 95% HPD values in Table 6.3. Mean effective female population size was larger for NL at ~ 317000 individuals while mean effective female population size for SL

University of Cape Town

was ~261000 individuals. The posterior distribution of t suggests that these lineages differentiated at about 1.27 Ma. This value is congruent with previous estimates (chapter 4) for the ages of of these two mitochondrial lineages.

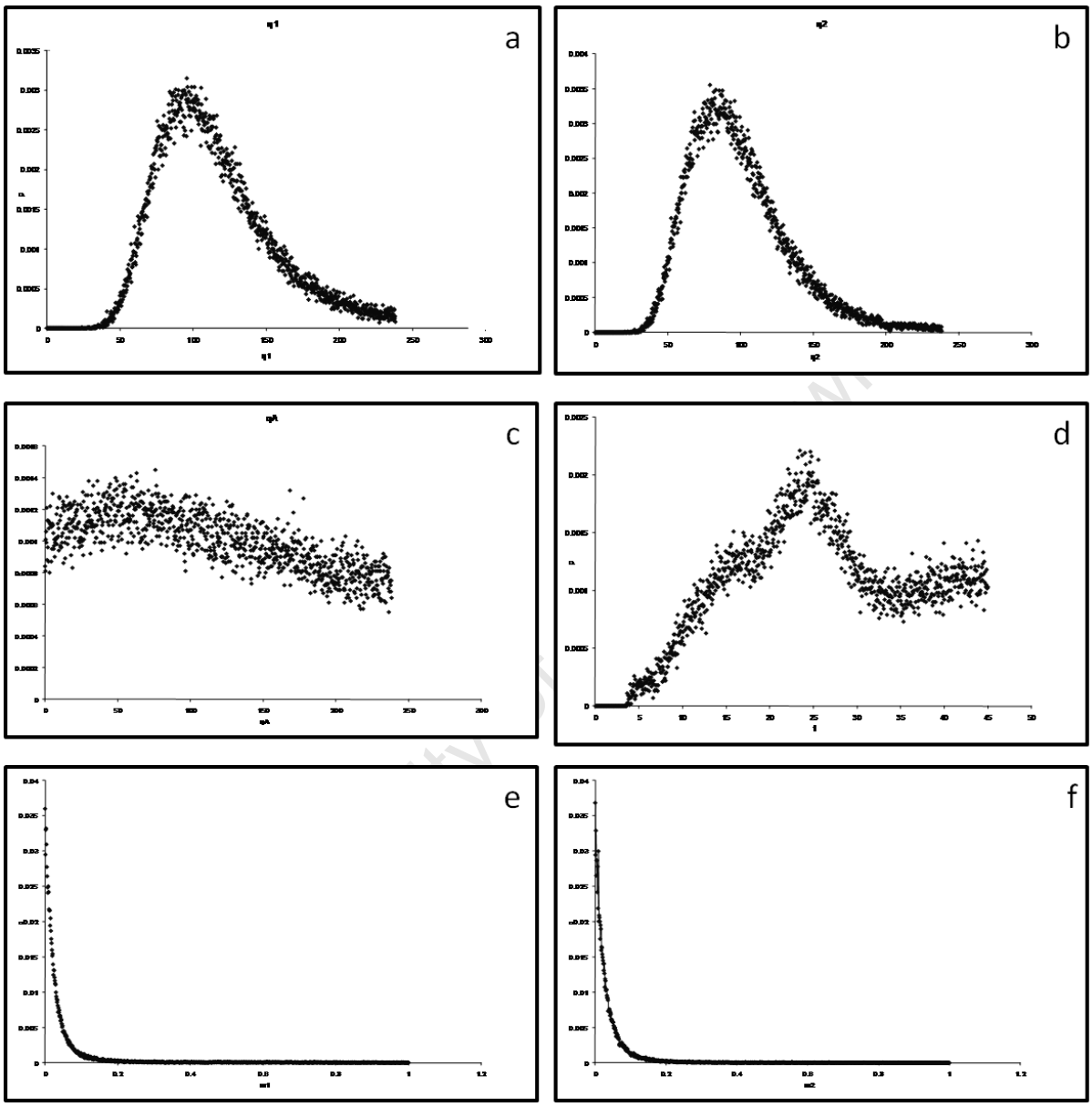

Figure 6.1- The posterior distribution curves of parameters estimated under a model of migration between NL and SL (a) Distribution of q1 (population size of NL). (b) Distribution of q2 (population size of SL). (c) Distribution of qa (ancestral population size). (d) Distributions of t (time since divergence). (e) Distribution of m1, the migration rate from NL into SL. (f) Distribution of parameter m2, migration rate from SL into NL.

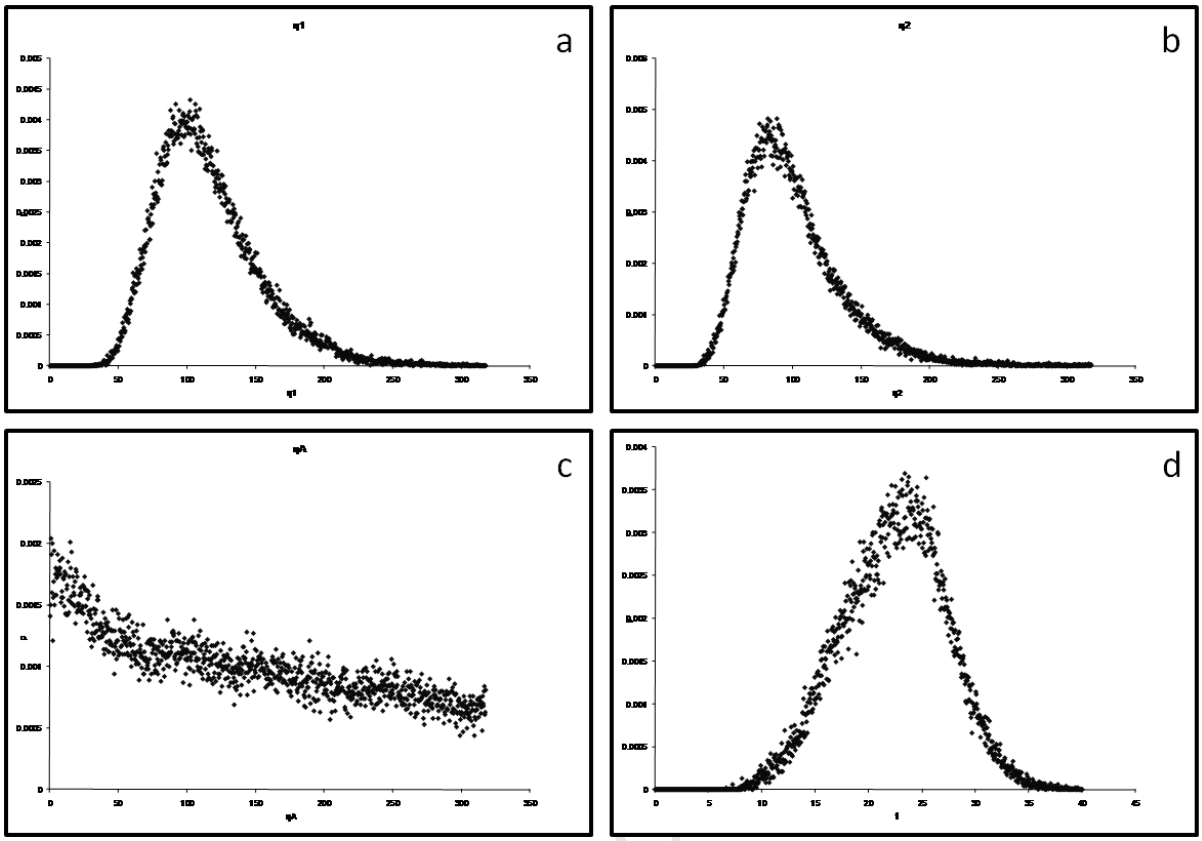

Figure 6.2- The posterior distribution curves of parameters estimated under a model of no migration between NL and SL (a) Distribution of q1 (population size of NL). (b) Distribution of q2 (population size of SL) (c) Distribution of qa (ancestral population size). (d) Distribution of t (time since divergence).

University of

Autocorrelations and Effective Sample Size Estimates			
q1	**q2**	**qa**	**t**
7766	39000	3416	632

Table 6.2a- ESS values for a 5000000 chain run under a model of no migration

Marginal Histograms-Summaries				
	q1	**q2**	**qA**	**t**
Minbin	29.082	23.778	0.179	3.140
Maxbin	317.676	178.838	317.676	39.980
HiPt	99.006	83.829	43.067	23.340
HiSmth	98.688	83.829	19.229	23.300
Mean	110.130	91.477	100.913	22.140
97Lo	60.230	49.980	4.609	10.300
97Hi	212.791	149.780	300.831	30.980
HPD90Lo	77.071?	71.409	0.178?	12.82
HPD90Hi	174.968?	138.178	262.372?	30.1

Table 6.2b- Summary values of the marginal histogram plots for a 5000000 chain run under a model of no migration.

University of Cape Town

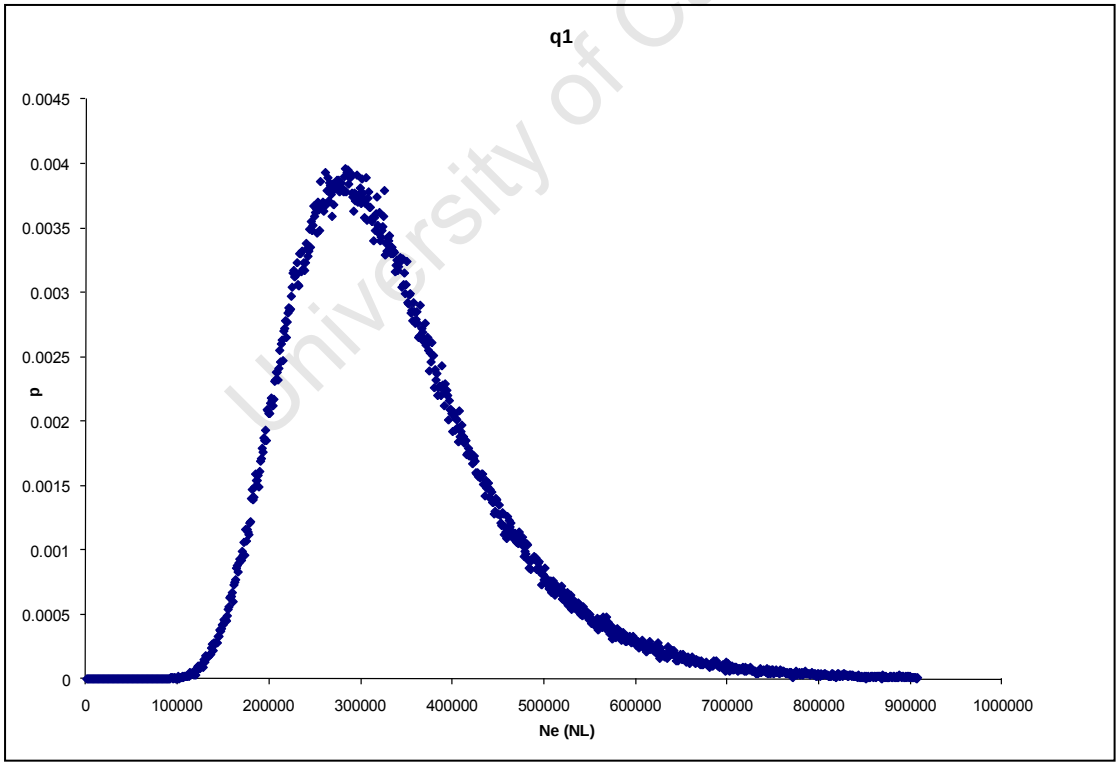

Figure 6.3a- Posterior distribution of population size for NL under a model of no migration.

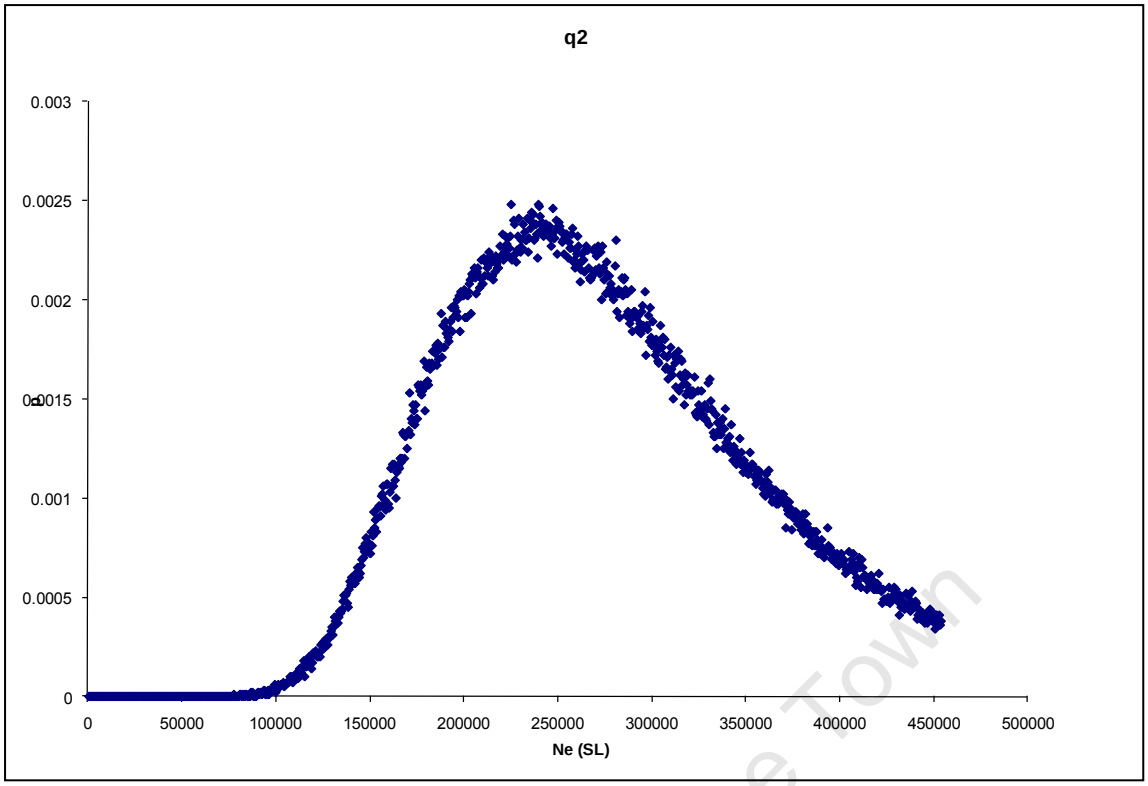

Figure 6.3b- Posterior distribution of population size for SL under a model of no migration.

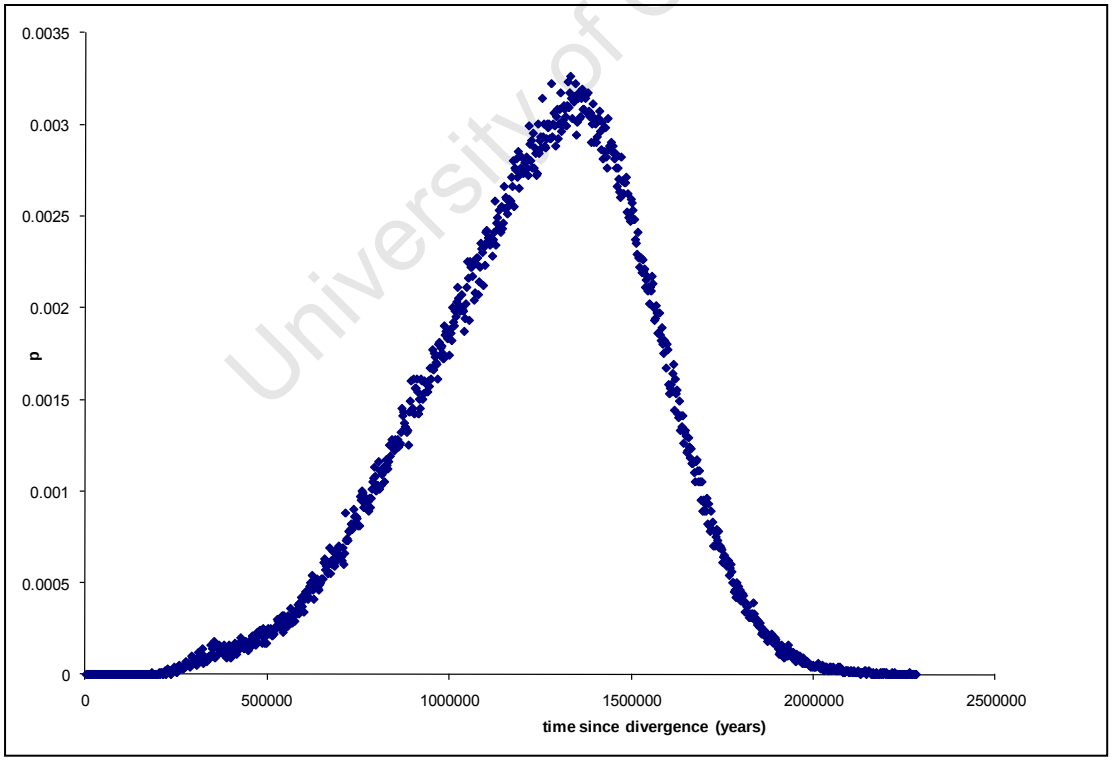

Figure 6.3c- Posterior distribution of time since divergence between NL and SL under a model of no migration.

University of Cape Town

	q1	q2	t
Mean	314676	261307	1267142
97% Lo	172087	142798	788771
97% Hi	607973	427942	1770287

Table 6.3- Table of demographic units estimated under a model of no migration between SL and NL.

Scenario 2: A test for changes in population size over time

Whilst today the distribution of local baboon troops may be limited by anthropological factors, chacma baboons as a species are distributed almost continuously across the southern African landscape. It is likely then, that gene flow across this region is also continuous. This type of geographic and genetic continuity reduces the probability of significant amounts of population subdivisioning within the species. Nonetheless, the phylogenetic reconstruction presented in chapter 4 provides clear evidence for the maintenance of two distinct mitochondrial lineages within chacma, the emergence of which has been convincingly linked to a vicariance event brought on by the expansion of the Kalahari Desert. Statistical proxies for demographic change suggest that both southern and northern chacma populations have experienced recent demographic expansions. Together these findings suggest that the current, ubiquitous, distribution of chacma baboons is a fairly recent phenomenon.

Results of the summary statistics and pairwise mismatch distributions reported in chapter 5 suggest recent expansion events for both the southern (SL) and the northern (NL) chacma populations. These summary statistics (based on distributions of haplotypes and numbers of segregating sites) do not however use all of the historical information contained in DNA sequence data (Sousa et al. 2009). This is also true for pairwise mismatch distribution curves. Coalescence methods however do, and are therefore able to test for demographic expansion in a more statistically robust way and provide a means by which to date demographic events. Here I assess changes in theta Θ, the effective population size scaled by the mutation rate, over time for each of the major lineages, SL and NL, using a Bayesian framework.

The Bayesian skyline plot (BSP) uses DNA sequence data, a model of nucleotide substitution and time since monophyly to estimate changes in effective population size

University of Cape Town

through time (Drummond et al., 2007). This algorithm implemented in BEAST v1.4.2 (Drummond and Rambaut 2007) uses standard MCMC sampling procedures and a coalescent framework to generate a posterior distribution of effective population size at equally spaced intervals through time (Drummond and Rambaut 2007). Unlike earlier skyline plots this method takes into account both the error inherent in phylogenetic reconstruction, and the stochastic error intrinsic to the coalescent process. In so doing, the BSP produces more accurate estimates of statistical uncertainty. This type of skyline plot is also unique in that it accommodates a flexible model of growth, allowing for arbitrary patterns of historical population growth. The BSP employs a 'relaxed clock' mutational model which allows the mutation rate to vary among lineages (Atkinson et al. 2009; Drummond and Rambaut 2006) and includes credibility intervals for the estimated effective population size at every point in time, back to the most recent common ancestor of the data set. These credibility intervals represent both phylogenetic and coalescent uncertainty (Atkinson 2009; Drummond and Rambaut 2007).

Here I assess changes in effective population size over time for each of two monophyletic lineages recovered within the Brown region dataset, NL and SL. I do this using Bayesian skyline plots (Drummond and Rambaut 2007) as implemented by the software BEAST v1.4.2 (Drummond and Rambaut 2007). In chapter 2, the HKY model (Hasegawa et al. 1985) with gamma-distributed rates among sites and a proportion of invariable sites, was determined to be the most appropriate model of nucleotide substitution for this marker and is employed here. For this analysis estimates of population size were grouped into 15 coalescent intervals (m=15) and the divergence date estimates for the MRCA of all haplotypes within each lineage, as calculated in chapter 5, were used as an upper time limit . A relaxed uncorrelated lognormal molecular clock was implemented with mean divergence dates and lognormal standard deviation values to reflect the 95% HPD obtained from the dating estimates for each clade. For NL mean divergence was set to 1.35 Ma, with a standard deviation of 0.18 and for SL mean divergence was set to 1.27 Ma, with a standard deviation of 0.25. The MCMC chains were run for 10 million generations, and sampled every 1000 steps. The first 10% of samples were discarded as conservative burn-in and each analysis was run twice and examined for convergence in TRACER 1.4 (Rambaut and Drummond 2007).

University of Cape Town

Results and interpretations

Bayesian skyline plots were used to visually illustrate changes in female effective population size (Nef) over time for each of the two mitochondrial Brown region lineages, NL (Fig. 6.4a) and SL (Fig. 6.4b). The skyline plot for NL shows a long period of sustained growth increasing the Nef by almost 2 orders of magnitude in the time period under investigation. The BSP for SL however, shows a two stage growth process. The first phase from 1.6 Ma to is a period of sustained growth at a constant rate. At approximately 17 ka the rate of growth increases significantly but is short lived (~10 kyr). This is interpreted as a sudden expansion event. Table 6.4 shows the absolute mean values for population sizes for NL and SL respectively. High ESS values indicate that the data is well sampled and that results are robust.

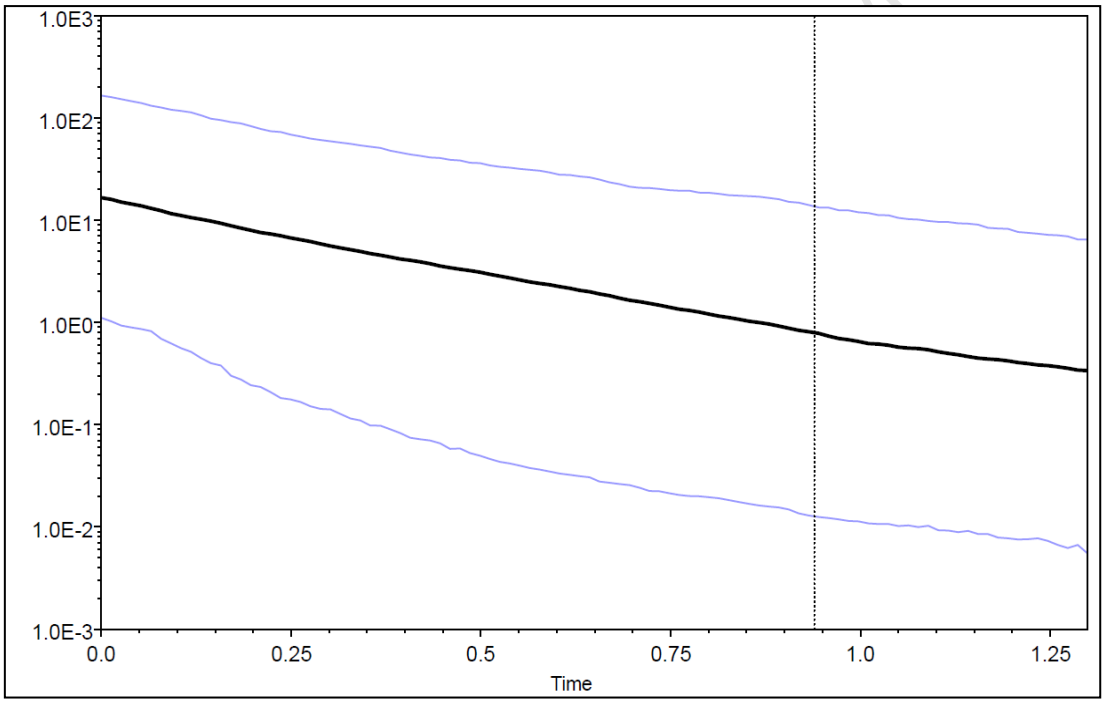

Fig 6.4a- The BSP for the NL northern population based on a 10 million chain MCMC analysis using 15 coalescent intervals. Y axis represents the female effective population size and the X axis is time in Ma. The black line represents the BSP and the blue lines the 95% HPD around the BSP. The dotted line represents the tree model root height. NL shows a long period of sustained growth for the period under investigation.

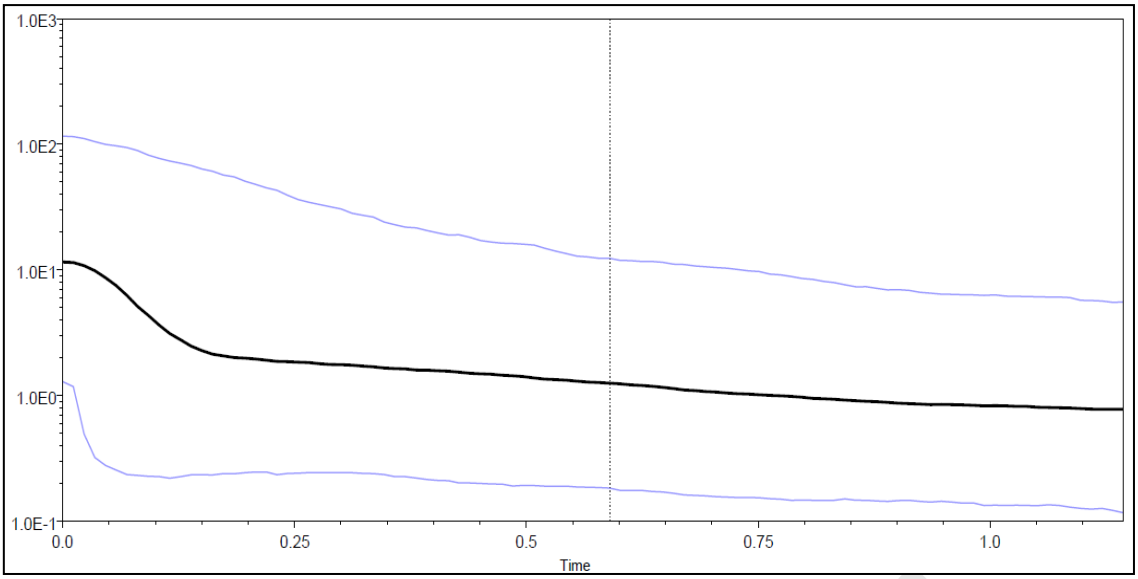

Fig 6.4b- The BSP for the SL southern population based on a 10 million chain MCMC analysis using 15 coalescent intervals. Y axis represents the female effective population size and the X axis is time in Ma. The black line represents the BSP and the blue lines the 95% HPD around the BSP. The dotted line represents the tree model root height. This plot shows a recent rapid increase in female effective population size from from 15-5ka.

	NL		SL	
Coalescent interval	**mean**	**ESS**	**mean**	**ESS**
skyline.popSize1	37.07	1371.32	28.91	792.8
skyline.popSize2	17.47	618.39	11.87	266.67
skyline.popSize3	12.32	778.97	7.32	222.09
skyline.popSize4	9.77	709.86	7.73	216.99
skyline.popSize7	7.63	427.73	4.37	174.07
skyline.popSize6	6.08	403.73	3.47	217.19
skyline.popSize7	4.88	720.73	2.84	290.83
skyline.popSize8	3.71	494.79	2.19	376.01
skyline.popSize9	2.66	792.21	2	419.38
skyline.popSize10	2.09	707.86	1.81	427.91
skyline.popSize11	1.47	462.1	1.6	799.84
skyline.popSize12	0.9	732.89	1.3	614.77
skyline.popSize13	0.76	707.03	0.97	421.71
skyline.popSize14	0.37	361.97	0.66	461.9
skyline.popSize15	0.27	247.47	0.41	632.4

Table 6.4- Table of effective population sizes for NL and SL for each of 15 coalescent time intervals. These mean values are graphed by the black BSP lines in Fig 6.4a and b.

University of Cape Town

Scenario 3- Evidence for long range dispersal during periods of climatic amelioration

Genealogical reconstructions strongly support a distinction between north western (NwC) and north eastern (NeC) chacma populations (Sithaldeen et al. 2009; Keller et al. 2010; chapter 5). It is estimated that NeC differentiated from the NL ~420kya, most likely as a minority fragment that became separated from NL in a single long range dispersal event across the Cape fold mountain range in the eastern part of South Africa. Diversification resulted from subsequent reduced gene flow between parent and daughter populations, possibly as forest barriers developed between groups during Marine Isotope Stage 11 (360-420kya) (Howard 1997; Muller and MacDonald 1997; Olson and Hearty 2009 ; Raynaud et al. 2005), when climatic warming continued for an unusually long time and likely resulted in the expansion of forest habitats (Lawes 1990) across the eastern Great Escarpment of South Africa. In this model a single ancestral population (NL) split and the majority fraction gave rise to the current day NwC and a minority fraction gave rise to NeC. Modern NeC then expanded from the founding population to achieve its current distribution while NwC an effective population size very close to the ancestral population.

This scenario can be modelled within the IM software package, however the basic IM model described under *Scenario 1* has certain limitations which do not fully accommodate this proposed series of events. The basic IM model assumes that all populations are constant in size and therefore cannot inform on issues of population size changes. For a model in which an ancestral population splits into two, an additional parameter (s) has been added. The relative sizes of the new populations are reflected in the parameter s, where $0 < s < 1$. *"At the time of the split, descendant population 1 has size s NA from which it moves to size N1 at the time of sampling. Similarly, population 2 begins with size (1 - s) NA from which it moves to size N2 at the time of sampling. Either population is free to either grow or shrink under this model"* (Hey 2007, pg1).

The mitochondrial control region evolves at a faster rate than the Brown region and is therefore more likely to capture the signature of more recent demographic change. Here I estimate model parameters using nine D-loop haplotypes from each of the two descendent populations NeC and NwC. NwC is assigned as 'population 1' and NeC as 'population 2'. The mitochondrial control region mutation rate estimated by Winney et al. (2004) for hamadryas baboons (0.13 changes per site per million years) is used. For a 463bp marker, this was converted into a mutation rate of 0.000647 (0.647×10^{-4}) mutations per year.

University of Cape Town

The probability that a minority fraction of an ancestral population (NL) founded NeC while the majority fraction gave rise to NwC is tested. Short preliminary runs (-l 1000000) with a 10% burn-in (-b 100000) with wide priors are used to explore the shape of posterior distribution curves (-m1 10, -m2 10, -q1 10, -t 10). As this model requires population fragmentation and population size change, the calculations are performed by using population size change (-j 9) and employing a lower and upper range limit for population split parameter s (-sl 0.5, -su 0.9). Priors were extensively tested in an attempt to produce better resolved distribution curves. Eventually a final run with the original set of widely distributed priors were used. Metropolis coupling was implemented using five chains with five chain swap attempts per step and a two-step heating increment. I used a burn-in period of 500000 and allowed the program to run for more 5000000 steps so that the lowest effective sample sizes (ESS; see Hey and Nielsen 2004) for each parameter were at least 500.

Parameter	Description
N1	Effective number of female chacma in present day NwC
N2	Effective number of female chacma in present day NeC
NA	Effective number of female chacma in ancestral NL
m1	Probability of migration from NeC to NwC, per gene copy per generation
m2	Probability of migration from NwC to NeC, per gene copy per generation
T	Time since founding of NeC
1-s	Fraction of NC that founded NeC

Table 6.5- Definition of the parameters estimated by IM in the analysis of isolation with migration between for a population fragmentation event where NL gives rise to NwC and NeC

University of Cape Town

Results and interpretations

It is clear from the distribution curves presented below (Fig 6.5), that most of the parameters estimated have failed to resolve adequately. Neither q1 or q2 are well resolved as the distributions do not approximate normal curves (Hey 2007). For q2 the right tails falling but fails to approach zero even if the prior is widened. Ancestral estimate qa seems to be slightly better resolved than q1 and q2 with both tails complete and with the highest probability approaching zero. Time since divergence (t) also appears to be better resolved than q1 and q2 with both tails complete. The fraction of NC that founded NwC (s) does not resolve satisfactorily (Hey 2007). Migration rates approach zero. These poor results are most likely due to an insufficient amount of data in the sample to address the question. Either the D-loop does not capture sufficient variation to address this issue or there is insufficient sampling of NeC. Results using the Brown region are considerably worse.

Extensive attempts were made to improve these results using longer runs (up to 100 million). Despite indications of good mixing and high ESS values no improvement in resolution was obtained.

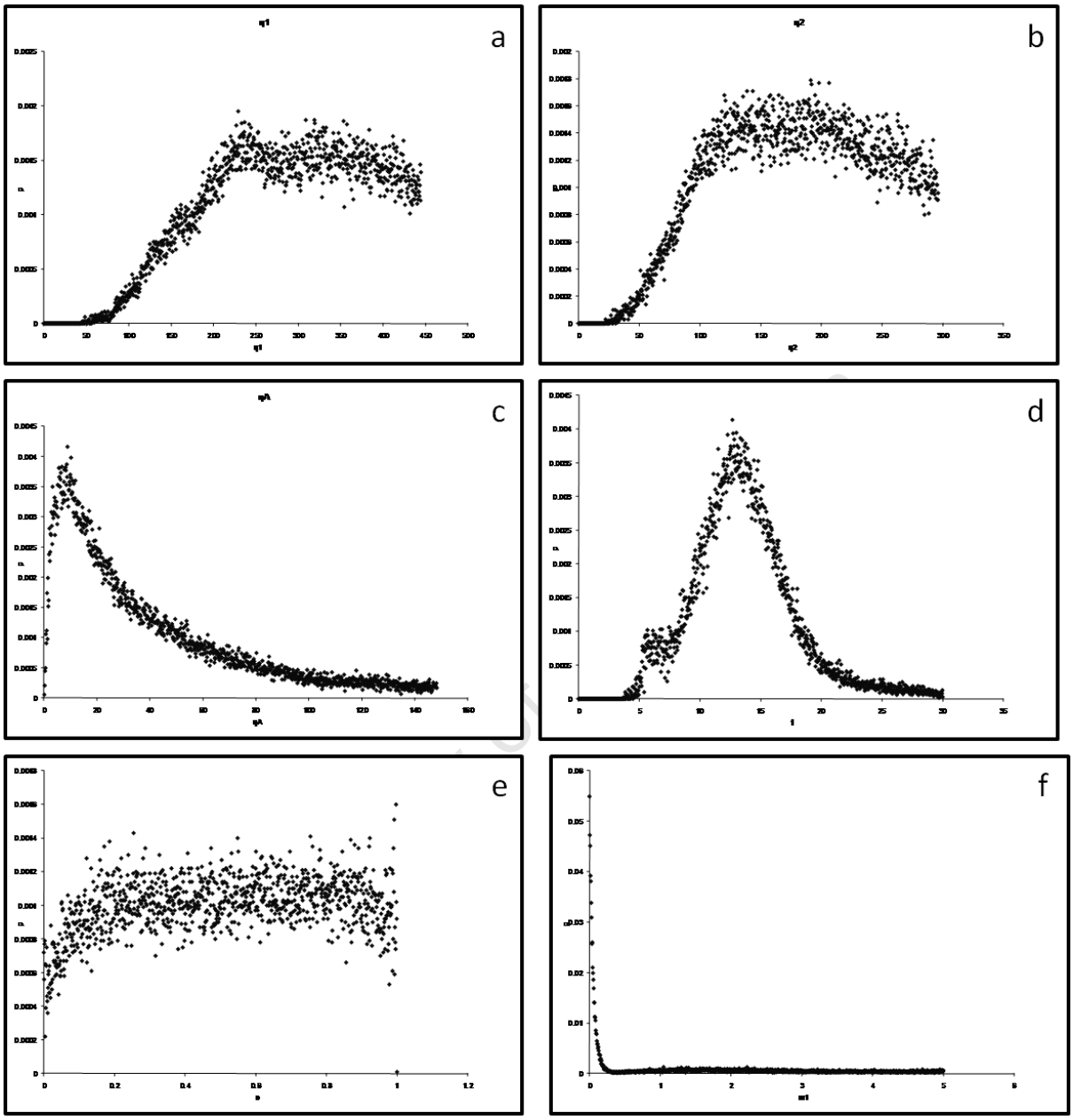

Figure 6.5- Posterior distribution curves for (a) q1 population size of NwC , (b) q2 population size of NeC (c) qa population size of ancestral NL . (d) t time since divergence (e) s that fragment of NL that gave rise to NwC (f) m1 the migration of NwC into NeC. The curve for m2 is identical.

Scenario 4- Continued differentiation within chacma baboons may be driven by either isolation by isolation or adaptation to local habitat

An important variable in the diversification process is the question of how regional variation within baboon species is accumulated during times of climatic stability. Here I consider modern climate as representative of periods of relative stability. Differentiation between baboon populations contributes to the accumulation of variation within the species. This may be driven by a multitude of factors including localized ecological or behavioural adaptations, or simply by distance between natal troop ranges. Here I test the fit of these two scenarios to the data by assessing their contribution to the structure that is observed within the sample. Both sets of analysis are performed using the D-loop sequence data as this, faster evolving marker, is more likely to retain variation related to more recent events than the Brown region.

Scenario 4.1- A test for isolation by distance

The Mantel Test (Mantel 1967) assesses the degree of correlation between two distance matrices. In this case the two matrices represent the pairwise geographic and genetic distances between individual sequences. Mantel tests were performed for the full data set of 132 D-loop sequences, including individuals from mixing localities; separate regressions were also generated for each of the two geographic populations NgP (N=70) and SgP (N=48), which did not include individuals from mixing localities. Allocations of individual sequences are tabled in Appendix 6A. Genetic distance was calculated as uncorrected p-distance between sequences. 1000 replicates were performed and significance is reported at the α=0.05 level. A weak positive correlation between the two matrices indicates that isolation by distance is not a significant factor in shaping the observed genetic structure.

Results

Plots of matrix correlation for the Mantel's tests are shown in Fig 6.6. Results reveal very little correlation between genetic and geographic distance. The highest correlation was for all 132 sequences (Fig 6.6a) (r=0.38, p=0.01). NgP is lower at r=0.21 but is significant (p=0.0009) (Fig 6.6b). The value of r for SgP is even lower (r= 0.13) and not significant (p=0.7) (Fig. 6.6c). The higher value for the full sample of 132 sequneces, is most likely due to the fact that it contains two distinct mitochondrial lineages that are, for the most part, also geographically separated. Results indicate that isolation by distance accounts for very little of the genetic structure observed across the whole sample but explains some of the pattern at smaller geographical scales.

University of Cape Town

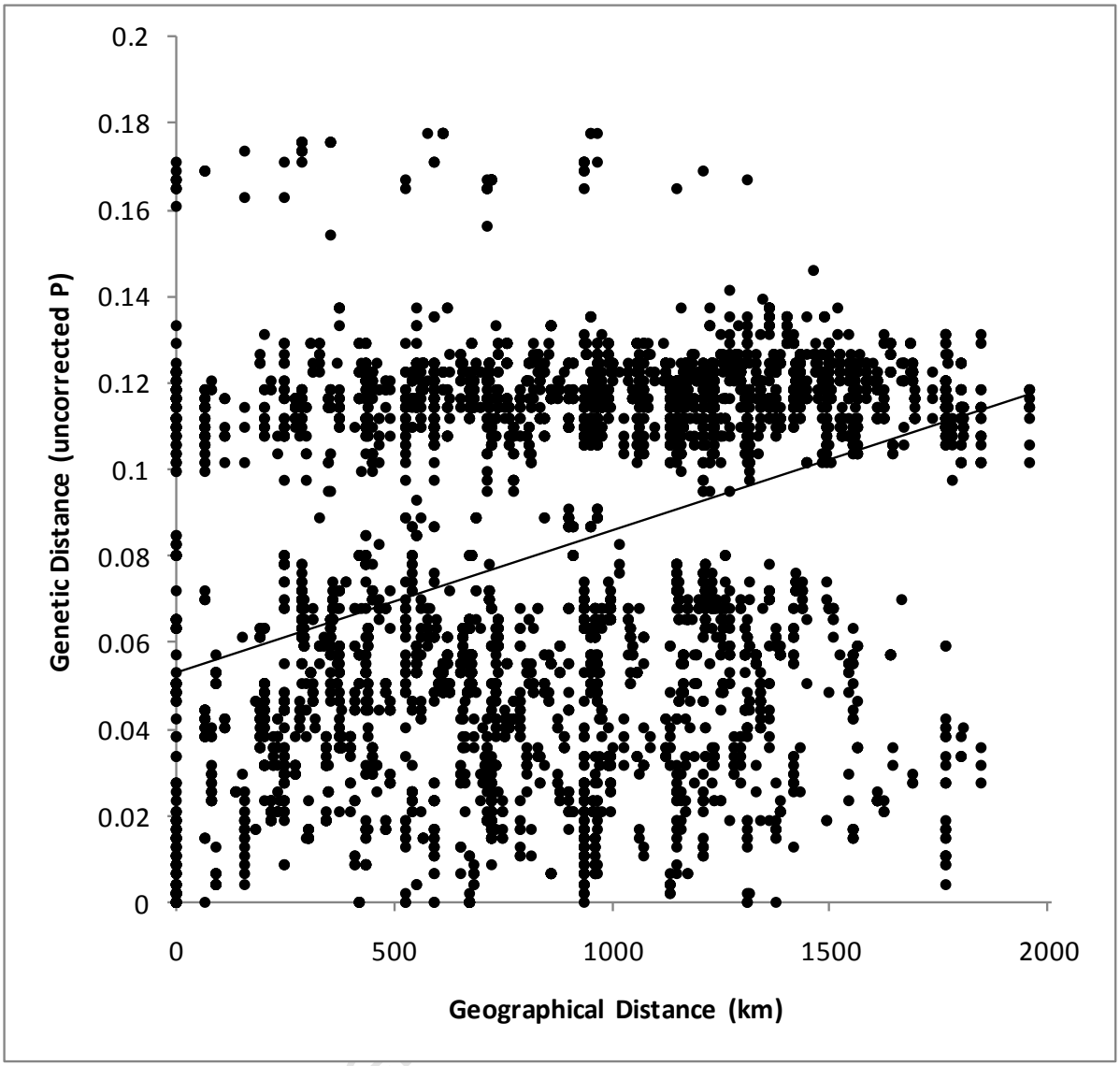

Figure 6.6a: Mantel test regression of genetic distance against geographic distance for 132 d-loop sequences; r =0.38; p = 0.0009.

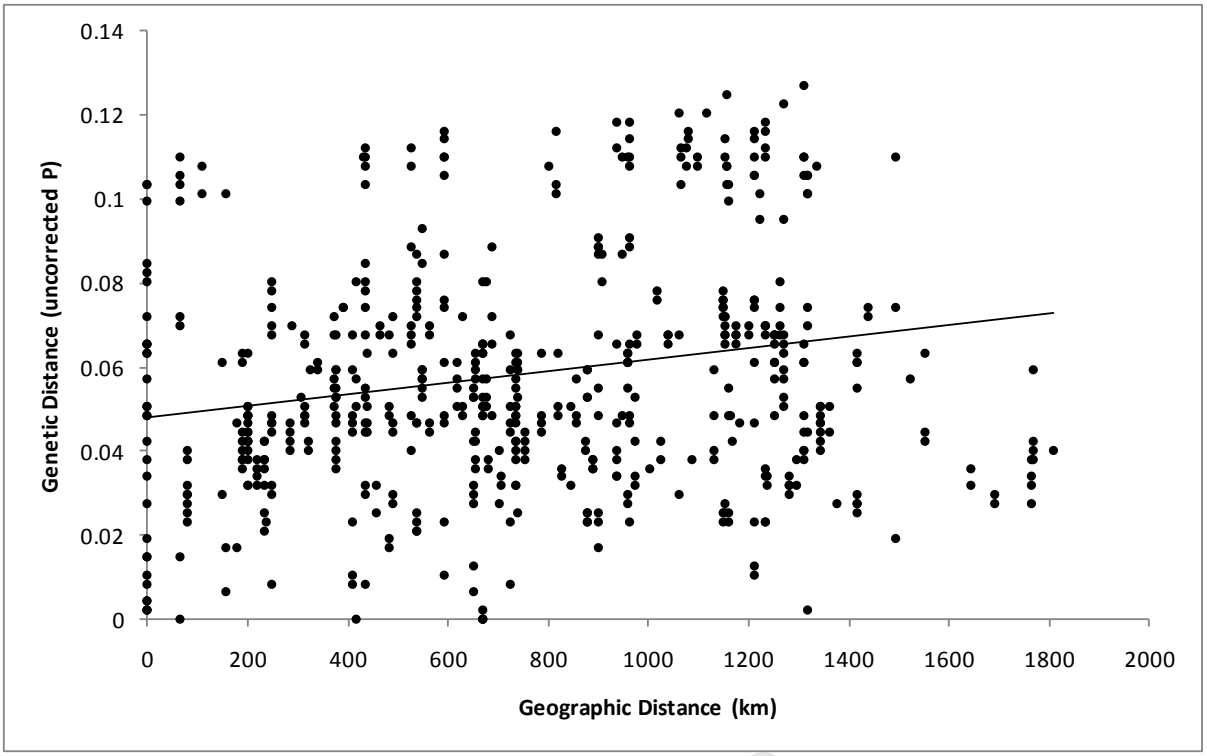

Figure 6.6b: Mantel test regression of genetic distance against geographic distance for for NgP: r = 0.21. p =0.0009.

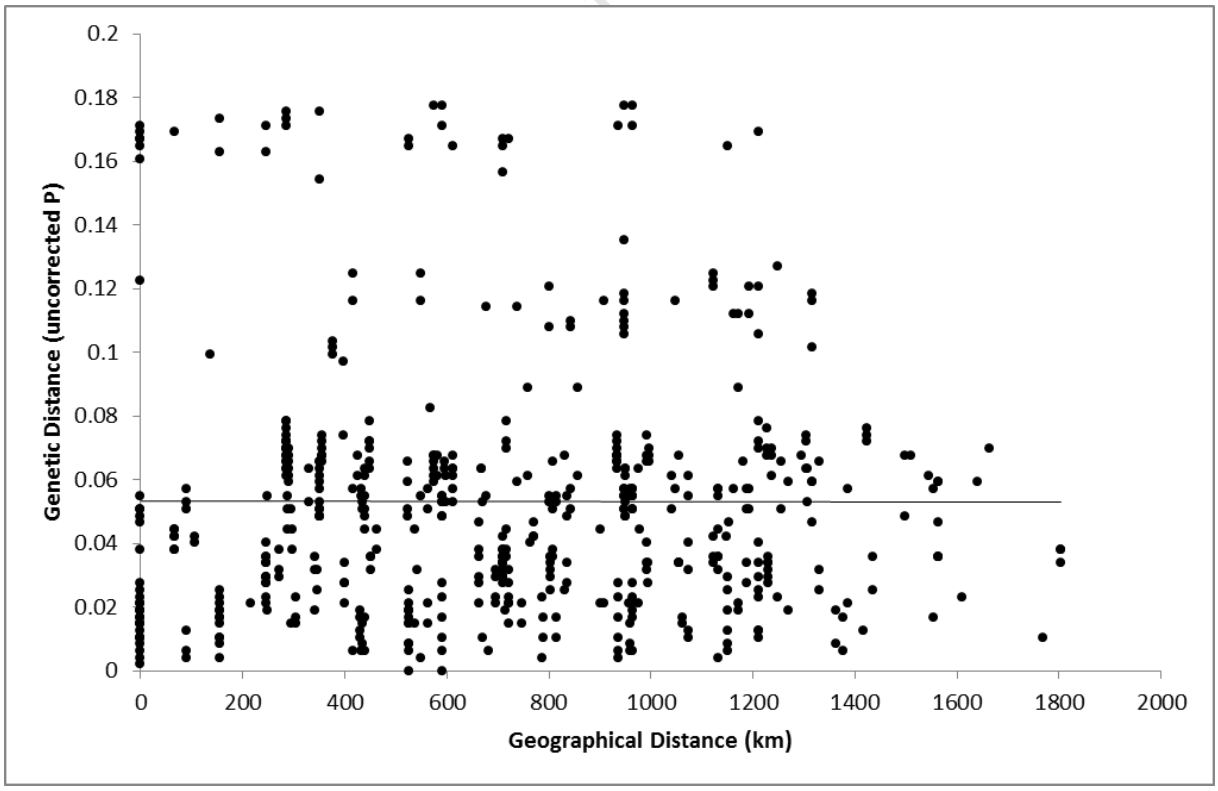

Figure 6.6c: Mantel test regression of genetic distance against geographic distance for for SgP, r = -0.1, p= 0.7.

Scenario 4.2- A test for signatures of ecologically limited dispersal

It is likely that environmental factors such as local resource availability play a significant role in constraining baboon dispersal and, as a result, in structuring baboon populations. In particular if male baboons choose dispersal routes based on the availability of familiar resources this will influence ecologically defined mitochondrial structuring within populations (Ballard and Dean 2001; Hiraoka and Hirao 1988; Taanman 1999). The recently developed genealogical sorting index (GSI) (Cummings et al. 2008) is a statistic that estimates the degree to which exclusive ancestry of individuals can be accounted for due to some variable, i.e. it provides an estimate of the extent to which individuals allocated to a user defined group e.g. biome, are monophyletic. This statistic then allows the user to estimate the contribution of that variable to the structure that is observed in the tree.

The GSI was developed to investigate the continuum of lineage sorting between its two endpoints, monophyly and complete paraphyly, and is aimed at furthering our understanding of the relationships between microevolutionary processes and phylogeny (Cummings et al. 2008). The degree of genealogical sorting (gs) for any group of terminal branches on a phylogenetic tree can be calculated. It is defined as "*the minimum number of nodes on a fully resolved tree required to unite a group, divided by the number of nodes actually uniting the group. Thus, the numerator represents the fully exclusive case (i.e., monophyly), the denominator represents the observed amount of exclusivity, and the quotient of these terms is a measure of relative exclusivity*" (Cummings et al. 2008, pg 2412). The maximum possible *gs* value for any group is therefore 1, reached when a defined group is monophyletic. The minimum would be the value for gs if all the nodes on the tree were required to unite the group under investigation. The gs value is normalised to make it possible to compare groups and trees and this normalisation generates the GSI:

$$GSI = \text{observed } gs - \min(gs) \,/\, \max(gs) - \min(gs).$$

Here I use this method to assess the contribution of each of three landscape factors to the mitochondrial structure that is observed in the South African baboon population. Tests are performed on a subset of the sample and include South African baboons **only** due to the availability of detailed landscape data for this political region. Although the D-loop evolves fairly rapidly on an evolutionary time scale, it is unlikely to be very useful when trying to identify differentiation at the decadal scale, especially given the sampling resolution of this study. This assessment is therefore performed at the broad regional scale.

University of Cape Town

The GSI analysis requires a user defined genealogical tree on which to overlay the grouping assignments. I therefore used D-loop sequence data to construct a Maximum Parsimony tree of all South African individuals characterised by unique haplotype and locality data combinations. An MP tree was generated in Mega v 4.0 (Tamura et al. 2007). MP analysis was run with CNI search factor of 2; a mini-heuristic search with level 100 was used to find the most parsimonious tree. Branch support was tested by 1000 bootstrap replicates and the final tree was converted into newick file format and uploaded to the GSI website (Cummings 2005; http://www.genealogicalsorting.org) on which all analyses was performed. The MP tree is presented in Fig 6.7.

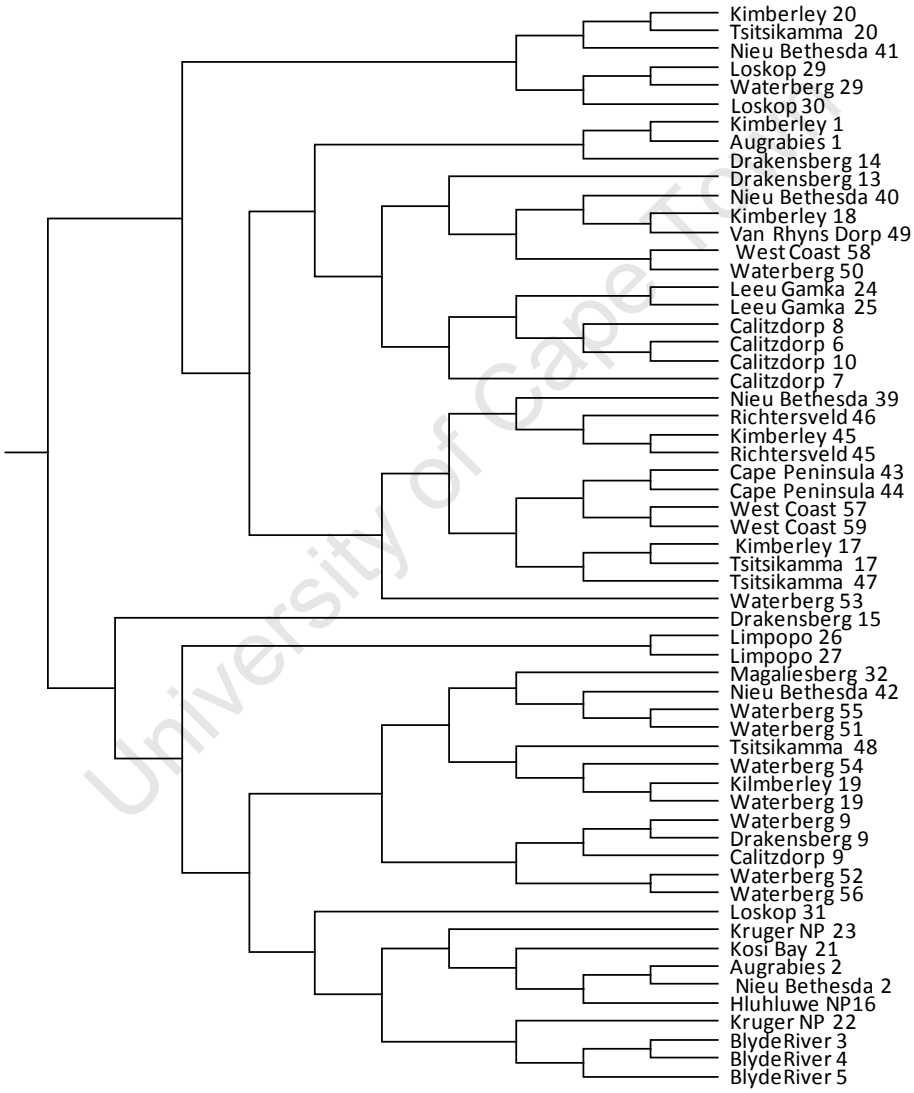

Fig 6.7-Maximum Parsimony tree of D-loop sequences of South African chacma baboon individuals. Each haplotype used to construct the tree is identified by the locality and haplotype number as tabled in Appendix 6A.

University of Cape Town

Using three landscape variables, I divided the sampling landscape into discrete categories. Individuals were then assigned to one of these categories and the GSI value for each grouping category was calculated. Once the GSI for a group as calculated, its statistical significance was estimated using permutation testing. In this context the permutation test holds the tree constant and permutes the group labels assigned to the tips of the tree (Maddison and Slatkin 1991), thus randomizing the common ancestry of members of the groups. A GSI value is calculated for each permutation and a distribution of values for GSI is generated. The frequency of GSI values equal to or greater than that which we observed from the original labelled tree is the p-value. In this case a significant p- value indicates that *GSI* values equal to or greater than the observed values are unlikely to be observed by chance alone (Cummings et al. 2008) and are therefore biologically meaningful. GSI values from 10000 permutations are shown in Table 6.5 and group assignments are tabled in Appendix 6A.

Variables

The Kalahari Desert poses a barrier to the distribution of chacma baboons in southern Africa. The Kalahari Basin is a low lying plateau in the interior of southern Africa and, although it sits at elevations between 600 and 1,600 m, it is essentially flat. Access to water and suitable shelter are also necessary to create a suitable baboon habitat. Sleeping site selection is an important part of baboon predator avoidance strategies and rocky outcrops with sheer faces are preferred (Cowlishaw 1997). The distribution of baboons across any landscape can therefore be limited by access to sleeping sites (Cowlishaw 1997) it is possible that part of the reason that the Kalahari region does not support baboons is that it lacks this fundamental baboon resource. It is my proposition then that both the shape of the landscape i.e. topographic features, and the distribution of resources on it, plays an important role in shaping contemporary gene flow patterns in baboons.

Here I also assess role of large landscape features and the distribution of resources on them, in shaping genetic structure in the South African baboon using three variables, biome, topography and drainage. An effective foraging strategy is fundamental to an animal's success (Hill and Dunbar 2002). As such foraging skills are taught and learnt early on, a model of ecological imprinting predicts that baboons will disperse into environments of familiar resources (Rodman 1995). It proceeds then that feeding ecology significantly influences the dispersal ecology of a species. As baboons are primarily vegetarian, vegetation type is a reasonable proxy for characterising feeding habitats. To test a role for vegetation in shaping the dispersal ecology, and thereby the distribution of mitochondrial

University of Cape Town

groups in chacma baboons I have plotted samples on a map of the vegetation biomes as defined by Mucina and Rutherford (2006) and each individual is assigned to a biome category (Fig 6.8a).

Topographically speaking, South Africa can be described as a central plateau that is bound by the main fold mountain ranges to the east, west and south. This rim of mountains falls steeply to the coasts. The major mountain chains of southern Africa are not continuous and this topographic relief of South Africa defines a number of important watersheds between high and low lying areas. Here I test the role of the distribution of mountain chains in South Africa on regional dispersal and gene flow in shaping local genetic structure of baboons. Haplotypes are plotted on a map of mountain ranges and each individual is assigned to a highland category accordingly (Fig. 6.8b). I also test the association between drainage basins and haplotype clustering by plotting samples on a map of the drainage basins of South Africa (Fig 6.8c). The coastal drainage system is further divided in southern, eastern and western systems.

University of Cape Town

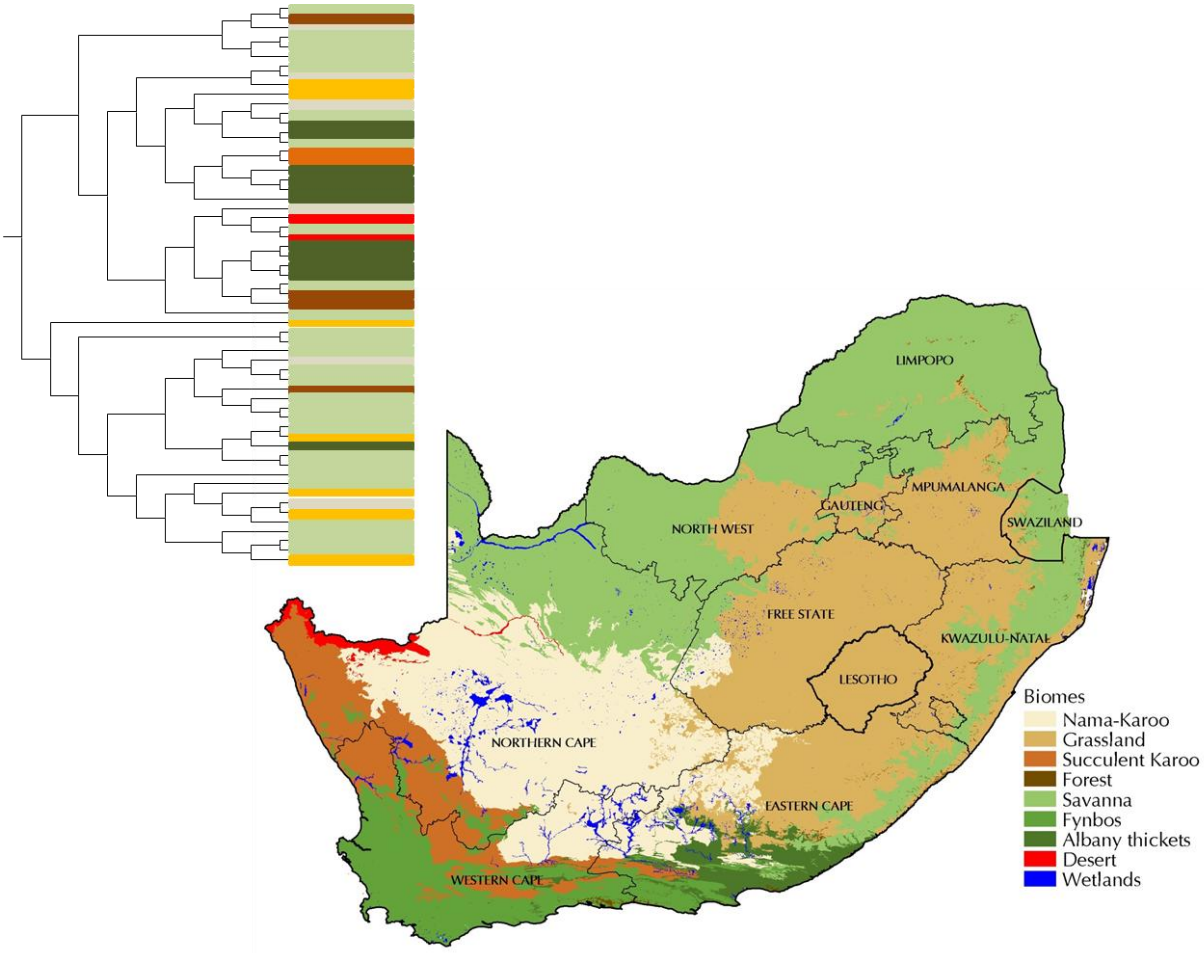

Figure 6.8a- Map of the biomes of South Africa (Mucina and Rutherford 2006). The tree insert is the MP tree (Fig 6.7) with individuals colour coded according to the biome from which they were sourced.

University of

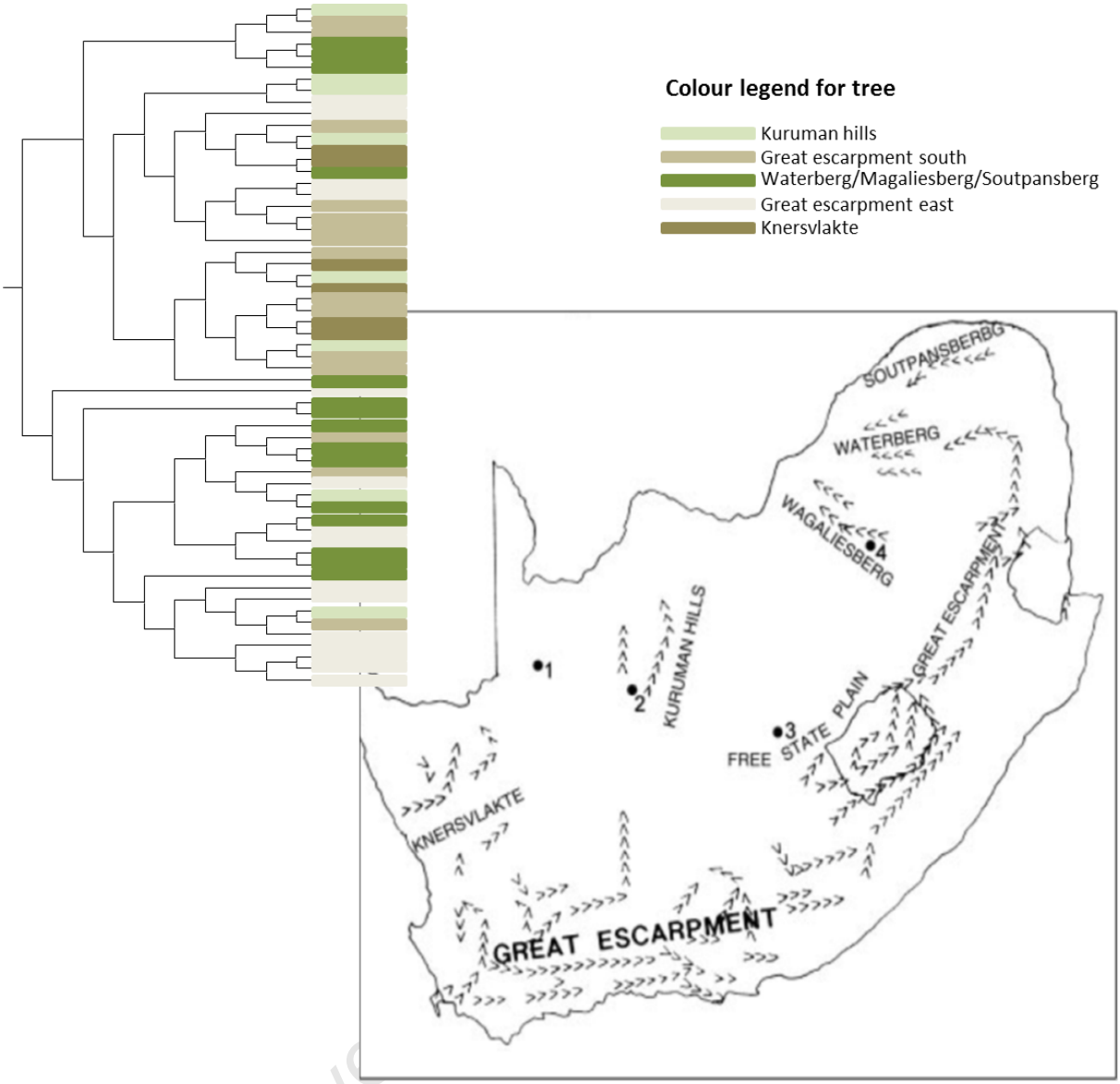

Figure 6.8b- Map of the distribution of South African mountain range topography (Mathee and Flemming 2002). The tree insert is the MP tree (Fig 6.7) with individuals colour coded according to the highland region from which they were sourced.

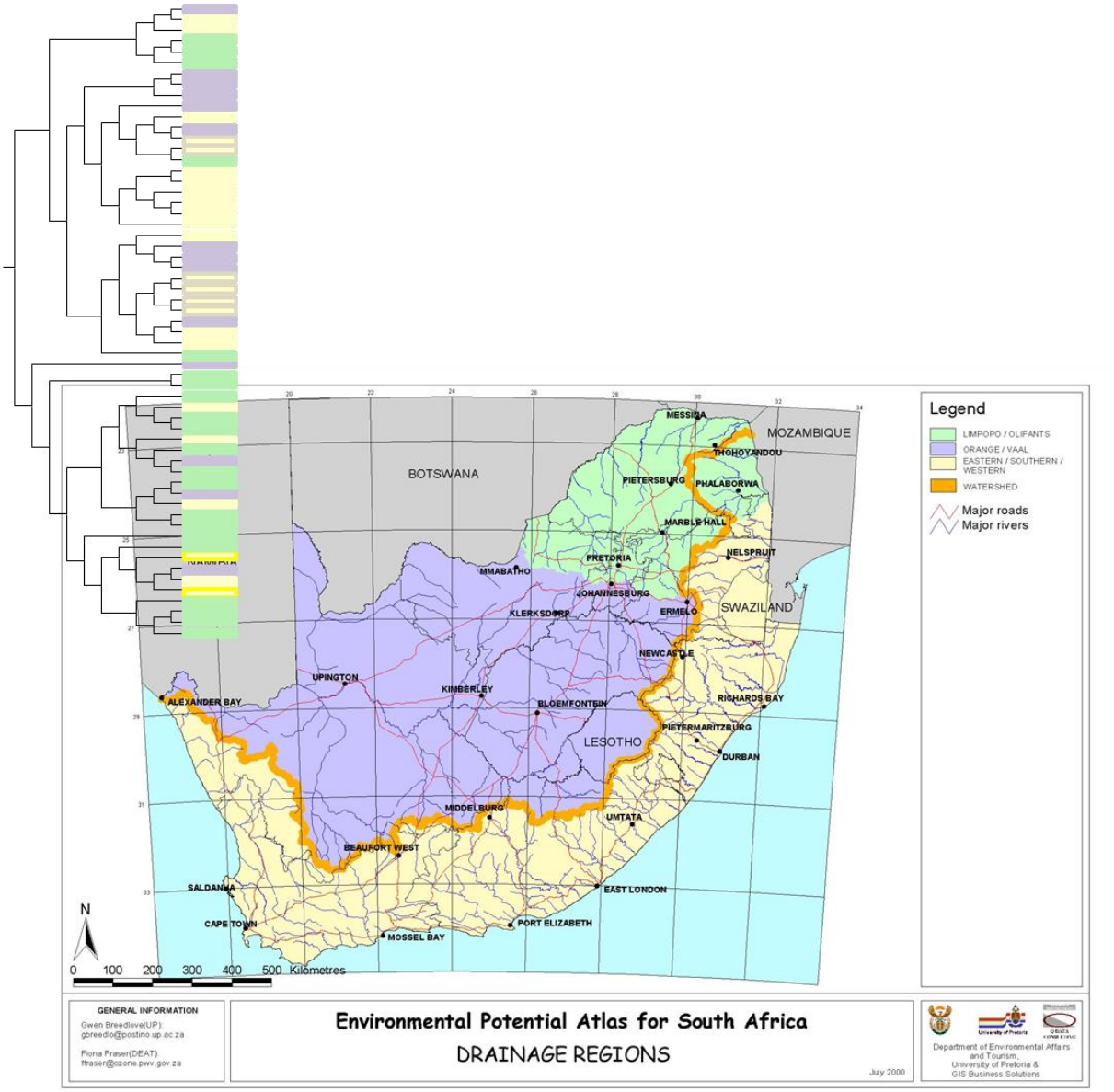

Figure 6.8c- Map of the main drainage basins of South Africa (source: The Dept of environmental affairs, SA). The tree insert is the MP tree (Fig 6.7) with individuals colour coded according to the drainage region from which they were sourced.

Results and interpretations

GSI values for catergories within all 3 variables were fairly low. Within biomes the highest value was 0.49 for individuals associated with the Desert biome but this is probably because there are only two individuals within the groups. Excluding this group, GSI values for all other biome catergories average ~0.2. Although certain groupings in each of the other two variable are slightly higher, e.g. Waterberg, Magaliesberg, Southpansberg = 0.348; Great escarpment East = 0.344 in the highland region category and Limpopo = 0.371 in the drainage basin category; overall both variables also average out to ~0.2. The results of the GSI analysis are somewhat inconclusive.This may well be due to the fact that I have

investigated population structure using a mitochondrial marker in a male dispersing species where migration is unlikely to leave a lasting signature. Nuclear markers may be more appropriate for this analysis.

BIOME			
Group	**# individuals**	**GSI value**	**p value**
savannah	23	0.213	0.078
succulentkaroo	11	0.208	0.088
NamaKaroo	7	0.123	0.699
grassland	9	0.261	**0.009**
fynbos	7	0.219	**0.035**
desert	2	0.491	**0.029**
HIGHLAND REGION			
Kuruman	7	0.100	0.933
Great escarpment, South	17	0.171	0.311
Waterberg, Magaliesberg, Southpansberg	16	0.348	**<0.001**
Great escarpment, East	13	0.344	**<0.001**
Knersvlakte	6	0.247	**0.016**
DRAINAGE BASIN			
Orange/Vaal	13	0.159	0.430
Coast South	17	0.185	0.215
Limpopo	21	0.371	**<0.001**
Coast West	6	0.296	**0.004**
Coast East	2	0.491	**0.028**

Table 6.6- GSI estimates for each of three regional landscape variables assessed above.

University of Cape Town

DISCUSSION

This discussion outlines a model for differentiation and diversification that is likely to have contributed to the pattern of phylogenetic structure observed in chacma baboons today. Although this model is built around the basic arguments of Jolly's (2001) original model for climate driven diversification for the genus *Papio*, it also incorporates more recent molecular and landscape data from both the literature and the results of this thesis to generate a hypothetical framework, on which this diversification process is overlaid.

The rise of a generalist template

The emergence and success of *Papio* in the Plio-Pleistocene coincides with the rise and success of many other generalist mammalian lineages, including *Homo* (Bowers and Harris 1994; Cain et al. 2000; Clark et al. 1998; Cwynar and MacDonald 1987; Potts 1996, 1999 Skellam 1951). This success has been attributed to the unpredictability in habitat change at the time which, according to Potts (1996, 1999), by and large favours selection for the generalist strategy. During the Plio-Pleistocene, Milankovitch oscillations created a period of particularly variable climate; shifts between warm and cool periods drove environmental and habitat change. In direct opposition to the idea that climate oscillations increase diversity through range fragmentation (Clarke and Crame 1997; Haffer 1969, 1997; Valentine 1984; Vrba 1992), Jansson and Dynesius (2002) proposed that in fact they reduce diversity by selecting for vagility and generalism in species. This trend towards generalism, in the long term underlies and, to some degree shapes evolutionary processes acting on shorter time scales (Bennet 1997). Given the rate of climate change for the time period, this is particularly relevant to the Pleistocene, as Bowers and Harris (1994) have shown that in constant or slowly changing environments the specialist species is favoured due to enhanced competitive abilities, but with moderate or rapid environmental change, as was experienced in the Pleistocene, generalist species are favoured over specialists by being able to track the shifting habitats.

Papio is considered to be a savannah or dry adapted primate (Jolly 1993) while the extinct *Theropithecus oswaldii* (2.0 – 1.5 Ma) was often associated with more well-watered environments (Jolly 1972). The rise of *Papio* and subsequent demise of *Theropithecus* in southern Africa has therefore been attributed to the overall drying out of southern Africa during the Plio-Pleistocene (Jablonski 1993; Jolly, 1993). This drying is linked to a global cooling trend which began ~3.0 to 2.5 Ma, and culminated in large scale glaciations at the poles and northern high latitudes ~ 2.4 Ma (Ruggiero 1994; Shackleton et al.1984; Thunell and Williams 1983). For southern Africa these glaciations had considerable impact on precipitation causing southern Africa to dry out (Meadows 2001). The subsequent changes in habitat are thought to have been unfavourable to *Theropithecus sp.* and led to localized extinction of the genus, thus allowing *Papio* to become the dominant Papionin in southern Africa (Jablonski 2002).

The current distribution of baboons (Fig. 2.6) indicates that while they may be largely distributed in savannah environments (Alberts and Altmann 2005) they are by no means restricted to them, and consequently today are regarded as niche generalists (Devore and

University of Cape Town

Washburn 1965). This characteristic of *Papio* supports an alternative explanation for the demise of *Theropithecus* in southern Africa. That is, that the large scale climate fluctuations characteristic of the Plio-Pleistocene favoured the generalist, which baboons clearly are, over the specialists like *Theropithicus oswaldii*, so that the success of *Papio* was not the *result* of localized extinction of *Theropithecus*, but rather the cause of it. This is based on the assumption that the ancestor to all *Papio* was in fact an ecological generalist, an assumption that is supported by isotopic data which shows that dietary variability within *P. robinsoni*, the most likely ancestor of modern baboons (Delson 1992; Heaton 2007; Lee Thorp and van de Merwe, 1993), was extremely high (Codron 2003).

The relationship between the genetic structure recovered for chacma baboons and modern landscape variables provide further support for the species as an ecological generalist (Devore and Washburn 1965). Tests for isolation by distance suggest that genetically, the chacma baboon is relatively continuous across its wide distribution. This is likely facilitated by the generalist feeding strategy employed by chacma baboons, allowing them to readily adapt to new resources so that males, and indeed entire troops, can disperse with relatively few local limitations. The fact that there is little correlation between biome and genetic boundaries in the GSI analysis further emphasizes this point. It would seem, at least from the results presented here, that while the shape of the landscape itself may play some role in maintaining differentiation within chacma baboons today, due to the generalist strategy employed by the species, the distribution of resources on this landscape does not.

Given the tendency of chacma baboons specifically, and *Papio* in general, towards geographic and genetic continuity, deep divergences are unexpected. Nonetheless habitat theory predicts that *all* organisms are vulnerable to changes in their habitat (Vrba 1992). As such they must respond to these changes, leading either to local extinction, or movement and adaptation to new available habitat. The second of these scenarios is more likely if the species is a generalist and highly adaptable, like baboons. In this instance one would predict lower rates of vicariance; however if these habitat changes are also accompanied by the generation of physical barriers to gene flow, this could lead to significant periods of independent evolution and eventual diversification.

Climate change generates landscape barriers to gene flow

Across Africa, baboons evolved and diversified within the Plio-Pleistocene, a period characterised by large scale cycles of colder and warmer climates. These cycles resulted in the expansion and retreat of ice at the poles (Denton 1985; Hays et al. 1976; Shackleton et al. 1984) and led to environmental change and habitat shifts around the globe. In the past

University of Cape Town

3.5myr there have been two major intensifications of the cyclic cold extremes. Global cooling from 3.0 Ma to 2.5 Ma culminated in the full glaciation of the poles and the onset of the 'modern ice age' (Shackleton et al. 1984). Signatures of this event have been documented around the world in both marine and terrestrial records (Thunell and Williams 1983; Shackleton et al.1984) and approximately 2.4 Ma, major ice sheets covered the northern hemisphere (Ruggiero 1994). As these ice sheets underwent repeated cycles of freezing and melting, they changed the volume of water in the oceans, resulting in fluctuations in sea levels. These ice sheets also affected the amount of atmospheric moisture, influencing rainfall regimes and other forms of precipitation and ultimately driving changes in the distribution of habitats around the globe.

Climate change drives habitat change and during the last 2.5 myr there have been ~ 20 glacial cycles that have affected the distribution of habitats in Africa. Any factor that reduces gene flow between two populations promotes diversification (and ultimately speciation) between them. This is particularly true for geographic barriers to gene flow e.g. forests or deserts (Ridley 1996). During cooler, drier periods, forests contract (Lawes 1990) allowing savannah corridors to open and thereby providing dispersal routes for savannah adapted species (while at the same time fragmenting and isolating forest adapted species). During wetter periods, forest expansions generate dispersal barriers to savannah adapted species and thereby lead to fragmentation of these lineages (Bonnefille et al. 1990; Hamilton and Taylor 1991; Maley 1996; Nicol 1999; Plana 2004; Zinner et al. 2009). Major climatic oscillations observed in the Pleistocene are therefore expected to correspond to lineage diversifications in many species (Vrba 1992; Hewitt 1996, 1999).

In 2009, Zinner et al. proposed a model for the diversification of *Papio* based on results from mitochondrial structure within the genus. Their findings support Jolly's (2001) original model and report a north-south split within baboons consistent with "Philopatry at the Frontier" hypotheses (Jolly 1993, 2001). From this the authors suggest that at ~2.1 Ma baboons dispersed north and south, out of southern Africa following the opening up of savannah corridors that resulted from global aridification. Once in the north, baboons dispersed over the complete northern savannah belt, from Eritrea to Senegal, and subsequent divergences are likely to be related to later expansion of forests, which, the authors propose, acted as barriers to gene flow between baboon populations. This scenario is essentially one of allopatric speciation i.e. any two populations separated for enough time will eventually diverge predominantly via genetic drift to the point of speciation. In contrast to the balance of forces required to produce sympatric or parapatric speciation, allopatric speciation requires only a degree of geographical isolation over time (Turelli et al. 2001) and the 'phenotypic

University of Cape Town

speciation' observed in baboons has been attributed to a series of such allopatric diversification events (Jolly 2001, Zinner et al. 2009). At 1.6 Ma chacma baboons experience an analogous population split attributed to the aridification of central southern Africa. In this case, the expansion of the Kalahari Desert is proposed as a geographic barrier to gene flow, isolating northern and southern chacma populations from each other and thereby contributing to the emergence of two distinct mitochondrial lineages within the species. Results of the IM analysis support this scenario of a period of complete or near complete isolation at ~1.3 Ma (0.8-1.7 Ma). When chacma baboons initially fragmented at 1.6 Ma both northern and southern populations likely shared many mitochondrial haplotypes, then after 300 kyr of independent evolution, lineage sorting reached completion. Based on the correlation between genetic and morphological structure within the species, it is likely that the differentiation of both characters was initiated by the same event and gave rise to southern *P.u. ursinus* and northern *P.u griseipes* phenotypes. This supports a strong positive relationship between mitochondrial and phenotypic diversification within *Papio.*

The role of landscape shape in structuring populations

A variable that has not been addressed in previous studies, but which may be significant in shaping baboons structuring, is that of landscape topography. I have already proposed that the distribution of rivers, which are known to shape structure in other primate species (Anthony et al. 2007; Arora et al. 2010), may have also affected baboons. In particular, the initial diversification of *P.u.ruacana* appears to be closely linked to the geomorphologic history of the Orange River. Today northern and southern chacma baboon lineages meet at the edges of the Kalahari Desert, yet the central Kalahari is still impenetrable to chacma baboons. The Kalahari ecoregion stretches across northwestern South Africa, southern Botswana and south eastern Namibia (Udvardy 1975). It is an intensely arid landscape characterised by extreme diurnal temperatures, infrequent rainfall (Lovegrove, 1993) and a nutrient poor substrate. Nonetheless the Kalahari ecoregion supports a number of large predator and ungulate species (Knight and Joyce 1997; Main 1987), and yet modern baboons are not found in the interior, suggesting that baboons avoid these extremely arid areas. Many Namibian baboons, however survive in extremely arid environments. The Tsoabis Leopard Park population in the Namib Desert (Davies and Cowlishaw 1997), for example, survives on transition vegetation (Geiss 1971) in a region where rainfall can be less than 85mm per annum (Colishaw 1997). Similarly baboons in the Kuiseb canyon are able to cope with an average annual rainfall of only 18mm (Colishaw 1997, Anderson, 1992). Evidently baboons can and do cope with extreme water stress, why then does the Kalahari Desert pose a barrier to dispersal of this species?

University of Cape Town

The Kalahari is a low lying plateau in the interior of southern Africa and is essentially flat. A useful predictor of where baboon troops can be found is the availability of appropriate sleeping sites (pers. field observations). Although troops have been reported to sleep in trees e.g. in the Tokai forest, on the Cape Peninsula (T. Hofmann pers. comm.), and in the Okavango Delta (pers. obs), they prefer to roost on rocky outcrops with sheer faces (Cowlishaw 1997). These provide a degree of protection from a number of predators, including leopards, when the troop is sleeping. Sleeping site selection is therefore an important part of the baboon survival strategy (Cowlishaw 1997). I propose then, that the Kalahari Desert does not serve as a barrier to the dispersal of baboons just because it is extremely arid, but that it also lacks an adequate supply of appropriate and broadly distributed sleeping sites. This suggests that topography itself may play a significant role in shaping the degree and direction of baboon diversification over relatively large spatial scales.

The fixation of morphological and behavioural differences outpaces the evolution of reproductive isolating mechanism

In the case of a widely distributed species becoming fragmented through habitat shifts and population contractions, gene flow will become reduced between highly dispersed populations (van Treuren et al. 1991; Young et al.1993). Because the now separated populations are only samples of the species, this process is usually accompanied by a decrease in intrapopulation variation and elevated levels of local inbreeding. Together these factors increase the effect of random genetic drift (Young et al. 1996) thereby driving diversification between populations. Climatic stress can also lead to the local extinction of demes within a species (Gilpin 1991; Raijmann et al. 1994) leading to further loss in diversity and even greater distances between surviving groups. The result is that in times of improved climate, the founding populations may be far removed from the mean genotype and phenotype pool of the ancestral population.

The process of genetic and phenotypic diversification is usually accompanied by the evolution of reproductive isolating mechanisms (Schluter 2001). In baboons there is no evidence for reproductive isolation, or at the least, the evolution of reproductively isolating mechanisms is incomplete, as evident in hybrid zones across the distribution of the genus in Africa (Nagel 1973; Jolly and Brett 1973; Sugawara 1979; Samuels and Altmann 1986; Phillips-Conroy et al. 1991, Jolly 1993; Alberts and Altmann 2001). This suggests that the speed at which phenotypic diversification has occurred in baboons outpaces the evolution of reproductive isolating mechanisms and is not unexpected for a phenotypically plastic, generalist organism (Pfennig 2010).

University of Cape Town

Van Valen (1965) hypothesised that niche expansion itself, in the absence of interspecific competition, is achieved by individual variation in the genome, i.e. allelic variation across a population. Selection for a generalist strategy therefore also selects for the accommodation of high levels of allelic variation within populations. This may have contributed to the speed at which the phenotypes differentiated within baboons. The existence of a high degree of allelic variation within a population implies that, for any given environment, there are differentially adapted individuals within that population. When such a population splits and the two resulting groups are exposed to different environments, the range of allelic variation within the groups, although less than in the original population is still significant. This increases the probability that there are already better adapted individuals within each group which will be favoured in new environments. Unequal reproductive success within groups shifts the mode of allelic frequencies away from the mode of the original population, and probably in different directions, so the two groups will begin to diverge from the ancestral group and each other (Avise et al. 1983; Avise 1988; Lynch 1987; Harrison 1991). Because selection operates on suites of traits, selection for one trait can influence selection on linked or correlated traits such as changes in behaviour and physiology (Bourdeau 2009; Cornwallis and Birkhead 2008) so that a whole suite of differences can become locally fixed, distinguishing the now divergent groups from each other.

The rate at which this divergence occurs may indeed be rapid because existing allelic variation buffers the new groups against extinction due to maladaptation (Pa´ and Miklos 1999; Price et al. 2003; Schlichting 2004; Pfennig et al. 2006). When an environment changes, a group can then return to optimal fitness, usually by the incremental accumulation of positive adaptations. But, if better adapted forms already exist in the groups, this allows the group to return to optimal fitness more quickly, simply by shifting the mode of allelic frequencies within it. Consequently, populations experiencing different environmental conditions or isolation can diverge rather rapidly (Pfennig 2010), and in the case of *Papio*, before reproductively isolating mechanisms have fully evolved. Chacma baboons are a case in point as no reproductive barriers are observed where individuals of the northern and southern lineages (pers. obs., Keller et al 2010).

Population contractions and expansions

Once climatic conditions become more favourableand geographic barriers reduce or disappear, it becomes possible for previously allopatric populations to disperse to points of contact. This can be achieved either through continuous population growth over time with groups slowly dispersing into newly available habitat until other populations are encountered

University of Cape Town

or, via sudden expansions in response to climatic amelioration. Sudden expansion of at least one of the groups will drastically reduce the time of genetic isolation and rapidly re-establish reproductive continuity between once isolated populations.

The current, continuous distribution of chacma baboons is linked to relatively recent population expansions. Bayesian skyline plots for both northern and southern populations indicate population growth over the last 1.3 Ma. For northern baboons this growth has occurred at a steady rate. Although the southern baboon population largely follows the same pattern, at ~15kya this population begins to expand relatively rapidly, a growth "spurt", that lasts for about 10kyr before the rate of expansion again slows down. Evidence for population expansions, by extension, can be taken as indirect evidence for preceding population contractions suggestive of glacial refuge use and accompanying genetic bottlenecks (Hewitt 1996).

During Pleistocene glacials, the inhospitable Kalahari was in direct contrast to the southern cape coast of South Africa. The contemporary south western cape (SWC) is a winter rainfall region which supports a highly diverse *fynbos* biome (Bond and Goldblatt 1984). A study of mammal communities in glacial and interglacial periods in the SWC shows little evidence for extreme habitat fluctuations between these climate phases (Rector and Reed 2010) suggesting that the SWC has been fairly stable at least 130kyr. In fact, although the bulk of southern Africa is dominated by summer rainfall regimes have been shown to cycle in synchronicity with global signatures of humidity (e.g. Partridge et al. 1997; Thomas and Shaw 2002), and evidence from SWC reveals that at least for the last glacial maximum (LGM) (Meadows and Baxter 1999) the SWC increased in humidity while the rest of southern Africa became increasingly arid. Lastly, the *fynbos* boasts the highest density of carbohydrate rich geophytic plant types of any region in the world (Cowling et al. 1997). Baboons are known to eat these geophytes when they are available (Cowling & Richardson 1995, Vincent 1985, G. Sampson, Zeekoevlei Archaeological Project, Dean and Milton, unpublished obs.). Together these factors could have made the SWC an attractive glacial refuge for baboons and the results of the spatial diversity analyses strongly suggest that this is a region into which baboons contracted during periods of harsh climatic conditions.

At approximately 15kya southern baboons expanded out of SWC to achieve a current female Ne ~260,000. Although the northern population is currently larger (female Ne 315,000) there is no evidence to suggest that this lineage underwent a rapid expansion event, further supporting independent histories of these populations in response to local conditions. The expansion of SL is very recent event and explains the low levels of mixing

University of Cape Town

observed between geographic lineages. This expansion coincides with the end of the LGM (18-17kya) which was a period of increased temperature and precipitation (Petit et al. 1999). Population expansion then, most likely began in response to landscape change further north, probably as the Kalahari Desert receeded. This period also marks the extinction of giant buffalo (*P. antiquus*) at ~14kya, suggesting significant environmental changes for the coastal areas at the time (Lewis 2008) which could have provided further impetus for seeking out new niches.

Long range dispersals

In Jolly's model (1993, 2001), diversification is largely linked to genetic bottlenecks resulting from population contractions, so that later founding populations represent a subset of variation of the ancestral population. This bottleneck effect can also be generated by long range dispersal events when small groups fission from the ancestral population and disperse in search of new habitats. The diversification of a north-eastern chacma clade recovered in both chapters 4 and 5, is interpreted here as resulting from long range dispersal during a period of global warming. This model assumes that NeC is generated from a minority fragment of NwC and so represents a founding or colonising group of the dominant clade. This model was tested but results of the IM analysis are inconclusive and further data is needed to assess the contribution of this variable, if any, to shaping the distribution of contemporary chacma baboon diversity.

The accumulation of variation during periods of climatic stability

Once further range expansion is no longer possible, and in the case of climatic stability, populations should approach a state of equilibrium. The ranges of a number of baboon species are parapatric and where species meet, introgression is likely to occur (Nagel 1973; Jolly and Brett, 1973; Sugawara 1979; Samuels and Altmann 1986; Phillips-Conroy et al. 1991; Jolly 1993; Alberts and Altmann 2001). There are however still at least five major baboon phenotypes (Jolly 1993) and this suggests that some mechanism of mating preference must exist at the core of each population (Van Valen 1965). This may be due to factors that cause groups to remain apart such as isolation by distance. Alternatively genetic differentiation may also result due to factors that promote within group cohesion such as adaptation to local habitat.

Dietary selection is fundamental to structuring the fine-scale ranging pattern of any primate implying that an animal's diet often defines the parameters of its habitat. Most primates therefore move within home ranges of reasonably consistent boundaries that last for multiple generations with little change (Rodmann 1999). Even in a generalist species ecological

University of Cape Town

imprinting may result in single individuals becoming specialised or behaviourally adapted to specific habitats according to early foraging experience (Bolnick 2005).

This process of ecological imprinting predicts non-random dispersal within a species as populations of con-specifics actively choose between different habitats. An individual with established preference can choose a new group more quickly. If the chosen habitat contains familiar resources, then that individual it is immediately fitter than it would be in a habitat, in which it would have to re-establish foraging patterns. Consequently ecological imprinting would provide a reproductive advantage to the individual. While this may lead to the dismissal of higher quality habitats, it also prevents dispersal into habitats with unknown potential, that may be ultimately limit reproductive success (Davis 2004). These choices maintain genetic variation across a heterogenous landscape and may result in population differentiation so that a species is composed of a mosaic pattern of populations with habitat linked differences (Immelman 1985). It is this variation on which selection can act during times of environmental change.

Unfortunately the molecular data collected for this thesis contains insufficient variation at the fine geographical scale to adequately address this issue. A fine scale geographic sample analysed using both rapidly evolving nuclear markers and mitochondrial markers will be considerably more informative about the habitat variables, if any, that may limit the dispersal of baboons during periods of climatic stability.

University of Cape Town

Conclusions

The pattern that is observed in chacma baboons today is one of a genetically and phenotypically differentiated taxon that has remained reproductively compatible across its entire distribution, very much like the pattern observed in *Papio* as a whole. To generate the reported patterns of phylogenetic complexity, population fragmentations are proposed to have been coupled with range contractions, periods of genetic isolation and possible genetic bottlenecks. Current population continuity is the result of both steady population growth and sudden population expansions during/over the last ~1.0 myr, and to some degree long range dispersal. This complex population history is closely tied to climate fluctuation and the data presented here provides evidence for the significant role of landscape change in shaping the evolution of this species, and baboons in general, over the last 2 million years.

Appendix 6A- Table of GSI assignments for each of three variables; biome, topography and drainage. MP tree ID's label each of the South African baboon D-loop sequences used to construct the Maximum Parsimony tree shown in Fig 6.7

MP tree ID	Biome	Topography	Drainage
Hluhluwe16	grassland	Great Escarpment East	CoastEast
KosiBay21	grassland	Great Escarpment East	CoastEast
LeeuGamka24	succulentkaroo	Great Escarpment East	CoastSouth
LeeuGamka25	succulentkaroo	Great Escarpment East	CoastSouth
BlydeRiver3	grassland	Great Escarpment East	Limpopo
BlydeRiver4	grassland	Great Escarpment East	Limpopo
BlydeRiver5	grassland	Great Escarpment East	Limpopo
Kruger22	savannah	Great Escarpment East	Limpopo
Kruger23	savannah	Great Escarpment East	Limpopo
Drakensberg13	grassland	Great Escarpment East	OrangeVaal
Drakensberg14	grassland	Great Escarpment East	OrangeVaal
Drakensberg15	grassland	Great Escarpment East	OrangeVaal
Drakensberg9	grassland	Great Escarpment East	OrangeVaal
Tsitsikamma17	forest	Great Escarpment South	CoastSouth
Tsitsikamma18	forest	Great Escarpment South	CoastSouth
Tsitsikamma20	forest	Great Escarpment South	CoastSouth
Tsitsikamma47	forest	Great Escarpment South	CoastSouth
Tsitsikamma48	forest	Great Escarpment South	CoastSouth
Calitzdorp10	fynbos	Great Escarpment South	CoastSouth
Calitzdorp6	fynbos	Great Escarpment South	CoastSouth
Calitzdorp7	fynbos	Great Escarpment South	CoastSouth
Calitzdorp8	fynbos	Great Escarpment South	CoastSouth
Calitzdorp9	fynbos	Great Escarpment South	CoastSouth
NieuBethesda2	NamaKaroo	Great Escarpment South	CoastSouth
NieuBethesda39	NamaKaroo	Great Escarpment South	CoastSouth
NieuBethesda40	NamaKaroo	Great Escarpment South	CoastSouth
NieuBethesda41	NamaKaroo	Great Escarpment South	CoastSouth
NieuBethesda42	NamaKaroo	Great Escarpment South	CoastSouth
Peninsula43	fynbos	Great Escarpment South	CoastWest
Peninsula44	fynbos	Great Escarpment South	CoastWest
VanRhynsDorp49	fynbos	Knersvlakte	CoastWest
WestCoast57	fynbos	Knersvlakte	CoastWest
WestCoast58	fynbos	Knersvlakte	CoastWest
WestCoast59	fynbos	Knersvlakte	CoastWest
Richtersveld45	desert	Knersvlakte	OrangeVaal
Richtersveld46	desert	Knersvlakte	OrangeVaal
Augrabies1	NamaKaroo	Kuruman	OrangeVaal
Augrabies2	NamaKaroo	Kuruman	OrangeVaal
Kimberley1	savannah	Kuruman	OrangeVaal
Kimberley17	savannah	Kuruman	OrangeVaal
Kimberley19	savannah	Kuruman	OrangeVaal
Kimberley20	savannah	Kuruman	OrangeVaal

`

University of Cape Town

Appendix 6A- Table of GSI assignments for each of three variables; biome, topography and drainage. MP tree ID's label each of the South African baboon D-loop sequences used to construct the Maximum Parsimony tree shown in Fig 6.7

MP tree ID	Biome	Topography	Drainage
Kimberley45	savannah	Kuruman	OrangeVaal
Limpopo26	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Limpopo27	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Loskop29	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Loskop30	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Loskop31	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Magaliesberg32	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Magaliesberg33	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Waterberg19	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Waterberg29	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Waterberg50	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Waterberg51	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Waterberg52	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Waterberg53	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Waterberg54	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Waterberg55	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Waterberg56	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo
Waterberg9	savannah	Waterberg/Magaliesberg/Soutpansberg	Limpopo

University of Cape Town

CHAPTER 7

SUMMARY AND CONCLUSIONS

The climate cycles of the Plio-Pleistocene and related vegetation shifts explain, to a significant degree, the distribution and diversity of many mammal lineages globally. These climatic fluctuations have also played a major role in the evolution of Old World monkeys, including *Papio*. The proposed model of diversification suggests that the pattern of diversity observed within modern chacma baboons is also closely linked to these climate cycles. The aim of this thesis was to assess how well this model fits the contemporary genetic structure observed within this species. This aim was achieved by testing specific hypotheses related to climate and landscape driven structuring of the southern African chacma baboon, through the application of phylogenetic and phylogeographic techniques.

The results of this thesis have shown that chacma baboons in southern Africa evolved approximately 2 Ma. Analyses of two mitochondrial markers have allowed for the detailed reconstruction of a possible population history for the species over the past two million years. Several major diversification events within the species are linked both spatially and temporally to major climate driven landscape change events. This supports a hypothesis of climate driven diversification within modern chacma baboons, and suggests that similar factors may have affected the genus more broadly. Coalescent modelling further confirms that landscape changes have produced periods of genetic isolation between baboon populations resulting in vicariance. A degree of association between mitochondrial and phenotypic clustering is also found, from which it is inferred that differentiation in both sets of characters is linked. The timeline of diversification also suggests that significant amounts of phenotypic diversification can be generated within baboons with no decrease in the reproductive compatibility of lineages even after 300 kyr.

Local geography is also shown to be an important variable in shaping structure within chacma baboons across the distribution of the species. The reconstruction of population dynamics identifies an event of population contraction and expansion, indicative of glacial refuge use for baboons. In particular, the south-western region of South Africa is identified as a potentially important region for sheltering past chacma populations during periods of hostile climatic conditions. Because the last 2myr is thought to represent ~20 glacial interglacial cycles, it is likely that a pattern of population contraction followed by expansion in chacma baboons repeated itself many times during this period. Considered in light of the phenotypic variation observed within chacma baboons today, it is possible that these

University of Cape Town

phenotypic variants represent refugial groups that were able to survive these cycles and whose survival was largely attributable to hospitable microhabitats created by local geographic features.

While this study has contributed significantly to our understanding of ecologically linked diversification in a large bodied, generalist primate, the findings also highlight several issues for further research. First, the study has shown the power of high resolution taxon sampling to inform on quite specific events shaping the evolutionary history of baboon species. This suggests that high resolution sampling of each of the other baboon species is required to generate a more accurate picture of the events shaping the evolution of *Papio* as a whole. As mitochondrial markers represent only the maternal history of the species, there are clear limitations to their application to understating population dynamics in a strictly male dispersing taxon such as *Papio,* there is therefore clearly also a need to add nuclear markers to these analyses.

Second, the results here hint that long range dispersals may also play a role in shaping genetic structuring and diversity within baboons. Unfortunately the data collected in this study does not capture sufficient population variability to satisfactorily test this proposal. This could be resolved with intensive sampling of baboons from the north-eastern and north-western chacma clades.

Third, the D-loop data has not been particularly informative on issues around contemporary population differentiation at the regional scale. It is likely that a combination of higher resolution sampling of a few target populations, as well as the use of faster evolving markers such as microsatellites, may prove useful in addressing how the shape of the landscape, and the distribution of resources on it, may or may not constrain gene flow between baboon populations.

Finally, there have been several references made in this thesis to a direct correlation between mitochondrial and phenotypic structure within chacma baboons. However, this is based on subjective measures of phenotypic diversity within chacma baboons and I would propose that a systematic assessment of morphological variation across the range of chacma baboons would prove useful in furthering our understanding of the relationship between genotype and phenotype.

University of Cape Town

A

Acocks, J. (1988). Veld Types of South Africa, 3rd Edition. *Memoirs of the Botanical Survey of South Africa* 57:146.

Ackermann, R.R., Bishop, J.M. (2010). Morphological and molecular evidence reveals recent hybridization between gorilla taxa. *Evolution* 64: 271-290.

Akaike, H. (1974). A new look at the statistical model identification*. IEEE Transactions on Automatic Control* 19 (6): 716–723.

Akaike, H (1980). *Likelihood and the Bayes procedure*. Bayesian Statistics (ed) J.M. Bernardo et al., Valencia: University Press. pp 143-166.

Alberts, S.C., Altmann, J. (2001). Immigration and hybridization patterns of yellow and anubis baboons in and around Amboseli, Kenya. *American Journal of Primatology* 53(4):139-154.

Alberts, S.C., Hollister-Smith, J.A., Mutata, R.S., Saiyalel, S.N., Muruthi, P., Warutere, J.K. Altmann, J. (2005). Seasonality and long-term change in a savanna environment. In: Brockman DK, van Schaik CP (eds) *Seasonality in Primates: Studies of living and extinct human and non- human primates*. Cambridge University Press.

Albrecht, G.H, Miller, J.M.A. (1993). Geographic variation in primates: a review with implications for interpreting fossils. In: Kimbel WH, Martin LB (eds) *Species, species concepts and primate evolution*. Plenum Press, New York, pp 123-151.

Aldrich-Blake, K., Richtsmeier, J.T. (1999). Patterns of craniofacial variability in the three savanna baboon groups. *American Journal of Physical Anthropology* Suppl 28: 82-83.

Altmann, S., Altmann, J. (1970). *Baboon Ecology: Africa Field Research*. University of Chicago Press, Chicago.

Altmann, S. A. (1998). *Foraging for Survival: Yearling Baboons in Africa*. University of Chicago Press, Chicago.

Altmann, J., Alberts, S.C., Altmann, S.A., Roy, S.B. (2002). Dramatic change in local climate patterns in the Amboseli basin, Kenya. *African Journal of Ecology* 40:248-255.

Anderson, C.M. (1982). Baboons Below the Tropic of Capricorn. *Journal of Human Evolution* 11: 205-217.

Anderson, C.M. (1992). Male investment under changing conditions among chacma baboons at Suikerbosrand. *American Journal of Physical Anthropology* 87: 479–469

Anthony, N.M., Clifford, S.L., Bawe-Johnson, M., Abernethy, K.A., Bruford, M.W., Jean Wickings, E. (2007). Distinguishing gorilla mitochondrial sequences from nuclear integrations and PCR recombinants: Guidelines for their diagnosis in complex sequence databases. *Molecular Phylogenetics and Evolution* 43(2):553-566.

University of Cape Town

Anthony, N.M., Johnson-Bawe, M., Jeffery, K.J., Clifford, S., Abernethy, K., Tutin, C.E.G., Lahm, S., White, L.J.T., Utley, J., Wickings, E., Bruford, M.W. (2007a). The role of Pleistocene refugia and rivers in shaping gorilla genetic diversity in central Africa. *Proceeding of the National Academy of Sciences, USA* 104: 20432-20436.

Arbogast, B.S., Edwards, S.V., Wakeley, J., Beerli, P., Slowinski, J.B. (2002). Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review Ecology and Systematics* 33: 707–740.

Arctander, P., Johansen, C., Coutellec-Vreto, M. (1999). Phylogeography of three closely related African bovids (Tribe *Alcelaphini*). *Molecular Biology and Evolution* 16(12):1727-1739.

Arora, N., Nater, A., van Schaik, C.P., Willems, E.P., van Noordwijk, M.A., Goossens, B., Morf, N., Bastian, M., Knott, C., Morrogh-Bernard, H., Kuze, N., Kanamori, T., Pamungkas, J., Perwitasari-Farajallah, D., Verschoor, E., Warren, K., Krutzen, M. (2010). Effects of Pleistocene glaciations and rivers on the population structure of Bornean orangutans (*Pongo pygmaeus*). *Proceedings of the National Academy of Sciences, USA* 107(50): 21376-21381.

Ascunce, M.S., Hasson, E., Mulligan, C.J., Mudry, M.D. (2007). Mitochondrial sequence diversity of the southernmost extant New World monkey, *Alouatta caraya*. *Molecular Phylogenetics and Evolution* 43(1): 202-215.

Atkinson, Q.D., Gray, R.D., Drummond, A.J. (2009). Bayesian coalescent inference of major human mitochondrial DNA haplogroup expansions in Africa. *Proceedings of the Royal Society B: Biological Sciences* 276 (1655): 367-373.

Avise, J.C., Shapira, J.F., Daniel, S.W., Aquadro, C.F., Lansman, R.A. (1983). Mitochondrial DNA differentiation during the speciation process in Peromyscus. *Molecular Biology and Evolution* 1: 38 –56.

Avise, J.C. (1986). Mitochondrial DNA and the evolutionary genetics of higher animals. . *Philosophical Transcripts of the Royal Society of London B* 312:325-342.

Avise, J.C. (1991). Ten unorthodox perspectives on evolution prompted by comparative population findings on mitochondrial DNA. *Annual Review in Genetics* 25: 45-69.

Avise. J.C., Walker, D. (1998). Pleistocene phylogeographic effects on avian populations and the speciation process. *Proceedings of the Royal Society of London B* 265: 457–463.

Avise, J.C. (2000). *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge.

Avise, J.C. (2004). *Molecular Markers, Natural History and Evolution*. 2nd edn. Sinauer, Sunderland, Massachusetts.

Avise, J.C. (2009). Phylogeography: retrospect and prospect. *Journal of Biogeography* 36*:* 3-15.

Awadalla, P., Eyre-Walker, A., Smith, J.M. (1999). Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science* 286 (5449): 2524-2525.

University of Cape Town

B

Ballard, J.W.O., Dean, M.D. (2001). The mitochondrial genome: mutation, selection and recombination. *Current opinion in Genetics and Development* 11(6):667-672.

Bandelt, H.J., Dress, A.W.M. (1992a). A canonical decomposition theory for metrics on a finite set. *Advances in Mathematics* 92:47-105.

Bandelt, H.J, Dress, A.W.M. (1992b). Split Decomposition: A new and useful approach to phylogenetic analysis of distance data. *Molecular Phylogenetics and Evolution* 1(3):242-252.

Bandelt, H.J., Dress, A.W.M. (1994). An order theoretic framework for overlapping clustering. *Discrete Mathematics* 136(1-3):21-37.

Banks, S.C., Hansen, B.D., Robinson, N.A., Taylor, A.C. (2002). Wombat coprogenetics: enumeration a common wombat population by analysis of faecal DNA. *Australian Journal of Zoology* 50:193-204.

Barraclough, T.G. (1998). Revealing the factors that promote speciation. *Philosophical Transcripts of the Royal Society of London B* 353: 241-249

Barraclough, T.G. (2006). What can phylogenetics tell us about speciation in the Cape flora? *Diversity and Distributions* 12:21-26.

Barrett. L., Henzi, S., Weingrill, T., Lycett, J., Hill, R. (2000). Female baboons do not raise the stakes but they give as good as they get. *Animal Behaviour* 59:763-770.

Barrett, L., Gaynor, D., Henzi, S.P. (2002). A dynamic interaction between aggression and grooming reciprocity among female chacma baboons. *Animal Behaviour* 63(6):1047-1053.

Bazin, E., Glemin, S., Galtier, N. (2006). Population size does not influence mitochondrial genetic diversity in animals. *Science* 312 (5773):570-572.

Beaumont, M.A., Rannala, B. (2004). The bayesian revolution in genetics. *Nature Reviews Genetics* 5:251-261.

Beerli, P., Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences, USA* 98:4563 - 4568

Belay, G., Mori, A (2006). Intraspecific Phylogeographic Mitochondrial DNA (D-Loop) Variation of Gelada Baboon, *Theropithecus Gelada*, in Ethiopia. *Biochemical Systematics and Ecology* 34:554-561.

Bennett, K.D. (1997). *Evolution and Ecology: The Pace of Life*. Cambridge: Cambridge Univ. Press.

Bensasson, D., Zhang, D-X., Hartl, D.L., Hewitt, G.M. (2001). Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution* 16(6):314-321.

Bergman, T., Beehner, J.C., Cheney, D.L., Seyfarth, R.M. (2003). Hierarchical classification by rank and kinship in baboons. *Science* 302:1234-1236.

University of Cape Town

Berlocher, S.H. (1998). *Can sympatric speciation be proven from phylogenetic and biogeographic evidence?* In: Howard DJ, Berlocher SH (eds) In Endless Forms: Species and Speciation. Oxford University Press, pp 99-113.

Bermingham, E., Moritz, C. (1998). Comparative phylogeography: concepts and applications. *Molecular Ecology* 7: 367-369.

Bernatchez, L. and Wilson, C.C. (1998). Comparative phylogeography of nearctic and palearctic fishes. *Molecular Ecology* 7:431-452.

Bininda-Emonds, O.R.P. and Stamatakis, A. (2007). Taxon sampling versus computational complexity and their impact on obtaining the tree of life. *Reconstructing the Tree of Life: Taxonomy and Systematics of Species Rich Taxa* 72:77-95.

Boelhouwers, J.C., Meiklejohn, K.I. (2002). Quaternary periglacial and glacial geomorphology of southern Africa: review and synthesis. *South African Journal of Science* 98(1/2):47.

Bolnick, D. I., Pressier, E. L. (2005). Resource competition modifies the stregth of trait mediated predator-prey interactions: a meta-analysis. *Ecology* 86 (10) 2771–2779.

Bonnefille, R., Roeland, J.C., Guiot, J. (1990). Temperature and rainfall estimates for the past 40,000 years in equatorial Africa. *Nature* 346:347-349.

Bourdeau, P.E. (2009). Prioritized phenotypic responses to combined predators in a marine snail. *Ecology* 90: 1659–1669.

Bowers, M.A., Harris, L.C. (1994). A large-scale metapopulation model of interspecific competition and environmental change. *Ecological Modelling* 72:251-273.

Brain, C.K. (1981). Hominid Evolution and Climatic Change. *South African Journal of Science* 77:104-105.

Brain, C.K. (1992). Deaths in a desert baboon troop. *International Journal of Primatology* 13(6):593-599.

Braun, E.L., Kimball, R.T. (2002). Examining basal avian divergences with mitochondrial sequences: Model complexity, taxon sampling, and sequence length. *Systematic Biology* 51:614–625.

Bremer, B., Jansen, R.K., Oxelman, B., Backlund, M., Lantz, H., Kim, K.J. (1999). More characters or more taxa for a robust phylogeny—case study from the coffee family (*Rubiaceae*). *Systematic Biology* 48:413–435.

Bristow, D., Ward, C. (1988). *Mountains of Southern Africa*. Struik Publishers, Cape Town, South Africa.

Brooks, D.R., McLennan, D.A. (1991). *Phylogeny, Ecology, and Behaviour: a Research Program in Comparative Biology*. Chicago University Press.

Brown, W.M., Prager, E.M., Wang, A., Wilson, A.C. (1982). Mitochondrial DNA sequences of primates: tempo and mode of evolution. *Journal of Molecular Evolution* 18:225–239.

University of Cape Town

Brown, D.M., Brenneman, R.A., Koepfli, K.P., Pollinger, J.P., Mila, B., Georgiadis, N.J., Louis, E.E., Grether, G.F., Jacobs, D.K., Wayne, R.K. (2007). Extensive population genetic structure in the giraffe. *BMC Biology* 5.

Brouat, C., Tatard, C., Ba, K., Cosson, J.F., Dobigny, G., Fichet-Calvet,E., Granjon, L., Lecompte, E., Loiseau, A., Mouline, K., Piry, S., Duplantier, J.M. (2009). Phylogeography of the Guinea multimammate mouse (*Mastomys erythroleucus*): a case study for Sahelian species in West Africa. *Journal of Biogeography* 36(12):2237-2250.

Brouns, G., De Wulf, A., Constales, D. (2003). Delaunay triangulation algorithms useful for multibeam echosounding. *Journal of Surveying Engineering* 129:79-84.

Brown, W.M., Prager, E.M., Wang, A., Wilson, A.C. (1982). Mitochondrial DNA sequences of primates: tempo and mode of evolution. *Journal of Molecular Evolution* 18:225-239.

Brown, R.M., Linkem, C.W., Siler, C.D., Sukumaran, J., Esselstyn, J.A., Diesmos, A.C., Iskandar, D.T., Bickford, D., Evans, B.J., McGuire, J.A., Grismer, L., Supriatna, J., Andayani, N. (2010). Phylogeography and historical demography of *Polypedates leucomystax* in the islands of Indonesia and the Philippines: Evidence for recent human-mediated range expansion? *Molecular Phylogenetics and Evolution* 57(2):598-619.

Buettner-Janusch, J. (1966). A Problem in Evolutionary Systematics: Nomenclature and Classification of Baboons, Genus *Papio*. *Folia Primatologica* 4:288-308.

Buhay, J.E. (2009). "COI-like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. *Journal of Crustacean Biology* 29(1):96-110.

Buneman, P. (1971). *The recovery of trees from measures of dissimilarity.* In: Hodson, F.R., Kendall, D.G., Tautu, P. (eds). Mathematics in the Archaeological and Historical Sciences. Edinburgh University Press, pp 387-395.

Burrell, A.S. (2008). *Phylogenetics and population genetics of central African baboons*. Phd theis. Department of Anthropology. New York University, New York, pp 323.

Burrell, A.S., Jolly, C.J., Tosi, A.J. and Disotell, T.R. (2009). Mitochondrial evidence for the hybrid origin of the kipunji, *Rungwecebus kipunji* (Primates: Papionini). *Molecular Phylogenetics and Evolution* 51:340-348.

Byrne, R.W., Whiten, A., Henzi, S.P., McCulloch, F.M. (1993). Nutritional constraints on mountain baboons (*Papio ursinus*)- implications for baboons socioecology. *Behavioral Ecology and Sociobiology* 33(4):233-246.

C

Cain, M.L., Milligan, B.G., Strand, A.E. (2000). Long-distance seed dispersal in plant populations. *Americal Journal of Botany* 87:1217-1227.

Castoe, T.A., de Koning, A.P.J., Kim, H.M., Gu, W.J., Noonan, B.P., Naylor, G., Jiang, Z.J., Parkinson, C.L., Pollock, D.D. (2009). Evidence for an ancient adaptive episode of convergent molecular evolution. *Proceedings of the National Academy of Sciences, USA* 106(22): 8986-8991.

Castoe, T.A., Jiang, Z.J., Gu, W.J., Wang, .Z.O., Pollock, D.D. (2008). Adaptive evolution and functional redesign of core metabolic proteins in snakes. *PLoS ONE* 3(5).

University of Cape Town

Cerling, T.E. (1992). Development of grasslands and savannas in East Africa during the Neogene. *Palaeogeography, Palaeoclimatoology, Palaeoecology* 97:241-247.

Chase, B.M. and Meadows, M.E. (2007). Late Quaternary dynamics of southern Africa's winter rainfall zone. *Earth-Science Reviews* 84(3-4):103-138.

Chen, W.J., Bonillo, C., Lecointre, G. (2003). Repeatability of clades as a criterion of reliability: A case study for molecular phylogeny of *Acanthomorpha (Teleostei)* with larger number of taxa. *Molecular Phylogenetics and Evolution* 26 262–288.

Cheney. D.L. and Seyfarth, R.M. (1995). The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: evidence for causal reasoning? . *Journal of Comparative Psychology* 109 (2):134-141.

Cheney, D.L., Seyfarth, R.M., Palombit, R. (1996). The function and mechanisms underlying baboon 'contact' barks. *Animal Behaviour* 52:507-518.

Cheney, D.L., Seyfarth, R.M. (1997). Reconciliatory grunts by dominant female baboons influence victims' behaviour. *Animal Behaviour* 54:409-418.

Cheney, D.L., Seyfarth, R.M. (1999). Recognition of other individuals' social relationships by female baboons. *Animal Behaviour* 58: 67-75.

Cheney, D.L., Seyfarth, R.M., Fischer, J., Beehner, J., Bergman, T., Johnson, S., et al. (2004). Factors affecting reproduction and mortality among baboons in the Okavango Delta, Botswana. *International Journal of Primatology* 25(2):410-428.

Chung, W.K., Steiper, M.E. (2008). Mitochondrial COII introgression into the nuclear genome of *Gorilla gorilla*. *International Journal of Primatology* 29:1341-1353.

Ciborowski, K.L., Consuegra, S., de Leaniz, C.G., Beaumont, M.A., Wang, J.L., Jordan, W.C. (2007). Rare and fleeting: an example of interspecific recombination in animal mitochondrial DNA. *Biological Letters* 3(5):554-557.

Clark, J.S, Fastie, C., Hurtt, G., Jackson, S.T., Johnson, C. et al. (1998). Reid's paradox of rapid plant migration: dispersal theory and interpretation of paleoecological records. *Bioscience* 48:13-24.

Clarke, A., Crame, J. A. (1997). *Marine Biodiversity: Patterns and Processes*, eds. Ormond, R. F. G., Gage, J. D. and Angel, M. V. Cambridge University Press, pp 122–147.

Codron, D.M. (2003). *Dietary Ecology of Chacma Baboons (*Papio ursinus *(Kerr, 1792)) and Pleistocene Cercopithecoidea in Savanna Environments of South Africa*. MSc thesis. Quaternary Science. University of Cape Town, Cape Town, p 190.

Collura, R.V., Stewart, C.B. (1995). Insertions and duplications of mtDNA in the nuclear genomes of old world monkeys and hominoids. *Nature* 378(6556):485-489.

Collura, R.V., Auerbach, M.R., Stewart, C.B. (1996). A quick, direct method that can differentiate expressed mitochondrial genes from their nuclear pseudogenes. *Current Biology* 6 (10):1337-1339.

University of Cape Town

Cornwallis, C.K. and Birkhead, T.R. (2008). Plasticity in reproductive phenotypes reveals status-specific correlations between behavioral, morphological, and physiological sexual traits. *Evolution* 62: 1149– 1161.

Cowling, R.M., Holmes, P.M., Rebelo, A.G. (1992). *Plant diversity and endemism*. In: Cowling RM (ed) The Ecology of Fynbos: Nutrients, Fire and Diversity. Oxford University Press. pp 62-112.

Cowlishaw, G. (1997). Refuge use and predation risk in a desert baboon population. *Animal Behaviour* 54:241-253.

Cowlishaw, G., Davies, J.G. (1997). Flora of the Pro-Namib Desert Swakop River Catchment, Namibia: community classification and implications for desert vegetation sampling. *Journal of Arid Environments* 36 (2):271-290.

Cowlishaw, G., Hacker, J.E. (1997). Distribution, diversity and latitude in African primates *American Naturalist* 150 (4):505-512.

Cronin, J.E., Sarich, V.M. (1976). Molecular evidence for dual origin of mangabeys among Old World monkeys. *Nature* 260:700-702.

Cummings, M.P., Neel, M.C., Shaw, K.L. (2008). A genealogical approach to quantifying lineage divergence. *Evolution* 62*:* 2411-2422

D

Davenport, T.R.B., Stanley, W.T., Sargis, E.J., De Luca, D.W., Mpunga, N.E., Machanga, S.J., Olson, L.E. (2006). A new genus of African monkey, *Rungwecebus*: morphology, ecology, and molecular phylogenetics. *Science* 312:1378-1381.

Davenport, T.R.B., De Luca, D.W., Jones, T., Mpunga, N.E., Machaga, S.J., Picton Phillips, G. (2008). The kipunji (*Rungwecebus kipunji)* of southern Tanzania: first census and assessments of distribution and conservation status. *Oryx* 42:352-359.

Davidge, C. (1978). Ecology of baboons (*Papio ursinus*) at Cape Point. *Zoologica Africana* 13: 329-350.

Davies, J.G., Cowlishaw, G. (1996). Baboon carnivory and raptor interspecific competition in the Namib Desert. *Journal of Arid Environments* 34(2):247-249.

Davis, R.E., Paker, W.D. (1998). Evidence that two reports of mtDNA cytochrome c oxidase "mutations" in Alzheimer's disease are based on nDNA pseudogenes of recent evolutionary origin. *Biochemical and Biophysical Research Communications* 244:877-883.

Davis, J.M. (2004). The effect of natal experience on habitat preferences. *Trends in Ecology and Evolution* 19:411-416.

Deacon, H.J., Lancaster. N. (1988). *Late Quaternary palaeoenvironments of southern Africa*. Claredon Press, Oxford.

Delson, E. (1975). *Palaeogeography and zoogeography of the old world monkeys*. In: Tuttle, R.H (ed) Primate functional morphology and evolution. Mouton Publishers, Paris, pp. 37-64.

University of Cape Town

Delson, E. (1984). Cercopithecid biochronology of the African Plio-Pleistocene: correlation among eastern and southern hominid-bearing localities. *Courier Forschungsinstitut Senckenberg* 69:199-218.

Delson, E. (1988). *Chronology of South African australopith site units*. In: Grine FE (ed) Evolutionary History of the "Robust" *Australopithecines*, Aldine de Gruyter, New York.

Delson, E., Dean, D. (1993). *Are* Papio baringensis *(R. Leakey, 1969) and* P. antiquus *(Iwamoto, 1982), species of* Papio *or* Theropithecus*?* In: Jablonski NG (ed) *Theropithecus*: the Rise and Fall of a Primate Genus Cambridge University Press., Cambridge.

Delson, E. (2000). *Cercopithecoidea.* In: Delson, E., Tattersall, I., Van Couvering, J.A., Brooks, A. (eds) Encyclopaedia and Human Evolution and Prehistory, 2 edn. Garland, New York.

de Menocal, P.B., Bloemendal, J. (1995). *Plio-Pleistocene climatic variability in subtropical Africa and the paleoenvironment of hominid evolution: a combined data-model approach.* In: Vrba ES, Denton GH, Partridge TC, Burckle LH (eds) Paleoclimate and evolution, with emphasis on human origins. Yale University Press. pp 262-288.

deMenocal, P.B. (1995). Plio-Pleistocene African Climate. *Science* 270(5233):53-59.

de Menocal, P.B. (2004). African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters* 220: 3-24.

Disotell, T.R. (1994). Generic level relationships of the *Papionin*i (Cercopithecoidea). *American Journal of Physical Anthropology* 94: 47-57.

Deuter, R., Pietsch, S., Hertel, S., Muller, O. (1995). A method for preparation of fecal DNA suitable for PCR. *Nucleic Acids Research* 23:3800-3801.

Devore, I., Hall, K.R.L. (1965). *Baboon ecology*. In: Devore I (ed) Primate Behavior: Field Studies of Monkeys and Apes Holt, Rinehart and Winston, Inc. USA, pp 20-52

DeWoody, J.A., Chesser, R.K., Baker, R.J. (1999). A translocated mitochondrial cytochrome b pseudogene in voles (*Rodentia: Microtus). Molecular Evolution* 48:380-382.

Disotell, T.R., Honeycutt, R.L., Ruvolo, M. (1992). Mitochondrial DNA phylogeny of the Old-World monkey tribe Papionini. *Molecular Biology and Evolution* 9(1):1-13.

Dollar, E.S.J. (1998) Palaeofluvial geomorphology in southern Africa: a review. *Progress in Physical Geography* 22: 325-349.

Doosey, M.H., Bart Jr, H.L., Saitoh, K., Miya, M. (2010). Phylogenetic relationships of catostomid fishes (*Actinopterygii: Cypriniformes*) based on mitochondrial ND4/ND5 gene sequences. *Molecular Phylogenetics and Evolution* 54(3):1028-1034.

Douady, C.J., Catzeflis, F., Raman, J., Springer, M.S., Stanhope, M.J. (2003) The Sahara as a vicariant agent, and the role of Miocene climatic events, in the diversification of the mammalian order *Macroscelidea* (elephant shrews). *Proceedings of the National Academy of Science, USA* 100(14):8325-8330.

Drummond, A.J., Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7:214.

University of Cape Town

Du Buy, H.G., Riley, F.L. (1967). Hybridization between the nuclear and kinetoplast DNAs of Leishmania enriettii and between nuclear and mitochondrial DNAs of mouse liver. *Proceedings of the National Academy of Sciences, USA* 57:790-797.

Ducroz, J.F., Volobouev, V., Granjon, L. (1998). A molecular perspective on the systematics and evolution of the genus *Arvicanthis* (*Rodentia, Muridae*): Inferences from complete *cytochrome b* gene sequences. *Molecular Phylogenetics and Evolution* 10(1):104-117.

Dunbar, R.I.M. (1992). Time - a hidden constraint on the behavioural ecology of baboons. *Behavioral Ecology and Sociobiology* 31(1):35-49.

Dupanloup, I., Schneider, S., Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* 11:2571-2581.

Dynesius, M., Jansson, R. (2000). Evolutionary consequences of changes in species geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences, USA* 97(16):9115-9120.

E

Eck, R., Dayhoff, M. (1966). *Atlas of protein sequence and structure*. National Biomedical Research Foundation, Silver Springs, Maryland.

Eck, G.G., Jablonski, N.G. (1987). *The skull of* Theropithecus brumpti *compared with those of other species of the genus* Theropithecus*. I*n: Coppens, Y., Howell, F.C. (Eds.), Les Faunes Plio-Pleistocenes de la Basse Vallee de l'Omo (Ethiopie), Tome 3, Cercopithecidae de la Formation Shungura. Cahiers de Paleontologie, Editions du Centre National de la Recherche Scientifique, pp. 11-122.

Edwards, S.V., Beerli, P. (2000). Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54:1839–1854.

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: Biological Sciences* 269(1504):1993-2006.

El-Zaatari, S., Grine, F.E., Teaford, M.F., Smith, H.F.(2005). Molar microwear and dietary reconstructions of fossil Cercopithecoidea from the Plio-Pleistocene deposits of South Africa. *Journal of Human Evolution* 49:180-205.

Engh, A.L., Beehner, J.C., Bergman, T.J., Whitten, P.L., Hoffmeier, R.R., Seyfarth, R.M.,et al. (2006). Female hierarchy instability, male immigration and infanticide increase glucocorticoid levels in female chacma baboons. *Animal Behaviour* 71:1227-1237.

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(2):479-491.

Excoffier, L., Yang, Z. (1999). Substitution rate variation among sites in mitochondrial hypervariable region I of humans and chimpanzees. *Molecular Biology and Evolution* 16(10):1357-1368.

Excoffier, L., Laval, G., Schneider, S. (2005). Arlequin version 3: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1.

University of Cape Town

Eyre-Walker, A., Smith, N.H., Smith, J.M. (1999). How clonal are human mitochondria? *Proceedings of the Royal Society of London B* 266(1418):477-483.

F

Felsenstein, J. (1981). Evolutionary trees from DNA sequences- a maximum likelihood approach. *Journal of Molecular Evolution* 17(6):368-376.

Felsenstein, J. (1988). Phylogenies from molecular sequences - inference and reliability. *Annual Review of Genetics* 22:521-565.

Fischer, J., Cheney, D.L., Seyfarth, R.M. (2000). Development of infant baboons' responses to graded bark variants. *Proceedings Royal Society London B* 267:2317-2321.

Fischer, J., Hammerschmidt, K., Cheney, D.L., Seyfarth, R.M. (2002). Acoustic features of male baboon loud calls: influences of context, age and individuality. *Journal of the Acoustical Society of America* 111(3):1464-1472.

Fischer, J., Kitchen, D., Seyfarth, R.M., Cheney, D.L. (2004). Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. B*ehavioral Ecology and Sociobiology* 56:140-148.

Fitch, W. (1971).Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology* 20:406-416.

Flagstad, Ø., Syvertsen, P.O., Stenseth, N.C., Jakobsen, K.S. (2001). Environmental change and rates of evolution: the phylogeographic pattern within the hartebeest complex as related to climatic variation. *Philosophical Transcripts Royal Society, London B* 268:667–677.

Fleagle, J.G., McGraw, W.S. (1999). Skeletal and dental morphology supports diphyletic origin of baboons and mandrills. *Proceedings of the National Academy of Sciences of the USA*, 96:1157-1161.

Fleagle, J.G., McGraw, W.S. (2002). Skeletal and dental morphology of African papionins: unmasking a cryptic clade. *Journal of Human Evolution*, 42(3):267-292.

Freedman, L. (1957). The fossil Cercopithecoidea of South Africa. *Annals of the Transvaal Museum*, 23,121-262.

Freedman, L. (1976). South African fossil Cercopithecoidea: a reassessment including a description of new material from Makapansgat, Sterkfontein, and Taung. *Journal of Human Evolution*, 5:297-315.

Frost, S.R. (2001). New early Pliocene Cercopithecidae (Mammalia: Primates) from Aramis, Middle Awash valley, Ethiopia. *American Museum Novitates*. 3350: 1-36.

Frost, S.R., Delson, E. (2002). Fossil Cercopithecidae From the Hadar Formation and Surrounding Areas of the Afar Depression, Ethiopia. *Journal of Human Evolution* 43: 687-748.

Frost, S.R., Marcus, L.F., Bookstein, F.L., Reedy, D.P., Delson, E. (2003). Cranial allometry, phylogeography, and systematics of large-bodied papionins (Cercopithecinae) inferred from geometric morphometric analysis of landmark data. *The Anatomical Record*. 275A (2) 1048-1072.

University of Cape Town

Fu, Y., Li, W. (1993). Statistical tests of neutrality of mutations. *Genetics* 133:693 - 709.

Fu, Y. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915 - 925.

Fu, Y-X , Li, W. H. (1999) Coalescing into the 21st Century: An Overview and Prospects of Coalescent Theory *Theoretical Population Biology* 56 (1):1-10.

G

Galley, C., Linder, H.P. (2006). Geographical affinities of the Cape flora, South Africa. *Journal of Biogeography* 33:236-250.

Galtier, N., Nabholz, B., Giamin, S., Hurst, G.D.D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology* 18 (22): 4541-4550.

Gantenbein, B., Fet, V., Gantenbein-Ritter, I.A., Balloux, F. (2005). Evidence for recombination in scorpion mitochondrial DNA (*Scorpiones : Buthidae*). *Proceedings of the Royal Society B* 272(1564):697-704.

Giess, W. (1971). A preliminary vegetation map of South West Africa. *Dinteria* 4:5-14.

Gilbert, C.C. (2007). Craniomandibular morphology supporting the diphyletic origin of mangabeys and a new genus of the *Cercocebus* / *Mandrillus* clade, *Procercocebus*. *Journal of Human Evolution* 53:69-102.

Gilbert, C.C. (2008). *African Papionin Phylogenetic History and Plio-Pleistocene Biogeography*. PhD thesis. Dept of Anthropology. Stony Brook University, New York.

Gissi, C., Reyes, A., Pesole, G., Saccone, C. (2000). Lineage-specific evolutionary rate in mammalian mtDNA. *Molecular Biology and Evolution* 17(7):1022-1031.

Goldblatt, P., Manning, J.C. (2002). Plant diversity of the Cape Region of southern Africa. *Annals of the Missouri Botanical Garden* 89:281-302.

Goudie, A.S., Allchin, B., Hegde, K.T.M. (1973). The former extensions of the Indian sand desert. *The Geographical Journal* 139: 243-257.

Graybeal, A. (1998). Is it better to add taxa or characters to a difficult phylogenetic problem? *Systematic Biology* 47:9-17.

Grossman, L.I., Wildman, D.E., Schmidt, T.R., Goodman, M. (2004). Accelerated evolution of the electron transport chain in anthropoid primates. *Trends in Genetics* 20(11):578-585.

Grove, A.T. (1958). The ancient erg of Hausaland and other similar formations on the southern side of the Sahara. *The Geographical Journal* 124 (4):526-533.

Groves, C.P. (1978). Phylogenetic and population systematics of the mangabeys (Primates: Cercopithecoidea). *Primates* 19:1-34.

Groves, C. (2001). *Primate Taxonomy*. Smithsonian Institution Press, Washington.

University of Cape Town

H

Haacke, W.D. (1989). Zoogeography of the Namib Desert reptiles as affected by dune formations. *Dunes 1989 Meeting*, Swakopmund, p 48.

Haffer J. (1969). Speciation in Amazonian forest birds. *Science* 165:131–37.

Haffer J. (1997). Alternative models of vertebrate speciation in Amazonia: an overview. *Biodiversity and Conservation* 6:451–76

Hall, B. (2004). *Phylogenetic trees made easy: a how-to manual.* 2nd edn. Sinauer Sunderland, Massachusetts.

Hamilton, W.J., Buskirk, R.E., Buskirk, W.H. (1975). Chacma baboon tactics during intertroop encounters. *Journal of Mammalogy* 56(4):857-870.

Hamilton III, W.J. (1985). Demographic consequences of a food and water desert chacma baboons, *Papio ursinus*. *International Journal of Primatology* 6 (5):451-462.

Hamilton, A.C., Taylor, D. (1991). History of climate and forests in tropical Africa during the last 8 million years. *Climatic Change* 19:65-78.

Harpending, H.C., Sherry, S.T., Rogers, A.R., Stoneking, M. (1993). The genetic-structure of ancient human-populations. *Current Anthropology* 34: 483-496

Harpending, H. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* 66(4):591 - 600.

Harris, E.E., Disotell, T.R. (1998). Nuclear gene trees and the phylogenetic relationships of the mangabeys (Primates: Papionini). *Molecular Biology and Evolution* 15:892-900.

Harrison, R. (1991). Molecular changes at speciation. *Annual Review of Ecology and Systematics* 22: 281– 308.

Hartigan, J.A. (1973). Minimum mutation fits to a given tree. *Biometrics* 29(1):53-65.

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

Hasegawa, M., Thorne, J.L., Kishino, H. (2003). Time scale of eutherian evolution estimated without assuming a constant rate of molecular evolution. *Genes and Genetic Systems* 78(4):267-283.

Heaney, L.R., Walsh, J.S., Peterson, A.T. (2005). The roles of geological history and colonization abilities in genetic differentiation between mammalian populations in the Philippine archipelago. *Journal of Biogeography* 32(2):229-247.

Heaton, T.H.E., Talma, A.S., Vogel, J.C. (1983). Origin and history of nitrate in confined groundwater in the western Kalahari. *Journal of Hydrology* 62:243-262.

Heaton, J. (2007). Taxonomy of the Sterkfontein fossil cercopithecinae: the papionin of members 2 and 4 (Gauteng, South Africa). Phd thesis. Dept of Anthropology. Indiana University.

University of Cape Town

Henzi, P., Lycett, J. E., Piper, S. E. (1997). Fission and troop size in a mountain baboon population. *Animal Behaviour* 53: 525–535

Henzi, S.P., Weingrill, T. (1999). Male Behavior and the evolutionary ecology of chacma baboons. *South African Journal of Science* 95(5):240.

Henzi, P., Barrett, L. (2002) Infants as a Commodity in a Baboon Market. *Animal Behaviour* 63:915-921.

Henzi, S.P., Barrett, L., Gaynor, D., Greef, J., Weingrill, T., Hill, R.A. (2003). Effect of resource competition on the long-term allocation of grooming by female baboons: evaluating Seyfarth's Model. *Animal Behaviour* 66:931-938.

Henzi, P., Barrett, L. (2003). Evolutionary ecology, sexual conflict, and behavioural differentiation among baboon populations. *Evolutionary Anthropology* 12(5):217-230.

Hernández F. M., Vrba, E.S. (2006). Plio-Pleistocene climatic change in the Turkana Basin (East Africa): Evidence from large mammal faunas. *Journal of Human Evolution* 50(6):595-626.

Herron, M.D., Waterman, J.M., Parkinson, C.L. (2005). Phylogeny and historical biogeography of African ground squirrels: the role of climate change in the evolution of *Xerus*. *Molecular Ecology* (1) 2773-2788.

Hewett-Emmett, D., Cook, C.N. (1978). *Atypical evolution of papionine alphahaemoglobins and indication that* Cercocebus *may not be a monophyletic genus*. In: Chivers DJ, Joysey KA (eds) Recent Advances in Primatology 3: Evolution Academic Press.

Hewitt, G.M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* 58: 247-276.

Hewitt, G.M. (2000). The genetic legacy of the Quaternary ice ages. *Nature* 405:907-913.

Hewitt, G.M. (2001). Speciation, hybrid zones and phylogeography- or seeing genes in space and time. *Molecular Ecology* 10:537-549.

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

Hey, J., Machado, C.A. (2003). The study of structured populations—new hope for a difficult and divided science. *Nature Reviews Genetics* 4: 535–543.

Hey, J., Nielsen, R. (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167: 747–760.

Hey, J., Nielsen, R. (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences, USA* 104: 2785–2790.

Hey, J. (2007) Introduction to the IM and IMa computer programs - March 5, 2007 Department of Genetics, Rutgers University.

University of Cape Town

Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H., Johnson, J.B., Rissler, L., Victoriano, P.F., Yoder, A.D. (2010). Phylogeography's past, present, and future: 10 years after. *Molecular Phylogenetics and Evolution* 54(1):291-301.

Hill, W.C.O. (1970). Primates comparative anatomy and taxonomy VIII: Cynopithecinae; *Papio, Mandrillus, Theropithecus*. Edinburg University Press.

Hill, R.A., Dunbar, R.I.M. (2002). Climatic determinants of diet and foraging behaviour in baboons. *Evolutionary Ecology* 16(6):579-593.

Hill, R.A., Barrett, L., Gaynor, D., Weingrill, T., Dixon, P., Payne, H., Henzi, S.P. (2003). Day length, latitude and behavioural (in)flexibility in baboons (*Papio cynocephalus ursinus*). *Behavioral Ecology and Sociobiology* 53(5):278-286.

Hillis, D.M. (1996). Inferring complex phylogenies. *Nature* 383:130-131.

Hillis, D.M. (1998). Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Systematic Biology* 47:3-8.

Hiraoka, J., Hirao,Y. (1988). Fate of Sperm Tail Components After Incorporation Into the Hamster Egg. *Gamete Research* 19(4):369-380.

Ho, S.Y.W., Phillips, M.J., Cooper, A., Drummond, A.J. (2005). Time Dependency of Molecular Rate Estimates and Systematic Overestimation of Recent Divergence Times. *Molecular Biology and Evolution* 22(7):1561-1568.

Ho, S.Y.W., Saarma, U., Barnett, R., Haile, J., Shapiro, B. (2008). The effect of inappropriate calibration : three case studies in molecular ecology. *PloS ONE* 3(2): e1615.doi: 1610.1371/journal.pone.0001615

Hoarau, G., Holla, S., Lescasse, R., Stam, W.T., Olsen, J.L. (2002). Heteroplasmy and evidence for recombination in the mitochondrial control region of the flatfish Platichthys flesus. *Molecular Biology and Evolution* 19(12):2261-2264.

Holland, B., Moulton, V. (2003). *Consensus networks: A method for visualizing incompatibilities in collections of trees*. In: G. B, Page R (eds) Proceedings of "Workshop on Algorithms in Bioinformatics" Springer, pp 165-176

Hooghiemstra, H., Melice, J.L., Berger, A., Shackleton, N.J. (1993). Frequency spectra and paleoclimatic variability of the high-resolution 30-145 kyr Funza I pollen record (Eastern Cordillera, Colombia). *Quaternary Science Review* 12:41-156.

Howard, W.R. (1997). A warm future in the past. *Nature* 388:418-419.

Huelsenbeck, J.P., Crandall, K.A. (1997). Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review Ecology and Systematics* 28:437-466.

Huelsenbeck, J.P., Ronquist, F. (2001). MR BAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 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.

University of Cape Town

I - J

Ibrahim, K.M., Nichols, R.A., Hewitt, G.M. (1996). Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* 77:282-291.

Immelmann, K. (1975) Ecological significance of imprinting and early learning. *Annual Review in Ecology and Systematics.* 6:15–37.

Irwin, D.E. (2002). Phylogeographic breaks without geographic barriers. *Evolution* 56: 2383–2394.

Jablonski, N.G. (1993). *Theropithecus*: *the rise and fall of a primate genus*. In: Jablonski, N. (ed) *Theropithecus*: the rise and fall of a primate genus. Cambridge University Press, Cambridge

Jablonski, N.G. (2002). Fossil Old World monkeys: the late Neogene radiation. In: Hartwig, W.C. (ed) The Primate Fossil Record. Cambridge University Press New York pp 255-299.

Jacobs, B.F. (2004). Paleobotanical studies from tropical Africa: relevance to the evolution of forest, woodland and savannah biomes. *Philosophical Transcripts of the Royal Society of London B* 359:1573-1583.

Jansson, R., Dynesius, M. (2002). The fate of clades in a world of recurrent climate change: Milankovitch Oscillations and Evolution. *Annual Review of Ecology and Systematics* 33: 741-777.

Johnson, K.P. (2001). Taxon sampling and the phylogenetic position of passeriformes: Evidence from 916 avian cytochrome b sequences. *Systematic Biology* 50:128-136.

Jolly, C.J. (1965). *The origins and specializations of the long-faced Cercopithecoidea*. University of London.

Jolly, C.J. (1972). The classification and natural history of *Theropithecus* (*Simopithecus*) (Andrews, 1916), baboons of the African Plio-Pleistocene. *Bulletins of the British Museum (Nat Hist) Geology* 22 1-122.

Jolly, C. (1993). *Species, subspecies, and baboon systematics*. In: W. H. Kimbel LBM (ed) Species, species concepts, and primate evolution. Plenum Press, New York, pp 67–107.

Jolly, C. (2001). A proper study for mankind: analogies from the Papionin monkeys and their implications for human evolution. *Yearbook of Physical Anthropology* 44:177-204.

Jones, T., Ehardt, C.L., Butynski, T.M., Davenport, T.R.B, Mpunga, N.E., Machaga, S.J. and De Luca, D.W. (2005). The highland mangabey *Lophocebus kipunji*: a new species of African monkey. *Science* 308 (5725):1161-1164.

Jukes, T.H., Osawa, S. (1993). Evolutionary changes in the genetic-code. C*omparative Biochemistry and Physiology B-Biochemistry & Molecular Biology* 106(3):489-494.

Jürgens, N. (1997). Floristic Biodiversity and History of African Arid Regions. *Biodiversity and Conservation* 6:495–514.

University of Cape Town

K

Kaneda, H., Hayashi, J-I., Takahama, S., Taya, C., Fischer Lindahl K., Yonekawa. H. (1995). Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis. *Proceedings of the National Academy of Sciences, USA* 92(10):4542-4546.

Karanth, K.P., Singh, L., Collura, R.V., Stewart, C-B. (2008). Molecular phylogeny and biogeography of langurs and leaf monkeys of South Asia (Primates: Colobinae). *Molecular Phylogenetics and Evolution* 46(2):683-694.

Keller, C., Roos, C., Groeneveld, L.F., Fischer, J., Zinner, D. (2010). Introgressive hybridization in southern African baboons shapes patterns of mtDNA variation. *American Journal of Physical Anthropology* 142:125-136.

Kerr, R. (1792). *The animal kingdom or zoological system of the celabrated Sir Charles Linnaeus.* J Murray.

Kimura, M. (1968). Evolutionary rate at molecular level. *Nature* 217:624-626.

Kimura, M. and Takahata, N. (1983). Selective constraint in protein polymorphism- study of the effectively neutral mutation model by using an improved pseudosampling method. *Proceedings of the National Academy of Sciences, USA* 80(4):1048-1052.

King, J.L. and Jukes, T.H. (1969). Non-Darwinian evolution. *Science* 164(3881):788.

Kingdon, J. (1997). *The Kingdon field guide to African mammals*. San Diego Academic Press. pp 464.

Kingman, J.F.C. (1982) On the Genealogy of Large Populations. *Journal of Applied Probability* 19: 27-43.

Kingman, J.F.C. (2000) Origins of the coalescent 1974-1982. *Genetics* 156: 1461-1463.

Kitchen, D.M., Cheney, D.L., Seyfarth, R.M. (2004). Factors mediating inter-group encounters in savannah baboons (*Papio cynocephalus ursinus*). *Behaviour* 141:197-218.

Kitchen, D.M., Cheney, D.L., Seyfarth, R.M. (2005a). Contextual factors meditating contests between male chacma baboons in Botswana: effects of food, friends and females. *International Journal of Primatology* 26(1):105-125.

Kitchen, D.M., Cheney, D.L., Seyfarth, R.M. (2005b). Male chacma baboons (*Papio hamadryas ursinus).* Discriminate loud call contests between rivals of different relative ranks. *Animal Cognition* 8:1-6.

Klein, R.G. (1999). *The Human Career: Human Biological and Cultural Origins, 2nd edn.* University of Chicago Press, Chicago.

Knight, M., Joyce, P. (1997). *The Kalahari: survival in a thirstland wilderness*. Struik Publishers, Cape Town.

Knoch, K., Schulze, A. (1956). *Niederschlag, Temperatur und Schwüle in Africa*. Falk Verlag, Hamburg.

Knowles, L.L., Maddison, W.P. (2002). Statistical phylogeography. *Molecular Ecology* 11: 2623-2635.

University of Cape Town

Knowles, L.L. (2004). Mini review: The burgeoning field of statistical phylogeography. *Journal of Evolutionary Biology*.17: 1-10.

Knowles, L.L. (2009). Statistical Phylogeography. *Annual Review of Ecology Evolution and Systematics* 40(1):593-612.

Kuhner, M.K., Yamato, J., Felsenstein, J. (1998). Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* 149: 429–434.

Kuhner, M.K. (2006). LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22: 768–770.

Kulongoski, J.T., Hilton, D.R. (2004). Climate variability in the Botswana Kalahari from the late Pleistocene to the present day. *Geophysical Research Letters* 31:L10204

L

Ladoukakis, E.D., Zouros, E. (2001). Direct evidence for homologous recombination in mussel (*Mytilus galloprovincialis*) mitochondrial DNA. *Molecular Biology and Evolution* 18(7):1168-1175.

Lamb, T., Bauer, A.M. (2003). *Meroles* revisited: complementary systematic inference from additional mitochondrial genes and complete taxon sampling of southern Africa's desert lizards. *Molecular Phylogenetics and Evolution* 29(2):360-364.

Lancaster, N. (1981). Palaeoenvironmental implications of fixed dune systems in southern Africa. *Palaeogeography, Palaeoclimatology, Palaeoecology* 33: 327-346.

Lawes, M.J. (1990). The distribution of the samango monkey (*Cercopithecus mitis erythrarchus* (Peters, 1852 and *Cercopithecus mitis labiatus* (I. Geoffroy, 1843) and forest history in southern Africa. *Journal of Biogeography* 17(6):669-680.

Leakey, M.G., Teaford, M.F., Ward, C.V. (2003). *Cercopithecidae from Lothagam. Lothagam: the dawn of humanity in eastern Africa*. Columbia University Press., New York.

Lee-Thorp, J.A., van der Merwe, N.J. (1993). Stable carbon isotope studies of Swartkrans fossils.*Transvaal Museum Monograph* 8:251-256.

Lee-Thorp, J.A., Sponheimer, M., Luyt, J. (2007). Tracking changing environments using stable carbon isotopes in fossil tooth enamel: an example from the South African hominin sites. *Journal of Human Evolution* 53:598-601.

Lesson, R.P. (1827). *Manuel de mammologie*, Paris.

Lewis, C.A. (2008). Late Quaternary climatic changes, and associated human responses, during the last ~ 45 000 yr in the Eastern and adjoining Western Cape, South Africa. *Earth-Science Reviews* 88(3-4):167-187

Li, W-H., Wu, C-I., Luo, C-C. (1985). *Evolution of DNA sequences in molecular evolutionary genetics*, Plenum Press.

Librado, P., Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.

University of Cape Town

Linder, H.P., Meadows, M.E., Cowling, R.M. (1992). History of the Cape Flora. In: Cowling RM (ed) *The Ecology of Fynbos: Nutrients, Fire and Diversity*. Oxford University Press, pp 113-134.

Linder, H., Hardy, C.R., Rutschmann, F. (2005). Taxon sampling effects in molecular clock dating: An example from the African Restionaceae. *Molecular Phylogenetics and Evolution* 35:569-582.

Linnaeus, C. (1766). *Systema Naturae* 12 edn. Holmiae: L. Salvii.

Lopez, J.V., Yuhki, N., Masuda, R., Modi, W., O'Brien, S. (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.

Lorenzen, E.D., Masembe, C., Arctander, P., Siegismund, H.R. (2009). A long-standing Pleistocene refugium in southern Africa and a mosaic of refugia in East Africa: insights from mtDNA and the common eland antelope. *Journal of Biogeography* 37(3):571-581.

M

Mabbutt, J.A. (1971). *Desert landforms*. Cambridge Massachusettes: MIT Press.

Maddison, W. P., Slatkin, M. (1991). Null models for the number of evolutionary steps in a character on a phylogenetic tree. *Evolution* 45:1184–1197.

Maier, W. (1971). Two new skulls of *Parapapio antiquus* from Taung and a suggested phylogenetic arrangement of the genus Parapapio. *Annals of the South African Museum* 59:1-16.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209-220.

Marshall, D.C., Hill, K.B.R., Fontaine, K.M., Buckley, T.R., Simon, C. (2009). Glacial refugia in a maritime temperate climate: Cicada (*Kikihia subalpina*) mtDNA phylogeography in New Zealand. *Molecular Ecology* 18(9):1995-2009.

Martin, A.P., Naylor, G.J.P., Palumbi, S.R. (1992). Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. *Nature* 357(6374):153-155.

Matthee, C.A., Robinson, T.J. (1996). Mitochondrial DNA differentiation among geographical populations of *Pronolagus rupestris*, Smith's red rock rabbit (Mammalia: *Lagomorpha*). *Heredity* 76(5):514-523.

Matthee, C.A., Robinson, T.J. (1997). Mitochondrial DNA phylogeography and comparative cytogenetics of the springhare, *Pedetes capensis* (Mammalia: Rodentia). *Journal of Mammalian Evolution* 4(1):53-73.

Matthee, C.A., Flemming, A.F. (2002). Population fragmentation in the southern rock agama, *Agama atra*: more evidence for vicariance in Southern Africa. *Molecular Ecology* 11:465-471.

Mau, B., Newton, M.A., Larget, B. (1999). Bayesian phylogenetic inference via Markov chain Monte Carlo methods. *Biometrics* 55(1):1-12.

University of Cape Town

McGrew, W.C., McKee, J.K., Tutin, C.E.G. (2003). Primates in caves: two new reports of *Papio* spp. *Journal of Human Evolution* 44(4):521-526.

McKee, J.K., Keyser, A.W. (1994). Craniodental remains of *Papio angusticeps* from the Haasgat cave site, South Africa. *International Journal of Primatology* 15:823-884.

Meadows, M.E. (1995). Biogeography. In: Adams WM, Goudie AS, R OA (eds) The physical geography of Africa. Oxford University Press, pp 161-172.

Meadows, M.E., Baxter, A.J. (1999). Late Quaternary Palaeoenvironments of the southwestern Cape, South Africa: a regional synthesis. *Quaternary International* 57-58:193-206.

Meadows, M.E. (2001). The role of Quaternary environmental change in the evolution of landscapes: case studies from southern Africa. *Catena* 42 (1):39-57.

Michaels, G.S., Hauswirth, W.W., Laipis, P.J. (1982). Mitochondrial DNA Copy Number in Bovine Oocytes and Somatic Cells. *Developmental Biology* 94:246-251.

Midgley, G.F., Hannah, L., Millar, D., Thuiller, W., Booth, A. (2003). Developing regional and species-level assessments of climate change impacts on biodiversity in the Cape Floristic Region. *Biological Conservation* 112(1-2):87-97

Miller, M.P. (2005). Alleles In Space: Computer software for the joint analysis of interindividual spatial and genetic information. *Journal of Heredity* 96:722-724.

Miya, M., Saitoh, K., Wood, R., Nishida, M., Mayden, R.L. (2006). New primers for amplifying and sequencing the mitochondrial ND4/ND5 gene region of the Cypriniformes (*Actinopterygii :Ostariophysi*). Ichthyol Research 53(1):75-81.

Moodley, Y., Bruford, M.W. (2007). Molecular Biogeography: Towards an Integrated Framework for Conserving Pan-African Biodiversity. *PLoS ONE* 2,5:e454

Moses, M.J. (1961). Spermiogenesis in crayfish (*Procambarus clarkii*): description of stages *Journal of Biophysical and Biochemical Cytology*. 10: 301- 2.

Mucina, L., Rutherford, M.C. (2006). The Vegetation of South Africa, Lesotho and Swaziland. South African National Biodiversity Institute, Pretoria.

Muller, R.A., MacDonald, G.J. (1997). Glacial cycles and astronomical forcing. *Science* 277:415-418.

Muwanika, V.B., Nyakaana, S., Siegismund, H.R., Arctander, P. (2003). Phylogeography and population structure of the common warthog (*Phacochoerus africanus*) inferred from variation in mitochondrial DNA sequences and microsatellite loci. *Heredity* 91:361-372.

N

Nabholz, B., Glemin, S., Galtier, N. (2008). Strong variations of mitochondrial mutation rate across mammals - the longevity hypothesis. *Molecular Biology and Evolution* 25(1):120-130.

Nabholz, B., Glemin, S., Galtier, N. (2009). The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evolutionary Biology* 9.

University of Cape Town

Nachman, M.W., Hoekstra, H.E., D'Agostino, S.L. (2003). The genetic basis of adaptive melanism in pocket mice. *Proceedings of the National Academy of Sciences, USA* 100(9):5268-5273.

Nagel, U. (1973). A comparison of anubis baboons, hamadryas baboons, and their hybrids at a species border in Ethiopia. *Folia Primatologica* 19: 104-165.

Nei, M., Li, W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases.*Proceedings of the National Academy of Sciences, USA* 76:5269-5273.

Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei, M., Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. Oxford University Press.

Nersting, L.G., Arctander, P. (2001). Phylogeography and conservation of impala and greater kudu. *Molecular Ecology* 10:711-719.

Newman, T.K. (1997). Mitochondrial DNA analysis of the intraspecific hybridization in *Papio hamadryas anubis*, *P.h. hamadryas* and their hybrids in the Awash National Park, Ethiopia. PhD thesis. New York University.

Newman, T.K., Jolly, C., Rogers, J. (2004). Mitochondrial phylogeny and the systematics of baboons (*Papio*). *American Journal of Physical Anthropology*:124: 17-27.

Nichol, J.E. (1999). Geomorphological evidence and Pleistocene refugia in Africa. *The Geographical Journal* 165(1):79-89.

Nielsen, R., Wakeley, J. (2001). Distinguishing migration from isolation. A Markov chain Monte Carlo approach. *Genetics* 158: 885-896

Nordberg, M. (2001) *Introduction to coalecsent theory.* In Balding, D., Bishop, M., Cannings, C., (eds) Handbook of Statistical Genetics. Chapter 7, Wiley Publishers.

Norton, G.W., Rhine, R.J., Wynn, G.W. and Wynn, R.D. (1987). Baboon diet: a five-year study of stability and variability in the plant feeding and habitat of the yellow baboons (*Papio cynocephalus*) of Mikumi National Park, Tanzania. *Folia Primatologica* 48:78-120.

Norton, J.E., Ashley, M.V. (2004). Genetic variability and population structure among wild Baird's tapirs. *Animal Conservation* 7:211-220.

Nyakaana, S., Arctander, P., Siegismund, H.R. (2002). Population structure of the African savannah elephant inferred from mitochondrial control region sequences and nuclear microsatellite loci. *Heredity* 89:90-98

O - P

O'Connell, S., Cowlsihaw, G. (1994). Infanticide avoidance, sperm competition and mate choice: the function of copulation calls in female baboons. *Animal Behaviour* 48:687-694.

Olson, L.E., Yoder, A.D. (2002). Using secondary structure to identify ribosomal numts: Cautionary examples from the human genome. *Molecular Biology and Evolution* 19(1):93-100.

University of Cape Town

Olson, S.L., Hearty, P.J. (2009). A sustained +21 m sea-level highstand during MIS 11 (400 ka): direct fossil and sedimentary evidence from Bermuda. *Quaternary Science Reviews* 28:271-285.

Page, S.L., Goodman, M. (2001). Catarrhine phylogeny: noncoding DNA evidence for a diphyletic origin of the mangabeys and for a human-chimpanzee clade. *Molecular Phylogenetics and Evolution* 18:14-25.

Pa´ l, C., Miklos, I. (1999) Epigenetic inheritance, genetic assimilation and speciation. J*ournal of Theoretical Biology* 200: 19–37

Palombit, R.A., Seyfarth, R.M., Cheney, D.L. (1997). The adaptive value of 'friendships' to female baboons: experimental and observational evidence. *Animal Behaviour* 54:599-614.

Palombit, R.A., Cheney, D.L., Seyfarth, R.M. (2001). Female-female competition for male 'friends' in wild chacma baboons, *Papio cynocephalus ursinus*. *Animal Behaviour* 61:1159-1171.

Pamilo, P., Viljakainen. L., Vihavainen, A. (2007). Exceptionally high density of NUMTs in the honeybee genome. *Molecular Biology and Evolution* 24(6):1340-1346.

Panchal, M., Beaumont, B.A. (2007).The automation and evaluation of nested clade phylogeographic analysis. *Evolution* 61, 1466–1480.

Partridge, T.C., Avery, D..M, Botha, G.A., Brink, J.S., Deacon, J., Herbert, R.S., Maud, R.R., Scott, L., Talma, A.S., Vogel, J.C. (1990). Late Pleistocene and Holocene climatic change in southern Africa. *South African Journal of Science* 86: 302-306.

Paulo, O.S., Dias, C., Bruford, M.W., Jordan, W.C., Nichols, R.A. (2001). The persistence of Pliocene populations through the Pleistocene climatic cycles: evidence from the phylogeography of an Iberian lizard. *Proceedings of the Royal Society of London Series B*: 268(1476):1625-1630.

Petit, J.R., Jouzel, J., Raynaud, D., Barkov, N.I., Barnola, J.M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., Delmotte, G., Kotlykov, V.M., Legrand, M., Lipenkov, V.Y., Lorius, C., Pepin, L., Ritz, C., Saltzman, E., Stievenard, M. (1999). Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* 399: 429–436.

Petit, R.J., Grivet, D. (2002). Optimal randomization strategies when testing the existence of a phylogeographic structure. *Genetics* 161: 469–471.

Pfennig, D.W. et al. (2006) Ecological opportunity and phenotypicplasticity interact to promote character displacement and speciescoexistence. *Ecology* 87: 769–779.

Pfennig, D.W., Wund, M.A., Snell-Rood, E.C., Cruickshank, T., Schlichting, C.D., Moczek, A.P. (2010). Phenotypic plasticity's impacts on diversification and speciation. *Trends in Ecology and Evolution* 25(8):459-467.

Philippe, H. and Douzery, E. (1994). The pitfalls of molecular phylogeny based on four species, as illustrated by the cetacea/artiodactyla relationships. *Journal of Mammalian Evolution* 2:133-152.

University of Cape Town

Phillips-Conroy, J.E., Jolly, C.J., Brett, F.L. (1991). Characteristics of Hamadryas-like male baboons living in Anubis baboon troops in the Awash hybrid zone, Ethiopia. *American Journal of Physical Anthropology* 86:353-368.

Pickering, T.R., Clarke, R.J., Heaton, J.L. (2004). The context of Stw 573, an early hominid skull and skeleton from Sterkfontein Member 2: taphonomy and paleoenvironment. *Journal of Human Evolution* 46(3):277-295.

Pickering. R., Kramers. J.D. (2010). Re-appraisal of the stratigraphy and determination of new U-Pb dates for the Sterkfontein hominin site, South Africa *Journal of Human Evolution* 59(1):70-86.

Pickford, M. (1993). *Climatic-Change, Biogeography and* Theropithecus. In: Jablonski, N. (ed) *Theropithecus*: the rise and fall of a primate genus. Cambridge University Press, Cambridge, pp 227-243.

Piganeau, G., Gardner, M., Eyre-Walker, A. (2004). A broad survey of recombination in animal mitochondria. *Molecular Biology and Evolution* 21(12):2319-2325.

Piggott, M.P., Taylor, A.C. (2003). Extensive evaluation of faecal preservation and DNA extraction methods in Australian native and introduced Species. *Australian Journal of Zoology* 51:341-355

Piggott, M.P., Bellemain, E., Taberlet, P., Taylor, A.C. (2004). A multiplex pre-amplification method that significantly improves microsatellite amplification and error rates for faecal DNA in limiting conditions. *Conservation Genetics* 5:417-420.

Pleines, T., Jakob, S.S., Blattner, F.R. (2009). Application of non-coding DNA regions in intraspecific analyses. *Plant Systematics and Evolution* 282(3-4):281-294.

Pocock, R.I. (1911). *Proceedings of the zoological society of London* 93(17).

Poe, S. (1998). The effect of taxonomic sampling on accuracy of phylogeny estimation: test case of a known phylogeny. *Molecular Biology and Evolution* 15:1086-1090.

Poe, S. (2003). Evaluation of the strategy of long-branch subdivision to improve the accuracy of phylogenetic methods. *Systematic Biology* 52:423-428.

Pollock, D.D., Zwickl, D.J., McGuire, J.A., Hillis, D.M. (2002). Increased taxon sampling is advantageous for phylogenetic inference. *Systematic Biology* 51:664-671.

Pons, J., Vogler, A.P. (2005). Complex pattern of coalescence and fast evolution of a mitochondrial rRNA pseudogene in a recent radiation of tiger beetles. *Molecular Biology and Evolution* 22:991-1000.

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

Posada, D., Crandall, K. (2001). Intraspecific gene genealogies: tree grafting into networks. *Trends in ecology and evolution* 16:37- 45.

Potts, R. (1996). Evolution and climate variability. *Science* **273**: 922-923.

Potts, R. (1998). Variability selection in hominid evolution. *Evolutionary Anthropology* 7: 81-96.

University of Cape Town

Price, T.D. et al. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proeedings of the Royal Soiety of London B* 270, 1433–1440.

Prinsloo, P., Robinson, T.J. (1992). Geographic mitochondrial DNA variation in the rock hyrax, *Procavia capensis*. *Molecular Biology and Evolution* 9(3):447-456.

Prinsloo P; Robinson, TJ (1992) Geographic mitochondrial DNS variation in the rock hyrax, Procavia capensis Mol. Bid. Evol. 9(3):447-456.

Q - R

Raaum, R.L., Sterner, K.N., Noviello, C.M., Stewart, C.B., Disotell, T.R. (2005). Catarrhine primate divergence dates estimated from complete mitochondrial genomes: concordance with fossil and nuclear DNA evidence. *Journal of Human Evolution* 48:237-257.

Raijmann, L.E.L. et al. (1994) Genetic variation and outcrossing rate in relation to population size in *Gentiana pneumonanthe* L. *Conservation Biology* 8: 1014-1026.
Ramos-Onsins, S. and Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolutio*n 19:2092 - 2100.

Rannala, B., Yang, Z. (1996). Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *Journal of Molecular Evolution* 43:304-311.

Rannala, B., Huelsenbeck, J.P., Yang, Z., Nielsen, R. (1998). Taxon sampling and the accuracy of large phylogenies. *Systematic Biology* 47:702-710.

Rasmussen, D. R. (1979). Correlates of patterns of range use of a troop of yellow baboons (*Papio cynocephalus*). Sleeping sites, impregnable females, births, and male emigrations and immigrations. *Animal Behaviour* 27(4): 1098-1112.

Raynaud, D., Barnola, J.M., Souchez, R., Lorrain, R., Petit, J.R., Duval, P., Lipenkov, V.Y. (2005). The record for marine isotopic stage 11. *Nature* 436:39-40.

Rector, A.L., Reed, K.E. (2010). Middle and late Pleistocene faunas of Pinnacle Point and their paleoecological implications. *Journal of Human Evolution* 59(3-4):340-357.

Reed, K.E. (1997). Early hominid evolution and ecological change through the African Plio-Pleistocene. J*ournal of Human Evolution* 32:289-322.

Richly, E., Leister, D. (2004a). Numts in sequenced eukaryotic genomes. *Molecular Biology and Evolution* 21:1081-1084.

Richly, E., Leister, D. (2004b). NUPTs in sequenced eukaryotes and their genomic organization in relation to NUMTs their genomic organization in relation to NUMTs. *Molecular Biology and Evolution* 21:1972-1980.

Rodman, P.S. (1999). Whither primatology? The place of primates in contemporary anthropology. *Annual Review of Anthropology* 28:311-339.

Rogers, A.R., Harpending, H.C. (1992). Population growth makes waves in the distribution of pairwise generic differences. *Molecular Biology and Evolution* 9:552 - 569.

Rokas, A., Ladoukakis, E., Zouros, E. (2003). Animal mitochondrial DNA recombination revisited. *Trends in Ecology and Evolution* 18(8):411-417.

University of Cape Town

Rowell, T.E. (1966) Forest living baboons in Uganda. *Journal of Zoology* 149:344-365.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19(18):2496-2497.

Ruggiero, A. (1994). Latitudinal Correlates of the Sizes of Mammalian Geographical Ranges in South America. *Journal of Biogeography* 21:545-559.

Russo, I.R.M., Chimimba, C.T., Bloomer, P. (2006). Mitochondrial DNA differentiation between two species of *Aethomys* (*Rodentia: Muridae*) from southern Africa. *Journal of Mammalogy* 87(3):545-553.

Russo, I.M., Chimimba, C.T., Bloomer, P. (2010). Bioregion heterogeneity correlates with extensive mitochondrial DNA diversity in the Namaqua rock mouse, *Micaelamys namaquensis* (*Rodentia: Murida*e) from southern Africa - evidence for a species complex. *BMC Evolutionary Biology* 10(307).

Rydin, C., Kallersjo, M. (2002). Taxon sampling and seed plant phylogeny. *Cladistics* 18:485-513.

S

Saitou, N., Nei, M. (1987).The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.

Samuels, A., Altmann, J. (1991). Baboons of the Amboselli Basin- demographic stability and change. *International Journal of Primatology* 12(1):1-19.

Sanderson, M.J. (1997). A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* 14(12):1218-1231.

Sanderson, M.J. (2002). Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Molecular Biology and Evolution* 19(1):101-109.

Schlichting, C.D. (2004) *The role of phenotypic plasticity in diversification*. In Phenotypic Plasticity: Functional and Conceptual Approaches (DeWitt, T.J. ed).

Schuster, M., Duringer, P., Ghienne, J.F., Vignaud, P., Mackaye, H.T., Likius, A. and Brunet, M. (2006). The age of the Sahara Desert. *Science* 311(5762):821-821.

Scott, L. (1999). Vegetation history and climate in the Savanna biome South Africa since 190,000 ka: a comparison of pollen data from the Tswaing Crater (the Pretoria Saltpan) and Wonderkrater. *Quaternary International* 57-58:215-223.

Segal, C. (2008) *Foraging behaviour and diet in chacma baboons in Suikerbosrand Nature Reserve.* MSc dissertation, University of the Witwatersrand, South Africa.

Ségalen, L., Lee-Thorp, J.A., Cerling, T. (2007). Timing of C4 grass expansion across sub-Saharan Africa. *Journal of Human Evolution* 53(5):549-559.

Shackleton, N.J. et al. (1984). Oxygen isotope calibration of the onset of ice-rafting and the history of glaciation in the North Atlantic region. *Nature* 307:620-623

University of Cape Town

Shaw, P.A. and Thomas, D.S.G. (1996). The Quaternary palaeoenvironmental history of the Kalahari, Southern Africa. *Journal of Arid Environments* 32(1):9-22.

Shoubridge, E.A. and Wai, T. (2007) Mitochondrion in the germline and early development. *Current Topics in Developmental Biology* 77:87-111.

Singleton, M. (2002).Patterns of Cranial Shape Variation in the Papionini (Primates: Cercopithecinae). *Journal of Human Evolution* 42:547-578.

Sithaldeen, R., Bishop, J. M., Ackermann, R. R. (2009). Mitochondrial DNA analysis reveals Plio-Pleistocene diversification within the chacma baboon. *Molecular Phylogenetics and Evolution* 53 (3): 1042-1048.

Slatkin, M., Hudson, R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129:555 - 562.

Smit, H.A., Robinson, T.J., Van Vuuren, B.J. (2007). Coalescence methods reveal the impact of vicariance on the spatial genetic structure of *Elephantulus edwardii* (Afrotheria, Macroscelidea). *Molecular Biology and Evolution* 2680-2692.

Smit, H.A., Watson, J., Jansen van Vuuren, B. (2010). Relative importance of habitat connectivity in shaping the genetic profiles of two southern African elephant-shrews. *Journal of Biogeography* 37:857-864.

Song, H., Buhay, J.E., Whiting, M.F., Crandall, K.A. (2008). Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Sciences, USA* 105(36):13486-13491.

Sorenson, M.D., Quinn, T.W. (1998). Numts: A challenge for avian systematics and population biology. *Auk* 115(1):214-221.

Sorenson, M.D., Oneal, E., Garcia-Moreno, J., Mindell, D.P. (2003). More taxa, more characters: the hoatzin problem is still unresolved. *Molecular Biology and Evolution* 20:1484-1498.

Sousa, V.C., Fritz, M., Beaumont, M.A., Chikhi, L., 2009. Approximate Bayesian computation without summary statistics: the case of admixture. *Genetics* 181: 1507–1519.

Steiper, M.E., Young, N.M. (2006). Primate molecular divergence dates. *Molecular Phylogenetics and Evolution* 41:384-394.

Stephens, M. (2001). *Inference under the coalescent*. In Handbook of Statistical Genetics, eds D. J. Balding, M. Bishop and C. Cannings, John Wiley

Stokes, S. (1998). Punctuated aridity. *Palaeogeography Palaeoclimatology Palaeoecology* 137:305-322.

Stone, G. (2000). Phylogeography, hybridization and speciation. T*rends in Ecology and Evolution* 15(9):354-355.

Strasser, E., Delson, E. (1987). Cladistic analysis of cercopithecid relationships. *Journal of Human Evolution* 16:81-99.

University of Cape Town

Stute, M., Talma, A.S. (1997). Isotope techniques in the study of past and current environmental changes in the hydrosphere and the atmosphere. *Isotopic techniques in the study of environmental change*. *International Atomic Energy Agency, Vienna*: 307-318.

Sugawara, K. (1979). Socioecological study of a wild group of hybrid baboons between *Papio anubis* and *P. hamadryas* in the Awash Valley, Ethiopia. *Primates* 20:21-56.

Sutovsky, P., Moreno, R.D., Ramalho-Santos, J. et al. (1999). Ubiquitin tag for sperm mitochondria. *Nature* 402:371-372.

Swart, B.L., Tolley, K.A., Matthee, C.A. (2009). Climate change drives speciation in the southern rock agama (*Agama atra*) in the Cape Floristic Region, South Africa. *Journal of Biogeography* 36(1):78-87.

Swofford, D.L., Olsen, J.L., Waddell, P.J., Hillis, D.M. (1996). Phylogenetic inference. In: Hillis, D.M., Moritz, C., Mable, B.K. (eds) *Molecular systematics*. Sinauer Associates, Sunderland, MA, pp 407–514.

Szalay, F.S., Delson, E. (1979). Evolutionary History of the Primates Academic Press. New York.

T

Taanman, J. (1999). The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta* 1410:103-123.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.

Tamura, K., Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* 10:512-526.

Tamura, K., Dudley, J., Nei, M., Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 24:1596-1599.

Templeton, A.R. (1998). Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology* 7(4):381-397.

Thompson, J.D., Gibson, T.J., Plewniak,F., Jeanmougin,F., Higgins.D.G., (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25(24):4876-4882.

Thorington, R.W., Groves, C.P. (1970). *An annotated classification of the Cercopithecoidea.* In: Napier JR, Napier PH (eds) Old World Monkeys. New York Academic Press, New York.

Thorne, J.L., Kishino, H., Painter, I.S. (1998). Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15(12):1647-1657.

Thorne, J.L., Kishino, H. (2002). Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* 51(5):689-702.

Tolley, K.A., Tilbury, C.R., Branch, W.R., Matthee, C.A. (2004). Phylogenetics of the southern African dwarf chameleons, *Bradypodion (Squamata: Chamaeleonidae*). *Molecular Phylogenetics and Evolution* 30(2):354-365.

University of Cape Town

Tolley, K.A., Burger, M., Turner, A.A., Matthee, C.A. (2006). Biogeographic patterns and phylogeography of dwarf chameleons (*Bradypodion*) in an African biodiversity hotspot. Blackwell Publishing Ltd, pp 781-793.

Tolley, K.A., Braae, A., Cunningham, M. (2010). Phylogeography of the clicking stream frog *Strongylopus grayii* (Anura, Pyxicephalidae) reveals cryptic divergence across climatic zones in an abundant and widespread taxon. *African Journal of Herpetology* 59(1):17-3.

Tosi, A.J., Morales, J.C., Melnick, D.J. (1999). Y-chromosome phylogeny of the macaques (Cercopithecidae: *Macaca*). *American Journal of Physical Anthropology* 28(Suppl:266).

Tosi, A.J., Disotell, T.R., Morales, J.C., Melnick, D.J. (2003). Cercopithecine Y-chromosome data provide a test of competing morphological evolutionary hypotheses. *Molecular Phylogenetics and Evolution* 27:510-521.

Tsaousis, A.D., Martin, D.P., Ladoukakis, E.D., Posada, D., Zouros, E. (2005). Widespread recombination in published animal mtDNA sequences. *Molecular Biology and Evolution* 22(4):925-933.

Tung, J., Charpentier, M.J.E., Garfield, D.A., Altmann, J., Alberts, S.C. (2008). Genetic evidence reveals temporal change in hybridization patterns in a wild baboon population. *Molecular Ecology* 17(8):1998-2011.

Turelli, M., Barton, N..H, Coyne, J.A. (2001). Theory and speciation. *Trends in Ecology and Evolution* 16(7):330-343.

U - V

Udvardy, M.D.F. (1975). *A classification of the biogeographical provinces of the world*. Switzerland: IUCN Occasional Paper 18.

Ujvari, B., Dowton, M., Madsen, T. (2007). Mitochondrial DNA recombination in a free-ranging Australian lizard. *Biological Letters* 3(2):189-192.

Ursprung, H., Schabtac, E. (1965). Fertilization in tunicates- loss of paternal mitochondrion prior to sperm entry. *Journal of Experimental Zoology* 159(3):379.

Valentine, J.W. (1984). Neogene marine climate trends: implications for biogeography and evolution of the shallow-sea biota. *Geology* 12:647–59.

Vandamme, A-M. (2003). *Basic concepts of molecular evolution.* In: Salemi M, Vandamme A-M (eds) The phylogenetic handbook: a practical approach to DNA and protein phylogeny. Cambridge University Press, pp 1-21.

Vanderkuyl, A.C., Kuiken, C.L., Dekker, J.T., Perizonius, W.R.K., Goudsmit, J. (1995) Nuclear counterparts of the cytoplasmis mitochndrial 12S ribosomal - RNA gene- a problem of ancient DNA and molecular phylogenies. *Journal of Molecular Evolution* 40(6):652-657.

van Hooft, W.F., Groen, A.F., Prins, H.H.T. (2002). Phylogeography of the African buffalo based on mitochondrial and Y-chromosomal loci: Pleistocene origin and population expansion of the Cape buffalo subspecies. *Molecular Ecology* 11: 267-279.

van Treuren, R. et al. (1991) The significance of genetic erosion in the process of extinction. Genetic differentiation in *Saluia pratensis* and *Scabiosa columbaria* in relation to population size *Heredity* 66: 181-189.

University of Cape Town

van Zinderen Bakker, E.M. (1969.) The 'Arid Corridor' Between South-Western Africa and the Horn of Africa. *Palaeoecology of Africa* 4:139-140.

van Zinderen Bakker, E.M. (1978) *Quaternary vegetation changes in southern Africa*. Dr. W. Junk Publishers.

Vianna, J.A., Bonde, R.K., Caballero, S., Giraldo, J.P., Lima, R.P., Clark, A., Marmontel, M., Morales-Vela, B., De Souza, M.J., Parr, L., RodrÍGuez-Lopez, M.A., Mignucci-Giannoni, A.A., Powell, J.A., Santos, F.R. (2006). Phylogeography, phylogeny and hybridization in *trichechid sirenians*: implications for manatee conservation. *Molecular Ecology* 15: 433-447.

Vigilant, L.A. (1986). *Control region sequences from African populations and the evolution of human mitochondrial DNA*. University of California, Berkeley

Vigilant, L. (2009). Elucidating Population Histories Using Genomic DNA Sequences. *Current Anthropology* 50(2):201-212.

Vignaud, P., Duringer,P., Mackaye, H.T., Likius, A., Blondel, C. et al. (2002). Geology and paleontology of the Upper Miocene Toros-Menalla hominid locality, Chad. *Nature* 418, 152-155.

Vogel, J.C. (1983). Isotopic evidence for past climates and vegetation of South Africa. *Bothalia* 14:391-394.

Vrba, S. (1992). Mammals as a Key to Evolutionary Theory. *Journal of Mammology* 73(1):1-28.

W

Wakeley, J. (2008). Complex speciation of humans and chimpanzees. *Nature* 452(7184):E3-E4

Watson, D.F. (1992). *Contouring: a guide to the analysis and display of spatial data*. Pergamon Press, New York.

Weingrill, T., Gray, D.A., Barrett, L., Henzi, S.P. (2004). Fecal cortisol levels in free-ranging female chacma baboons: relationship to dominance, reproductive state and environmental factors. *Hormones and Behaviour* 45:259-269.

Whiten, A., Bryne, R.W., Barton, R.A., Waterman, P.G. and Henzi, S.P. (1991). Dietary and foraging strategies of baboons. *Philosophical Transactions of the Royal Society of London B* 334(1270):87-197.

Wildman, D.E., Bergman, T.J., Al-Aghbari,A., Sterner,K.N., Newman, T.K. and Phillips-Conroy, J.E. (2004). Mitochondrial evidence for the origin of hamadryas baboons. *Molecular Phylogenetics and Evolution* 32(1):287-296.

Wilfert, L., Kaib, M., Durka, W., Brandl, R. (2006). Differentiation between populations of a termite in eastern Africa: implications for biogeography. *Journal of Biogeography* 33:1993-2000.

Williams-Blangero, S., Vanderberg, J.L., Blangero, J., Konigsberg, L. and Dyke, B. (1990). Genetic differentiation between baboon subspecies: relevance for biomedical research. *American Journal of Primatology* 20: 67-81.

University of Cape Town

Williams, S.T. and Knowlton, N. (2001). Mitochondrial pseudogenes are pervasive and often insidious in the snapping shrimp genus *Alpheus*. *Molecular Biology and Evolution* 18(8):1484-1493.

Williams, F.L., Ackermann, R.R. and Leigh, S.R. (2007). Inferring Plio-Pleistocene southern African biochronology from facial affinities in Parapapio and other fossil papionins. *American Journal of Physical Anthropology* 132: 163-174.

Winney, B.J., Hammond, R.L., Macasero, W., Flores, B., Boug, A., Biquand, V., Biquand, S., Bruford, M.W. (2004). Crossing the Red Sea: phylogeography of the hamadryas baboon, *Papio hamadryas hamadryas*. *Molecular Biology and Evolution* 2819-2827.

Wright,S. (1943). Isolation by distance. *Genetics* 28:114 - 138.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics* 15:323-354.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19:395-420.

X - Z

Xu, J.P. (2005). The inheritance of organelle genes and genomes: patterns and mechanisms. *Genome* 48(6):951-958.

Yang, Z.H., Rannala, B. (2006) Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution* 23(1):212-226.

Zawada, P.K. (1997). Palaeoflood hydrology: method and application in flood-prone southern Africa. *South African Journal of Science* 93:111 -132.

Zhang, D.X., Hewitt, G.M. (1996). Highly conserved nuclear copies of the mitochondrial control region in the desert locust *Schistocerca gregaria*: Some implications for population studies. *Molecular Ecology* 5(2):295-300.

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.

Zinner,D. Greoneveld, L.F., Keller, C., Roos, C. (2009) Mitochondrial phylogeography of baboons (Papio spp.) - indication for introgressive hybridization? BMC Evolutionary Biology 9(83):doi:10.1186/1471-2148-1189-1183

Zinner, D., Arnold, L.M., Roos, C. (2009a) Is the New Primate Genus *Rungwecebus* a Baboon? PLoS ONE 4(3):e4859. doi:4810.1371/journal.pone.0004859

Zischler, H., Geisert. H., Castresana, J. (1998). A hominoid-specific nuclear insertion of the mitochondrial D-loop: implications for reconstructing ancestral mitochondrial sequences. *Molecular Biology and Evolution* 15(4):463-469.

Zunino, G.E., Bravo, S., Murad, F. F., Reisenman, C. (1996). Characteristics of two types of habitat and the status of the howler monkey (*Alouatta caraya*) in Northern Argentina. *Neotropical Primates* 4(2):48-50.

University of Cape Town

Zwickl, D.J. and Hillis, D.M. (2002). Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology* 51:588-598.