Genetic structure of the savannah elephant population (*Loxodonta africana* (Blumenbach 1797)) in the Kavango-Zambezi Transfrontier Conservation Area

By

Alida de Flamingh

A dissertation submitted in partial fulfilment of the requirements for the degree

Magister Scientiae (Zoology)

In the Faculty of Natural and Agricultural Science

University of Pretoria

2013

Genetic structure of the savannah elephant population (*Loxodonta africana* (Blumenbach 1797)) in the Kavango-Zambezi Transfrontier Conservation Area

by

Alida de Flamingh

Supervisor: Professor Rudi J. van Aarde
Conservation Ecology Research Unit
Department of Zoology and Entomology
University of Pretoria
Pretoria
South Africa

Co-Supervisor: Dr. Catherine L. Sole
Scarab Research Group
Department of Zoology and Entomology
University of Pretoria
Pretoria
South Africa

ABSTRACT

Earlier studies investigated the genetic structure of fragmented or isolated elephant populations by comparing the genetic characteristics of pre-defined populations. This study aimed to determine if there was genetic evidence for spatial structuring in a continuous elephant population in the Kavango-Zambezi Transfrontier Conservation Area (KAZA-TFCA).

I sequenced one mtDNA gene region for 88 individuals and genotyped 100 individuals for 10 nuclear microsatellite loci. Bayesian Clustering Algorithms incorporated in the program Geneland were used to identify groups of genetically similar individuals. An Analysis of Molecular Variance (AMOVA) determined if these groups (henceforth referred to as sub-populations) were significantly differentiated. I used a Geographic Information System (GIS)

landscape genetic toolbox to identify areas in the landscape with high genetic divergence between individual samples to determine if there were identifiable genetic barriers in the landscape.

There were three significantly differentiated mtDNA sub-populations (F_{st} = 0.787), and two nDNA sub-populations that were not significantly differentiated (Fst = -0.02; Rst = -0.045), implying obstructed mtDNA, but high nDNA gene flow across the study region. Also, gene flow was apparent between Chobe and Kafue National Parks, where telemetry data has as of yet not recorded inter-population movements between these parks.

The three mtDNA sub-populations were geographically differentiated and followed political boundaries as apparent sub-populations in Botswana, Zambia and Zimbabwe. The differences between mtDNA and nDNA genetic structuring may be explained by i) historical events that shaped the current genetic structure (e.g. through founder-effects and persistent poaching hotspots) and ii) intrinsic variables that influence genetic structure at a local scale (e.g. through resource dependencies and social behaviour). The KAZA elephant population has a genetic diversity (mtDNA diversity as the pairwise number of differences (π) = 2.59; nDNA diversity as the mean alleles/locus and H_e = 7.5, 0.71) higher than other southern African populations, and inter-population movements may be responsible for maintaining this genetic diversity.

I recommend continued support for conservation initiatives that aim at maintaining and restoring connectivity between populations through landscape linkages, which in so doing may ensure inter-population gene flow and uphold the current genetic state of the KAZA-TFCA elephant population.

To my parents, for your support, encouragement and perpetual love!

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Professor Rudi van Aarde, for his encouragement, trust and imperative support throughout my studies. Thank you Prof for giving me this chance, and for always seeking the best in me as a scientist. Your encouragement has greatly influenced, and I am sure will continue to influence my life as a scientist.

Catherine Sole, thank you for your invaluable support and patience, and for always encouraging me to seek out and investigate everything I don't know about genetics. Thanks for allowing me to follow my own mind!

Many thanks to all my friends and colleagues in CERU, specifically Morgan Trimble, Tamara Lee, Teri Ott, Carrie Roever, Lillian Scholtz, and Jo Fourie, without your help I would not have been able to tackle this MSc. To my departmental friends, please know that I appreciate every single one of you, thank you for your support and encouragement. I would also like to thank Nicolene Smit and the DNA Sequencing Facility at the University of Pretoria for their respective parts in the laboratory work.

Thank you to the International Fund for Animal Welfare, the Zambian Wildlife Authority, the Department of Wildlife and National Parks (Botswana), Elephants Without Borders, and the Conservation Foundation (Zambia) for funding this project and facilitating sample collection. Thank you also to the University of Pretoria and the National Research Fund for my personal grants.

Thank you to my wonderful parents. To my dad, thank you for always enthusiastically entertaining my sometimes endless scientific rants, and to my mom, for everything you do to give me courage to continue pursuing my passions. To the rest of my family, Tilly and Willem Buys especially, thank you for all the encouragement, love, support and enthusiasm!

DISCLAIMER

I hereby declare that all of the work is my own and that it has never been submitted for any other degree, or to any other university. Furthermore, I affirm that I have thanked and acknowledged all those that have helped or contributed to this thesis.

Alida de Flamingh

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

GLOSSARY OF ACRONYMS AND DEFINITIONS

ACRONYMS

BCA ……..Bayesian Clustering Algorithms

CERU…….Conservation Ecology Research Unit, University of Pretoria, Department of Zoology and Entomology

CITES…….The Convention on International Trade in Endangered Species of Wild Fauna and Flora

IBD……...……..Isolation By Distance

IUCN …….The International Union for the Conservation of Nature

KAZA-TFCA…….Kavango-Zambezi Transfrontier Conservation Area

mtDNA ……..Mitochondrial (mt) deoxyribonucleic acid (DNA)

nDNA…….Nuclear (n) deoxyribonucleic acid (DNA)

DEFINITIONS

Female-based sub-population	Sub-populations that were identified using mtDNA. A comparison between mtDNA inferred populations and nDNA inferred populations are used to identify gender-biased population structuring.
Gene dispersion	The dispersal of genes through gene flow.
Genetic divergence	The evolutionary change in allele frequencies between populations or individuals (Allendorf et al., 2013). In this thesis, genetic divergence is used as a measure of the genetic differentiation between populations and individuals.
Genetic diversity	The extent of genetic variation in a population or species, or across a group of species. Measures of genetic diversity include

	heterozygosity, allelic diversity and heritability (Frankham et al., 2004).
Genotype error	Errors that arise during the genotyping process that cause the resulting genotype to be different to the actual genotype of the sample.
Isolation By Distance:	The theory of Isolation By Distance (first proposed by Sewell Wright, 1943) suggests that the geographic distance between individuals or populations can explain the genetic differences between them, where populations or individuals that are further apart geographically will have a higher degree of genetic divergence than populations or individuals that are close to each other.
Male-based sub-populations	Sub-populations that were identified using nDNA microsatellites. A comparison between mtDNA inferred populations and nDNA inferred populations are used to identify gender-biased population structuring.
Philopatry	The behaviour where an individual remains with its natal group.
Population fragmentation	Population fragmentation occurs when a previously continuous population is broken up into fragments. In this thesis I refer to populations that are genetically and/or spatially fragmented. Genetic population fragmentation is characterized by limited gene flow between fragments, while spatial population fragmentation occurs when these fragments are not connected geographically. Unless stated otherwise, population fragmentation refers to populations that are spatially fragmented.

CHAPTER 1
INTRODUCTION

Genetic diversity is sustained by gene flow between populations, described by Allendorf and Luikart (2007) as "the cohesive force that holds together geographically separated populations into a single evolutionary unit – the species". Consequently, populations that are spatially fragmented may also be genetically fragmented (Gerlach and Musolf, 2000). Genetic fragmentation may accelerate inbreeding and reduce the ability of populations to adapt to environmental changes, thus increasing the risk of extinction (Frankham et al., 2004). Conservation aims at minimising the risk of extinction and therefore also at maintaining genetic diversity.

African savannah elephants are discontinuously distributed across the continent (van Aarde and Jackson, 2007). The discontinuous distribution of Africa's elephants may be a response to anthropogenically induced landscape transformation (Hoare and du Toit, 1999; Osborn and Parker, 2003; Graham et al., 2009), obstruction of dispersal opportunities (e.g. Loarie et al., 2009), and extirpation of local populations by hunters (Whitehouse and Harley, 2001). Elephants avoid intensely transformed landscapes (Hoare and du Toit, 1999; Graham et al., 2009) and mostly live in protected areas (Douglas-Hamilton et al., 2005; Lee, 2013), albeit that some 80% of their present distributional range extends beyond the borders of protected areas (Hoare, 1999).

Studies on the genetic characteristics of elephant populations lend support to the implied behavioural, spatial and demographic structuring of populations (Archie et al., 2008; Chiyo et al., 2011; Ahlering et al., 2012; Archie and Chiyo, 2012). The genetic structure of elephant

populations have been studied at a continental scale, nationally, and within defined protected areas (Appendix I). Continental scale studies ascribe genetic variation between populations to historical climatic conditions, Pleistocene refugia and long-term geographic barriers that shaped the genetic structure of populations through allopatric isolation (Eggert et al., 2002; Nyakaana et al., 2002). Locally, genetic variation has been accredited to gender differences in space use and social behaviour (Archie et al., 2006; Archie et al., 2008).

Most of these earlier studies were conducted on spatially fragmented elephant populations (Nyakaana and Arctander, 1999; Okello et al., 2008b; Munshi-South, 2011). Studying the genetic structure of fragmented populations detracts from inferences of the forces at work that may give rise to genetic structure under natural conditions. Such forces could include geographic barriers (Rueness et al., 2003; Eriksson et al 2004) historical phylogeography and glacial refugia (Nyakaana et al., 2002; Schmitt et al., 2005) as well as the distribution of key resources (Chamaillé-James et al., 2008; Harris et al., 2008), vegetation heterogeneity (Young et al., 2009) and density dependent dispersal (Young and van Aarde, 2010).

Many of these earlier studies assume *a priori* populations or sub-populations, where the genetic characteristics of pre-defined populations and sub-populations are compared. Consequently, the findings of these studies may not hold for elephants that form part of a continuous population where their distributional range precludes the *a-priori* delineation of sub-populations.

The elephants in the Kavango-Zambesi Transfrontier Conservation Area (KAZA-TFCA) form part of such a widespread population (Chase and Griffin, 2006; Chase, 2007). Added to this, these elephants have not been exposed to culling like many populations elsewhere in southern Africa. Although elephants, especially those on the periphery of the KAZA-TFCA, have been

subjected to poaching[1], the intensity of poaching in the core of the KAZA-TFCA is less than that of elephant populations elsewhere (Burn et al., 2010). Therefore, on a landscape level, the elephant population in the KAZA-TFCA may represent a naturally structured population.

In this thesis I investigate the genetic structure of this naturally structured elephant population. This thesis is comprised of five chapters. The first chapter serves as a general introduction to the genetic structure of Africa's elephant populations. In the second chapter I describe the study area (the Kavango-Zambezi Transfrontier Conservation Area) and provide a home range analysis for 12 elephants to illustrate the roaming range of elephants in the areas. In the third chapter I provide a DNA extraction procedure that, when combined with the described genotype error eliminating protocols increases amplification success and genotype accuracy. In the fourth chapter I compare the mtDNA-based and nDNA-based genetic structure of the KAZA-TFCA elephant population. I also investigate gene flow across the KAZA-TFCA and identify genetic discontinuities that support the different genetic sub-populations of male and female elephants. In the final chapter I validate the genetic structure of this population.

AIM OF THIS STUDY

The aim of this study was to determine if elephants in the KAZA-TFCA are structured into genetically distinct sub-populations. Furthermore, I intended to ascertain whether these sub-populations, if present, had gender specific genetic structures. To my knowledge, these questions were the first to investigate if a naturally structured elephant population that occurs across a continuous landscape, where space is not a limiting resource, adheres to detectable genetic structuring.

[1] as is evident in the largest ivory seizure on record in 2002 that included poached ivory from the Kafue National Park in Zambia (Wasser et al., 2007)

WORKING HYPOTHESES

- There will be high gene flow for both nDNA and mtDNA across the entire KAZA-TFCA.
- The genetic structure of female-based sub-populations will show a higher degree of genetic divergence than male-based sub-populations.

PREDICTIONS AND EXPECTATIONS

I appraise these hypotheses by expecting the following,

- I expect elephants in the KAZA-TFCA to form part of a continuous, widespread population (Chase and Griffin, 2006; Chase 2007).
- I expect female elephants to restrict their roaming to smaller areas than male elephants (Hall-Martin, 1987; Jackson and Erasmus, 2005).
- I expect gender differences in genetic population structures, where female-based sub-populations show higher degrees of genetic divergence between sub-populations than males-based sub-populations (Nyakaana and Arctander, 1999; Nyakaana et al., 2002; Okello et al., 2008b; Ishida et al., 2012).

CHAPTER 2

STUDY REGION

THE KAVANGO-ZAMBEZI TRANSFRONTIER CONSERVATION AREA (KAZA-TFCA)

The Kavango-Zambesi (KAZA) region spans five southern African countries – Botswana, Namibia, Angola, Zambia and Zimbabwe (Figure 1) and includes communal, state and privately owned land, where land-use ranges from agriculture to nature-based tourism.

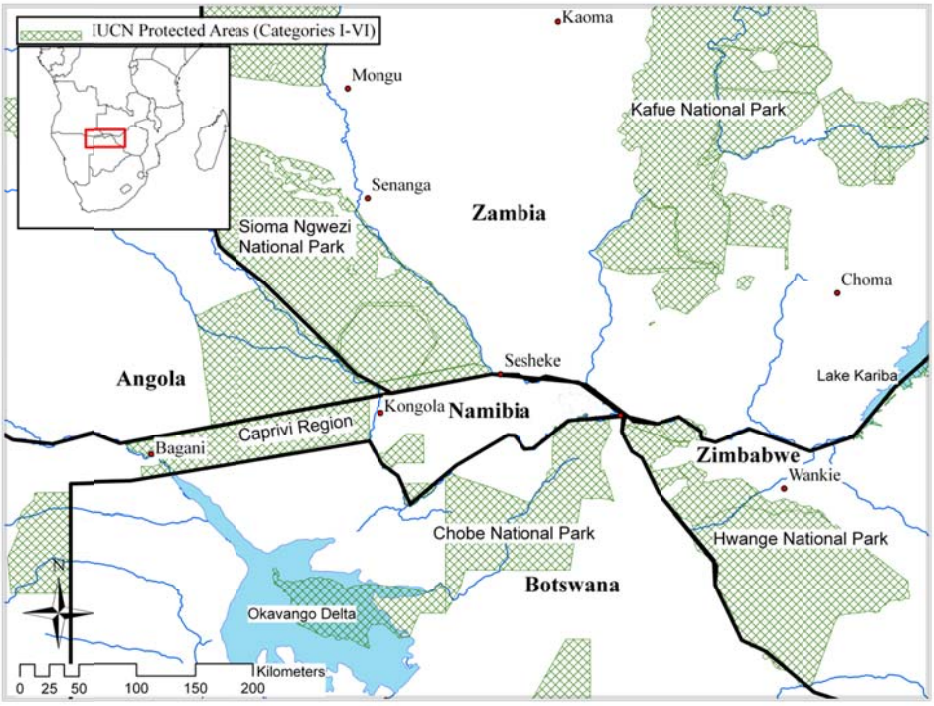


Figure 1. The greater Kavango-Zambesi region spans across Botswana, Namibia, Angola, Zambia and Zimbabwe, and includes a range of IUCN protected areas (Categories I to VI, Appendix II).

The KAZA region experiences distinct rainfall seasons, where the wet season lasts from November to March in the central region (Botswana) and from December to March in the northern region (South Kafue). Average annual rainfall[2] ranges from 430mm in Botswana (Okavango – Seronga) to 783mm in Kafue (South Kafue – Ngoma weather station) (Jackson and Erasmus, 2005).

Included in the greater Kavango-Zambezi region is the Kavango-Zambezi Transfrontier Conservation Area (KAZA-TFCA), which was formally established in 2006[3] (TCC, 2006; Metcalfe and Kepe, 2008). The KAZA-TFCA initiative, however, awaits implementation. The KAZA-TFCA includes approximately 278 000km^2 (Metcalfe and Kepe, 2008), equalling the size of entire African countries like Gabon (267 668km^2) and Burkina Faso (272 967km^2). The elephant populations in the five countries that form part of the KAZA-TFCA are some of the largest in Africa, with about 130 000 elephants living in northern Botswana, 16 000 in north-eastern Namibia, and 50 000 in north-western Zimbabwe (inferred from Chase and Griffin (2005), Chase (2011), and The Elephant Database (2011)).

The KAZA-TFCA is an initiative that aims at conserving biodiversity in a way that benefits the livelihoods of people (Hanks, 2001). Elephants feature strongly in proposed management options for this region, specifically in those options that aim at preserving animal movement corridors and decreasing human-wildlife conflict (TCC, 2006). These management options are important, as elephants in this region roam beyond the boundaries of protected areas (Appendix III). The

[2] Average annual rainfall was calculated using 20 years of data in Botswana, and 17 years of data in South Kafue. Wet seasons are defined by those months receiving more than 10 % of the total annual rainfall (Jackson and Erasmus, 2005)

[3] The original concept of the Kavango-Zambesi Transfrontier Conservation initiative was conceived in 2003 by the tourism and wildlife sector ministers from Angola, Botswana, Namibia, Zambia, and Zimbabwe (Metcalfe and Kepe, 2008). The memorandum of understanding, however, was only signed in 2006 after the completion of a pre-feasibility study, where after the five partner countries agreed to jointly work on the implementation of the KAZA-TFCA (TCC, 2006).

development of functional linkages between existing protected areas in the region is in agreement with proposed population management strategies that call for the establishment of 'megaparks' to stabilise regional population growth through demographic and spatial responses to resource heterogeneity (van Aarde et al., 2006; van Aarde and Jackson, 2007; Young et al, 2009).

Earlier studies show that elephants in the KAZA region cross international boundaries (Chase and Griffin, 2006; Chase 2007), where suitable habitat link elephant populations from several protected areas (Roever et al., 2013a). Here anthropogenic activities may have altered local, but not international movements (Hanks, 2003; Chase, 2007; Chase and Griffin, 2009; Roever et al., 2013b). My analysis of the roaming patterns for 12 elephants tracked in the region supports earlier notions that some breeding herds as well as bulls cross international borders[4]. My analysis also supports previous findings on gender differences in home-range sizes (Jackson and Erasmus, 2005; Legget, 2006; Shannon et al., 2006); with bulls having significantly larger annual home ranges than breeding herds (Appendix V).This suggests that females restrict their roaming to smaller areas than males.

Based on the apparent distribution of elephants in the region I expect that the elephants in the KAZA-TFCA will be part of a continuous, wide ranging population, with high gene glow across the entire region for both nuclear and mitochondrial DNA. Despite this continuous distribution I expect that the gender differences in home range sizes will be reflected in the genetic structure of the KAZA-TFCA's population, where the genetic structure of female-based sub-populations will have a higher degree of genetic differentiation than male-based sub-populations. These

[4] Refer to Appendix (IV) for a description of the spatial data and methods used, and Appendix III for individual maps of the 12 kernel home ranges.

expectations are based on previous studies on the genetic structure of elephant populations that show high gene flow across landscapes and gender differences in genetic structure of populations (e.g. Archie et al., 2008; Okello et al., 2008b).

Elephant faecal samples were collected from 120 distinct locations across the KAZA-TFCA (Figure 2). To minimise the possibility of collecting more than one sample from a single individual, samples were once off collected at least 5km apart along a unidirectional transect.

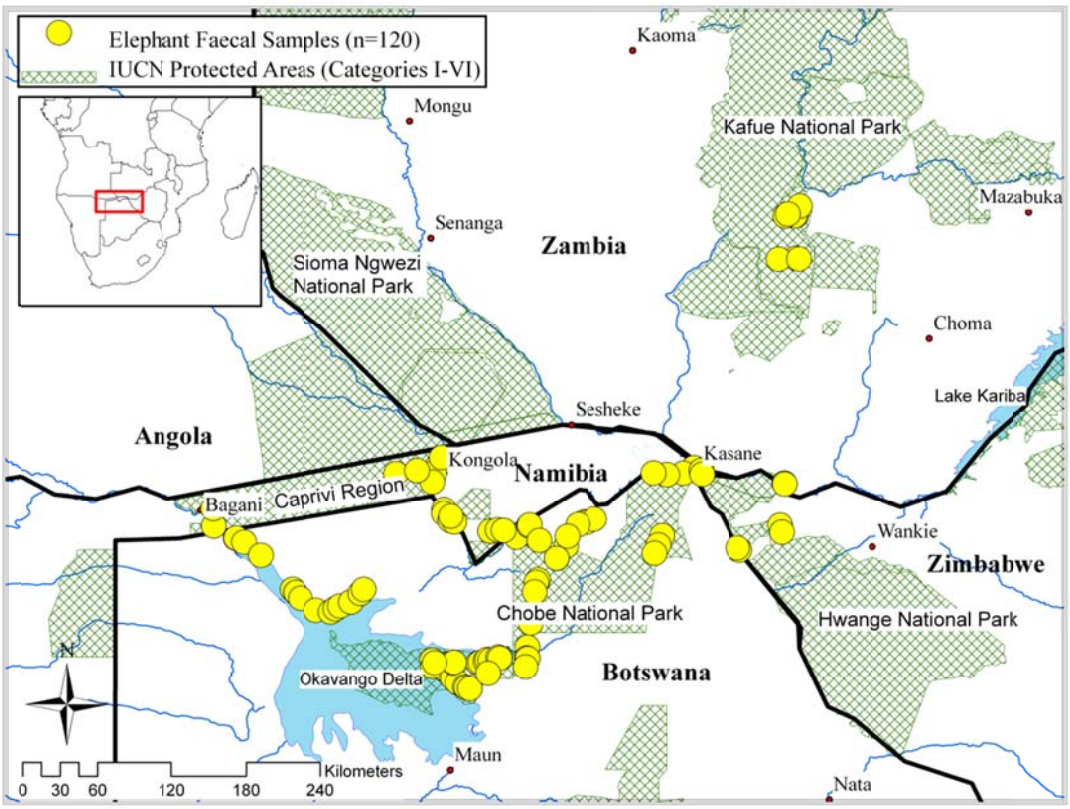


Figure 2. The locations across parts of the KAZA-TFCA where faecal samples (n=120) were collected during September 2010.

CHAPTER 3
MOLECULAR METHODOLOGY
USING FAECAL DNA TO NON-INVASIVELY ESTIMATE GENOTYPES

INTRODUCTION

Throughout the last few decades, genetic techniques and applications have advanced considerably. In just half a century we have gone from discovering the double helix structure of DNA (deoxyribonucleic acid (Watson and Crick, 1953)), to the possibility of sequencing an entire diploid mammalian genome in less than a day (Branton et al., 2008). These advances have no doubt been augmented by the concurrent development in electronic computational abilities (Brooks, 1994).

One such advance is the ability to now extract DNA from non-traditional sources such as fossil remains and faecal samples, which often contain DNA that has been severely degraded, through UV and temperature degradation and chemical or enzymatic degradation (Pääbo et al., 2004; Murphy et al., 2007). The use of faecal samples for DNA extraction has not only increased the sources of genetic material available for research, but has also reduced the efforts needed to collect genetic material. Using faecal samples as a DNA source facilitates studies on animals that are elusive or dangerous (in general - Piggott and Taylor, 2003; bears - Bellemain et al., 2005; elephants - Eggert et al., 2003), difficult to catch (birds - Regnaut et al., 2006; bats - Puechmaille and Petit, 2007) or species that are endangered (marsupials - Piggott et al., 2006). Examples of this are not limited to but include using faecal DNA for genetic studies on the African savannah elephant (*Loxodonta africana*), that often occur in areas that are difficult to access (Georgiadis et al., 1994; Eggert et al., 2003; Nyakaana et al., 2002; Comstock et al., 2003; Ishida et al., 2012).

Additionally, faecal samples can be used to create a geo-referenced genetic database to track illegal ivory (see Wasser et al., 2008; Ishida et al., 2012).

The most pertinent problem with the use of faecal DNA is the low DNA quantity and quality due to the presence of PCR inhibitors (Monteiro et al., 1997), which often leads to incorrect genotyping of samples (Taberlet et al., 1999). With the use of mtDNA sequencing techniques, genotype errors can be reduced or eliminated by manually viewing each individual sequence and the base pair mutations that it contains, and also by comparing study sequences to published sequences on genetic databases such as the National Centre for Biotechnology Information (NCBI) genetic sequence database Genbank.

Microsatellite genotype errors can arise at any point during the genotyping process (sampling, DNA extraction, molecular and data analysis) and are caused by a multitude of factors such as chance occurrences, human caused errors (due to improper training and supervision – Paetkau, 2003) or errors that arise due to technical discrepancies (Bonin et al., 2004). Genotype errors can be defined as those errors that arise during the genotyping process that cause the resulting genotype to be different to the actual genotype of the sample, and the per-locus genotype error is the proportion of single locus genotypes with one or both alleles scored incorrectly (Taberlet et al., 1999; Bonin et al., 2004; Hoffman and Amos, 2005). Potential genotype errors include null allele amplification (Dankin and Avise, 2004) and stuttering (Shinde et al., 2003), false allele amplification (Taberlet et al., 1996) and allelic dropout (Taberlet et al., 1999). Null allele genotype errors arise when one (or both) of the two alleles (for diploid organisms) does not amplify during Polymerase Chain Reaction (PCR). If null allele amplification occurs and only one allele amplifies, the resulting genotype will indicate that the individual or organism has homozygote alleles at that specific locus (Dankin and Avise, 2004). Excessive null allele

amplification will present as a higher than expected homozygote presence in a population, which can be calculated by determining deviations from the Hardy-Weinberg equilibrium hypothesis (Gomes et al., 1999; Xu et al., 2002; Leal, 2005). Genotype errors related to stuttering and false allele amplification occur when there are slight changes in the size of the amplified alleles during PCR due to Taq DNA polymerase slippage (Taberlet et al., 1996; Shinde et al., 2003).

Several approaches to correct or account for genotype errors have been suggested and broadly speaking, these approaches can be divided into two categories. The first category includes correctional methods that focus on pre-genescan and pre-analysis corrections, for example, the multiple-tubes approach (Navidi et al., 1992; Taberlet et al., 1996) where individual samples are amplified independently multiple times. The second category focuses on post-genescan correctional methods where genotype errors are statistically calculated and corrected for. For instance, the program MICROCHECKER (Van Oosterhout et al., 2004) compares observed allele frequencies to expected allele frequencies through the Hardy-Weinberg equilibrium equation (HWE). Deviations from the expected HWE are used to determine the appropriate genotype error correctional factors and also enable the identification of loci where genotype errors are high.

In lieu of these technological advances and the associated progress of genetic techniques and applications, there seems to remain some doubt regarding the use of faecal DNA for accurate and easy genotyping (Taberlet et al., 1999; Fernando et al., 2003; McKelvey and Schwartz, 2004; Stenglein et al., 2010; Panasci et al., 2011). Because of these uncertainties, I thought it necessary to investigate the use of faecal samples as an effective source of mtDNA and nDNA, and an efficient DNA source despite the genotyping errors associated with degraded DNA. The ability to use faecal material as an effective and efficient source of DNA will not only help researchers,

but may also assist conservation efforts that rely on information from non-invasive DNA sampling methods (Taberlet et al., 1999). Non-invasive faecal sampling is relatively easy and sometimes the only option for collecting DNA samples. To investigate the efficacy and efficiency of using elephant faecal DNA as DNA source, I applied a novel approach to existing molecular techniques. I investigated the efficacy of using faecal DNA as DNA source through calculating the ability to positively extract and sequence long-stranded (600-700bp fragments) (Frantzen et al., 1998) elephant specific mtDNA from faecal samples. I also investigated the efficiency of using faecal DNA as DNA source by calculating the genotype error associated with each microsatellite marker, and the efforts needed to acquire accurate genotypes.

METHODS AND MATERIALS

Sample collection and preservation

Elephant faecal samples were collected from the Kavango-Zambesi Transfrontier Conservation Area (KAZA-TFCA) (Chapter 2). Samples were collected in a similar manner to the collection protocols described by Georgiadis et al., (1994) and Fernando et al., (2003). Elephant dung samples (golf ball size) were collected from the outer layer of a fresh (recently defecated) elephant dung ball because it contains the highest concentration of elephant epithelial cells (Marrero et al., 2009). Each sample was then stored at room temperature in a plastic honey jar containing absolute ethanol as preservative (Fernando et al., 2003; Vidya and Sukumar, 2005) until processing and DNA extraction.

DNA extraction

The contents of each sample bottle were homogenised. Approximately 2ml of this mixture was placed in a 2ml microcentrifuge tube, and centrifuged for 30 minutes, after which the supernatant was discarded. This process was repeated three times, and the remaining pellet was used as

starting material for DNA extraction. The Qiagen DNA Stool Mini Kit (Southern Cross Biotechnology, Cape Town, SA) was used to extract DNA following a modified protocol for isolation of DNA from stool for human DNA analysis (see Appendix VI for a detailed description of the modified protocol steps). Modifications to the protocol included increasing the vortex, centrifuge and incubation times. More specifically, the initial vortex time was extended to 15 minutes after which the sample was centrifuged for 30 minutes. I increased both the vortex and incubation times to five minutes when adding the InhibitEX tablet, and centrifuged this suspension for an increased time of 6 minutes to pellet the stool particles. I incubated the sample for 30 minutes at 70°C. I decreased the amount of elution buffer used in the final step of the protocol to 100µl, and increased the incubation and centrifuge times in this final step to 10 minutes. I applied this DNA extraction procedure to 120 elephant faecal samples in total.

Mitochondrial DNA amplification and analysis

Positive extraction of elephant specific DNA was tested for the 120 samples by amplifying a 509 nucleotide base pair (bp) fragment of the 5 prime end of the D-loop control region with primers LafCR1 and LafCR2 (Nyakaana and Arctander, 1999). The PCR was set up using a Qiagen Multiplex PCR kit (1000 reactions) (Southern Cross Biotechnology, Cape Town, SA) following the standard protocol for multiplex PCRs. All PCRs included a no-template control to test for possible contamination. Cycling conditions encompassed a DNA Polymerase activation step of 15 minutes at 95°C followed by a three step cycle of denaturation (30 seconds at 94°C), annealing (90 seconds at 57°C) and extension (90 seconds at 72°C). The three step cycle was repeated 40 times and was followed by a final extension of 10 minutes at 72°C. All PCRs were done using a Veriti 96 well Thermal Cycler from Applied Biosystems (Foster City, CA).

The PCR product was then tested for positive amplification by means of gel electrophoreses. The products were combined with a loading buffer (10mM Tris-HCl (pH7.6), 0.03% bromophenol blue, 0.3% xylene cyanol FF, 60% glycerol and 60 mM ECTA) and sizing agent (GeneRuler^{TM}100bp Plus DNA Ladder). Samples were run for 25 minutes at 100 Volts on a GelXLultra (Labnet international inc.) electrophoreses machine using a 1.5% Agarose gel and viewed using a Viber Lourmat ultra violet viewer in the freeware computer program DScaler (version 4.1.15 © 2005). Positive samples were then purified using the NucleoSpin®Extract ll purification protocol based on the manufacturer's specifications (Machery-Nagel, GmbH and Co, KG, Düren). Samples were then cycle sequenced using Big Dye Terminator ver. 3.1 (Applied Biosystems, Foster City, CA) and subjected to the Sodium-Acetate precipitation (see Appendix VII for the respective reagents and protocols used in Cycle Sequencing and Sodium Acetate Precipitation). Samples were sequenced on an Applied Biosystems 3130xl (Life Technologies, Carlsbad, US) automated sequencer at the DNA Sequencing facility of the University of Pretoria, South Africa. Sequences were viewed and processed using the program CLC Bio Genomics workbench (CLC bio, Cambridge, MA). Sequence alignment and data formatting was done using ClustalW (Thompson et al., 1994) and Webprank (Goldman Group Software©, European Bioinformatics Institute 2011). DnaSP version 5.10.01 (Librado and Rozas, 2009) was used to investigate haplotype diversity.

Nuclear microsatellite amplification

The 120 extracted DNA samples were used to genotype 12 fluorescently labelled microsatellite markers. Two microsatellite genotype panels were compiled to include the 12 markers; Panel 1 contained five of the markers that were multiplexed into one PCR setup (Multiplex One), and panel 2 contained seven markers that were divided into three multiplexed PCR setups (Multiplex Two, Multiplex Three and Multiplex Four respectively) (Table 1). The PCRs were done using

the Qiagen Multiplex PCR kit (1000 reactions) (Southern Cross Biotechnology, Cape Town, SA) and included no-template controls to test for possible contamination. Cycling conditions followed the protocol stipulated in the Qiagen Multiplex PCR kit for the Amplification of microsatellite loci or small amplicons (up to 0.5kb) using multiplex PCR, and included a DNA Polymerase activation step of 15 minutes at 95°C followed by a three step cycle of denaturation (30 seconds at 94°C), annealing (90 seconds at the annealing temperature of the markers in the PCR setup) and extension (60 seconds at 72°C). The three-step cycle was repeated 40 times for Panel 1, and 38 times Panel 2. Decreasing the number of cycle repeats was necessary to eliminate background stutter found in some of the amplified markers in Panel 2. The three-step cycle was followed by a final extension of 35 minutes at 60°C to allow for the generation of A-overhangs by the HotStarTaq DNA polymerase which is necessary for high-resolution capillary and gel electrophoresis analysis. All PCRs were done using a Veriti 96 well Thermal Cycler from Applied Biosystems (Foster City, CA). The multiplex PCR products were then tested for positive amplification by agarose gel electrophoresis using a 2% Agarose gel and a run time of 16 minutes with the same reagents and lab equipment specified for mtDNA amplification tests.

After identifying the samples that amplified positively for the different PCR setups, I created a dilution gradient (PCR product: SABAX water – 1:100, 1:50, 1:25, and 1:10) for a subset of samples to test which dilution will generate the clearest marker visibility when subjected to a genescan fragment analysis. Multiplex One, Multiplex Three and Multiplex Four produced the clearest marker visibility at a dilution of 1:25 (PCR product:SABAX water), while Multiplex Two worked best at a 1:50 dilution. Subsequent genescan fragment analysis for all the samples was done using the above-specified dilutions.

Table 1. Microsatellite marker names and individual marker characteristics as found in the published data and the generated data.

					Published data			Project data			
Marker Name	**Locus Name**	**Repeat Motif**	**Size (bp)**	**Ta-°C**	**No of alleles**	**HO[5]**	**HE[6]**	**Panel**	**No of alleles**	**HO**	**HE**
LaT08[7]	1	(TAGA)16	166–234	56	10	0.818	0.854	1	10	0.828	0.851
Lat13[4]	2	(CATC)21	234–262	56	8	0.576	0.699	1	7	0.559	0.699
Lat17[4]	3	(GGAT)15… (GGAT)10	323–355	56	8	0.848	0.867	1	8	0.780	0.835
Lat18[4]	4	(CCAT)22	286–318	56	8	0.788	0.826	1	8	0.655	0.829
Lat24[4]	5	(GGAT)22	211–231	56	6	0.879	0.785	1	9	0.778	0.858
FH1[8]	6	(CA)12	81	55	3	0.217	0.258	2	6	0.545	0.569
LA3[9]	7	(CA)10	165–171	55	3	0.521	0.527	2	3	0.857	0.503
FH39[5]	8	(CA)18	242	60	7	0.826	0.742	2	10	0.632	0.779
FH102[5]	9	(CT)11(CA)14	179	60	3	0.435	0.492	2	9	0.546	0.559
FH19[5]	10	(CA)15	185	60	7	0.739	0.743	2	9	0.541	0.689
LA5[6]	11	(CA)13	130–154	52	7	0.575	0.377	2	7	0.526	0.618
Lat25[4]	12	(CCAT)15	298–318	52	6	0.758	0.755	2	6	0.628	0.779

[5] Observed Heterozygocity (HO)
[6] Expected Heterozygocity (HE)
[7] Archie et al., 2003
[8] Comstock et al., 2000
[9] Eggert et al., 2000

Genescan fragment analysis was done using an ABI Applied Biosystems 3130xl (Life Technologies, Carlsbad, US) automated sequencer for panel one and an Applied Biosystems 3500xl (Life Technologies, Carlsbad, US) automated sequencer for panel 2, at the DNA Sequencing facility of the University of Pretoria, South Africa. The resulting genescan fragment analysis protocol included the DS 33 Filter set G5 using the 500bp size standard GeneScan^{TM} - LIZ500^{TM} (Applied Biosystems – manufactured in Warrington WA1 4SR, United Kingdom) at a 14:1000 ratio of Formamide (provided by the DNA Sequencing Facility). Genescan fragment analysis results were viewed and scored using the computer program GeneMapper ® Software Version 3.7 (Applied Biosystems, Foster City, USA). Scored data was then exported in table format to Excel spreadsheets for further analysis.

I used the program Arlequin (Excoffier and Lischer, 2010) to calculate mtDNA diversity as the average pairwise differences (π), and nDNA diversity as the average number of alleles per locus and the Expected Heterozygocity averaged across all loci (He).

Pre- and post-genescan genotype error methodology

Pre-genescan genotype eliminating methods included repeating the genotyping process (PCR and fragment analysis) for samples that amplified as homozygote genotypes, samples that did not amplify clearly (where marker profiles showed stutter and unclear allele peaks) and samples where amplified alleles were unique (not repeated in the population). Samples that amplified as homozygotes two or more times were accepted as true homozygotes. Sample genotyping was repeated up to 4 times per locus, after which samples were scored as missing data if the genotypes were still unclear. For fine scale parentage analysis I suggest increasing the genotype

repeats to at least seven, however, for the purpose of this large scale population study we only repeated genotypes up to four times.

Post-genescan genotype error estimation was done using the program MICROCHECKER© version 2.2.3 (Van Oosterhout et al., 2004), to test for microsatellite null alleles and scoring errors related to stuttering, large allele dropout and false allele amplification. MICROCHECKER© identifies markers that show an excess of homozygotes (null allele amplification), evidence for scoring error due to stuttering and evidence for large allele dropout. The input variables for MICROCHECKER© included the marker types, of which five were tetranucleotide markers, five dinucleotide markers and two compound markers (Table 1). The MICROCHECKER© analysis was based on a 95% Confidence Interval (CI) for 10 000 permutations per analysis. The probabilities for the observed number of homozygotes per allele are calculated using a cumulative binomial distribution (Weir, 1996) and the p-values of all homozygote size classes were calculated by ranking these observed values to randomized genotypes (expected values). These significance values are then combined to calculate Fisher's probability test to identify genotype errors. All data were first subjected to the pre-genescan correctional method, thus providing a high quality datasheet for the degraded DNA samples as a starting datasheet for post-genescan genotype error corrections.

The relationship between the amount of missing data per locus and the amplicon length (using the largest allele size (in base-pairs) as the amplicon size) was investigated by calculating the correlation coefficient and slope of linear regression. I determined if dinucleotide amplicon length differed significantly from tetranuclotide amplicon length, and compared the average

missing data for simple[10] microsatellite markers (dinucleotide markers (n=5)) and complex microsatellite markers (tetranucleotide (n=5) and compound markers (n=2)) by doing two unpaired two-tailed t-tests. All statistical analyses were done using GraphPad Prism version 3.00 for Windows (GraphPad Software, San Diego California, USA).

RESULTS

Eighty-eight of the 120 extracted DNA samples (73.3%) were successfully sequenced and amplified for a 509 base pairs (bp) sequence. By comparing the amplified sequences to previously published sequences in the genetic sequence database GenBank the 13 resulting haplotype sequences (Genbank accession numbers KC179715-KC179727) were verified as African elephant DNA sequences belonging to the genus *Loxodonta.*

One-hundred of the 120 extracted DNA samples (83.3%) were successfully genotyped for 12 microsatellite loci. The average percentage of missing data per locus was 7.6% and values ranged between 1% and 22%.

The amount of missing data (%) increased significantly with marker amplicon length ($F=9.33$; $r^2=0.483$; $P<0.05$), suggesting that 48.3% of the variability in missing data could be explained by differences in the marker amplicon length (Figure 3). I found no significant difference ($p>0.5$) between the length of dinucleotide (mean = 171.4; SD=59.94; SE = 26.81) and tetranucleotide (mean = 232.6; 64.47; 28.83) amplicons. Simple microsatellite markers (mean =2.6, SD= 1.52; SE=0.68) had significantly less missing data ($p<0.05$) than complex markers (mean =11.29; SD=6.13; SE=2.32) (Figure 4).

[10]In the context of this chapter, I use the word "simple" to refer to the less complicated repeat motif of dinucleotide markers as compared to more complex repeat motifs of tetranucleotide or compound microsatellite markers.

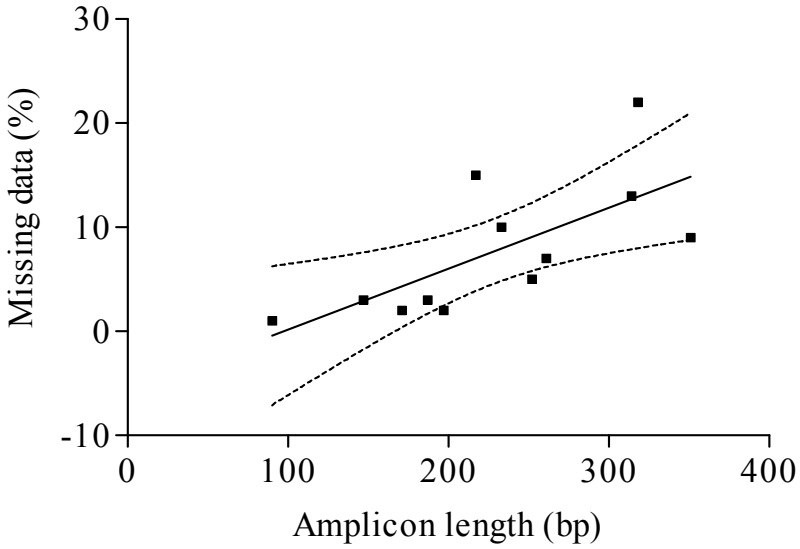


Figure 3. Based on a linear regression, the amount of missing data (%) per locus is significantly positively related ($F=9.333$; $r2 = 0.4827$; $P<0.05$) to the amplicon length of each of the 12 microsatellite markers (base pair size of the largest allele per locus).The regression line (least squares line) is indicated by the solid black line (Slope = 0.059±0.02) with the 95% confidence indicated by the dotted lines

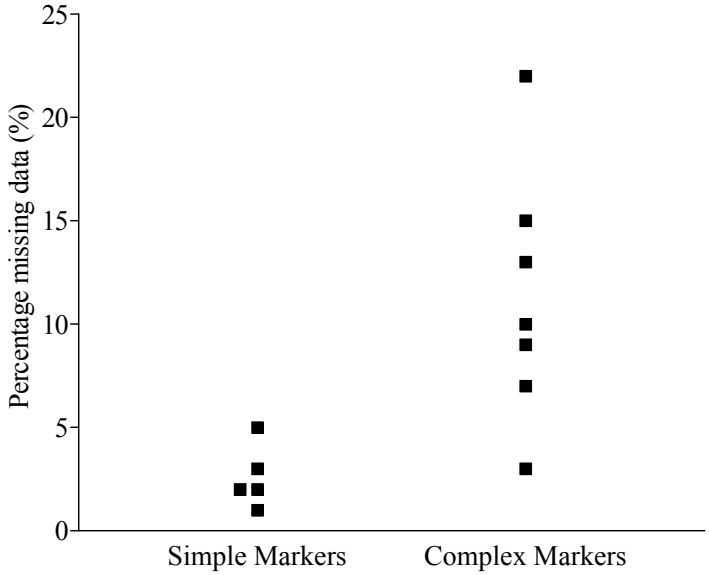


Figure 4. Simple microsatellite markers (dinucleotide markers) (mean =2.6, SD= 1.52; SE=0.68) had significantly less missing data ($p<0.05$) than complex microsatellite markers (tetranucleotide and compound markers) (mean =11.29; SD=6.13; SE=2.32), when comparing (unpaired t-test) genotype results of 100 individuals for five dinucleotide markers, five tetranucleotide markers and two compound markers.

Based on genotype error estimations, five of the alleles showed evidence of an excess of homozygotes (Table 2), but no locus showed any evidence of scoring error due to stuttering or large allele dropout. Only two of the loci (Locus 4 and Locus 10) showed an excess of homozygotes at more than one size class (allele).

Table 2. The total observed and expected homozygotes (as a percentage of the alleles) calculated using a cumulative binomial distribution (Weir 1996) (observed) and randomised genotypes (expected) for five microsatellite loci that showed possible null allele amplification through a significant excess of homozygotes (were Fisher's combined probability < 0.05).

Locus	Expected homozygotes (% of alleles)	Observed homozygotes (% of alleles)	Fisher's combined probability
Locus 2	28.32	41	<0.01
Locus 4	15.41	30	<0.01
Locus 8	20.8	34	<0.01
Locus 10	30.40	45	<0.01
Locus 12	17.62	29	<0.01

DISCUSSION

The ability to extract mitochondrial specific DNA from faecal samples may vary greatly, with successful extractions occurring as many as 92% of the time when extracting mtDNA with an long amplified sequence length (600-700bp fragment) (Frantzen et al., 1998) to only 8.3% of the time (one of 12 samples) when amplifying a shorter mtDNA strand (398bp) (Kohn et al., 1995). However, Frantzen et al. (1998) found that about 31% of faecal samples do not yield any DNA, even after repeated DNA extraction attempts.

The successful extraction of nuclear DNA from faecal DNA also varies greatly. For instance, in one study the amplification success rates for nuclear DNA ranged from 26% to 88% (Murphy et al., 2002). Successful DNA extraction depends on the sample quality (Taberlet et al., 1999),

preservation and storage methods (Murphy et al., 2002; Nsubuga et al., 2004) and extraction procedure (Piggot and Taylor, 2003; Wehausen et al., 2004). Congruent with studies on Asian elephants (*Elephas maximus*) (Fernando and Lande, 2000; Fernando et al 2000; Fernando et al., 2003), it is apparent from the present amplification success rates of 73.3% for mtDNA and 83.3% for nDNA, respectively, that elephant faeces can effectively be used as a DNA source. By combining an extremely simple but precise collection and preservation method, with a modified extraction protocol, I was able to effectively use faecal DNA as a DNA source.

The variability in amplification success rate between mtDNA and nDNA may be due to the more effective amplification of short strands of degraded DNA (Taberlet et al., 1999) by the microsatellite markers (the largest allele lengths for markers range from 90bp to 351bp), as compared to the longer amplicon needed for the mtDNA (509bp). Even within nDNA amplification alone, the effect of amplicon length could clearly be seen. I found a significant positive association between marker amplicon length and the amount of missing data, where the longer the marker amplicon the lower the amplification success (Figure 3). My results support those of Ishida et al. (2012) where they redesigned previous microsatellite primers to produce shorter amplicons, and found that the redesigned primers were significantly more effective at amplifying faecal DNA than the original primers.

Repeat motif may also have a considerable influence on the percentage of missing data. I found that simple markers (dinucleotide markers) had significantly less missing data than complex markers (tetranucleotide and compound markers) (Figure 4), suggesting that complex markers may have more difficulty amplifying when only short fragmented strands of degraded DNA are available. There was no significant difference between amplicon lengths of the dinucleotide and tetranucleotide markers, suggesting that the significant difference in amplification success

between simple and complex markers cannot solely be attributed to size differences between simple and complex markers, but rather supports the likelihood that the repeat motif also influences the amplification success (Taberlet et al., 1996). To maximise amplification success for severely degraded DNA such as that found in faecal samples, it may thus be necessary to use relatively simple microsatellite markers containing alleles with small amplicon sizes.

Another factor that should influence marker choice is the presence of genotype errors. Genotype errors often go undetected because they are unobtrusive (Bonin et al., 2004), and although the influence of genotyping errors are especially detrimental to finer-scale individual identification studies (Taberlet and Luikart, 1999) their influence on population structure studies is not negligible (Falush et al., 2007). Programs like MICROCHECKER© determine if there are possible genotype errors in a dataset through calculating if there is an excess of homozygotes, but the possibility remains that the excess homozygotes may actually be present in a population due to factors like inbreeding (Allendorf et al., 2013 – section 6.3 pg 101), the Wahlund effect (Allendorf et al., 2013 – section 9.1.1 pg 95), or sex-linkage (Allendorf et al., 2013 – section 5.5 pg 90-91). When calculating the probability of having genotype errors in a dataset, MICROCHECKER considers the possibility of all loci showing an excess of homozygotes, and if so, the program will indicate that the population might not be in panmixia, meaning that the excess homozygotes may actually represent the genetic characteristics of the population (Van Oosterhout et al., 2004). My study population included five loci that showed possible null allele amplification by having an excess of homozygotes, and thus not conforming to Hardy-Weinberg Equilibrium, but no locus showed evidence of scoring error due to stuttering or large allele dropout. Only two of these loci (Locus 4 and Locus 10) showed excess of homozygotes at more than one size class and these two loci were also the microsatellite markers where differences

between the total expected homozygotes, and the total observed homozygotes were the greatest. I therefore suggest that these two markers should be excluded from subsequent population structure analysis as it is highly likely that the high presence of excessive homozygote amplification at more than one size class (allele) may indeed indicate that these markers are prone to null allele amplification, especially when compared to the other microsatellite markers of which only three show excessive homozygote amplification at a single size class (allele). I was thus able to use 10 of the 12 proposed microsatellite markers in subsequent population structure analysis, indicating that faecal DNA may be an efficient DNA source to use for nDNA microsatellite genotyping.

Although faecal samples provide an effective way of acquiring non-invasive DNA samples, there seems be a trade off between the ease of collecting samples, the time and money spent on extracting the DNA and the accuracy of the genotypes obtained from the degraded DNA. Nevertheless, by focussing efforts on appropriate sample collection and storage, and by following the stringent extraction and amplification protocols, the accuracy of the genotypes obtained may be of such quality as to correctly and effectively be used in population structure analysis. In summary, faecal samples do not only provide an effective way to obtain both mtDNA and nDNA, but also an efficient way to genotype multiple loci if genotype error eliminating protocols are implemented.

DATA ACCESSIBILITY

DNA sequences: Genbank accession numbers KC179715-KC179727

CHAPTER 4
THE GENETIC STRUCTURE OF THE KAZA-TFCA ELEPHANT POPULATION

INTRODUCTION

The genetic structure of elephant populations have been studied at a continental scale (Comstock et al., 2002; Eggert et al., 2002; Nyakaana et al., 2002; Ishida et al., 2012), within countries (Nyakaana and Arctander, 1999; Munshi-South, 2011), and within defined protected areas (Archie et al., 2006; Wittemyer et al., 2013). Continental-scale studies that included samples from several countries have been based on different molecular genetic techniques (Comstock et al., 2002; Eggert et al., 2002; Nyakaana et al., 2002; Ishida et al., 2012). Ishida et al. (2012), for instance, used 653 samples from 22 locations in 13 different countries. These studies improved our understanding of the structuring of populations and the findings are also of relevance in the combat of the illegal ivory trade (Wasser et al., 2004; Wasser et al., 2007; Wasser et al., 2008; Ishida et al., 2012). Country specific genetic population structure studies mainly focussed on the genetic relatedness of elephants within and between protected areas (Nyakaana and Arctander, 1999; Chiyo et al., 2011; Munshi-South, 2011). Some of these studies detected historical linkages between populations and identified possible geographic linkages between populations (Nyakaana and Arctander, 1999; Eggert et al., 2007; Okello et al., 2008b; Epps et al., 2013), estimated effective population sizes (Eggert et al., 2003), and studied the interactions between relatedness, social organisation and behaviour of elephants in national parks (Gobush et al., 2009; Chiyo et al., 2011; Munshi-South, 2011; Archie and Chiyo, 2012)[11].

[11] See Chapter 1 for more details on previous studies that investigated genetic structuring in elephant populations.

Earlier studies on the genetic characteristics of populations lend support to the implied behavioural, spatial and demographic structuring of populations (McRae and Beier, 2007; Guillot et al., 2008; Chiyo et al., 2011; Ahlering et al., 2012; Archie and Chiyo, 2012) and have been used to explain associations within populations (Archie et al., 2008). For instance, female elephants live in "core social groups", with fissions and fusions happening between these groups (Archie et al., 2006; Archie et al., 2008; Fishlock and Lee, 2012), where associations within and between core social groups is a function of genetic relatedness (Archie et al., 2006; Moss and Lee, 2011). While female elephants mostly remain with their natal social groups, males disperse during puberty (Evans and Harris, 2008). Female core social groups restrict their roaming to smaller areas than males (Hall-Martin, 1987; Jackson and Erasmus, 2005), as is also the situation in my study region (see Chapter 2). These gender differences in roaming may possibly influence gene flow across a landscape, where the primary source of gene flow across geographic areas is expected to be predominantly male-mediated (Nyakaana and Arctander, 1999; Archie et al., 2008). In agreement with the work of others (Archie et al., 2008; Okello et al., 2008b) I therefore expect to find a more defined genetic structure among female-based sub-populations compared to male-based sub-populations.

Most past studies on the genetic structuring of elephant populations were conducted on fragmented populations[12] (Nyakaana and Arctander, 1999; Okello et al., 2008b; Munshi-South, 2011). This detracts from inferences where natural factors instead of anthropogenic factors may influence genetic structure. Examples of natural factors include landscape forces such as mountains ranges (Rueness et al., 2003), rivers (Eriksson et al., 2004; Anthony et al., 2007), geographic distances (Anthony et al., 2007), and historical phylogeography and glacial refugia

[12] Defined in Glossary

(Nyakaana et al., 2002; Schmitt et al., 2005; Anthony et al., 2007). Natural factors that influence the distribution of elephants, and consequently also the genetic structure of a population, include resource availability (Chamaillé-James et al., 2008) and vegetation heterogeneity (Young et al., 2009) and density dependent dispersal (Young and van Aarde, 2010). Anthropogenic induced factors, on the other hand, include landscape transformation (Hoare and du Toit, 1999; Osborn and Parker, 2003; Graham et al., 2009), obstruction of dispersal opportunities (e.g. Loarie et al., 2009), and extirpation of local populations by hunters (Whitehouse and Harley, 2001).

The elephant population in my study area is one of the few that inhabits a landscape that has not been fragmented. Here anthropogenic disturbances may have altered local movement (Chase and Griffin, 2009) but apparently do not limit the international movement of elephants (Roever et al., 2013b). Elephants in this area may form part of a larger population that spans the five countries included in the KAZA-TFCA (Chase and Griffin, 2006), an area some 60 times greater than the size of the average annual elephant home range in the area. Elephants in this region have not been exposed to management interferences such as culling nor have they been subjected to intensive poaching like elephant populations elsewhere (Burn et al., 2010)[13]. The elephant population of the KAZA-TFCA therefore may be considered a naturally structured population.

In this chapter I aimed to establish if this population is genetically structured. I compared mitochondrial DNA (mtDNA – female inherited) to nuclear DNA microsatellite markers (nDNA – male and female inherited) of 120 samples collected in the KAZA-TFCA (Chapter 2). I used mtDNA haplotype analyses, an Analysis of Molecular Variance (AMOVA) and Bayesian

[13] However, it should be noted that elephants in the periphery of our study area (e.g. Kafue National Park) have been subjected to intensive poaching, as is evident in the largest ivory seizure on record in 2002 (Wasser et al., 2007). This may also hold for elephants in north-western Zimbabwe (personal communication, Prof. Sam Wasser).

Clustering Algorithms (BCAs) to determine the genetic structure of elephant populations in this region and inferred the influence that male and female mediated gene flow have on this structure.

METHODS

Data

The sampling, DNA extraction and amplification procedures have been described in Chapter 3. The mitochondrial DNA (mtDNA) database consisted of 88 mtDNA D-loop control region sequences (primers LafCR1 and LafCR2; see Nyakaana and Arctander, 1999) containing 13 haplotype sequences, with 25 polymorphic sites across sequences (Table 3). The nuclear DNA (nDNA) database consisted of 100 individuals that were genotyped at 10 different microsatellite loci (Table 1, excluding Locus 4 and Locus 10 (Chapter 3)).

Table 3. Haplotype names, definitions, relative frequency of occurrence and GenBank accession numbers for 13 mtDNA D-loop haplotypes found in 88 individuals.

Haplotype list	Haplotype definition	Relative frequency of occurrence	GenBank accession numbers
Haplotype 1	AATCCCAGGGTGCTATGAAGAAAAC	0.0227	KC179715
Haplotype 2	AAACCCAGGGTGCTACGAAGAAAAC	0.0114	KC179716
Haplotype 3	AATCCCAGGGTGCTACGAAGAAAAC	0.0114	KC179717
Haplotype 4	GATCCCGGGGTGCTATGAGAATAGT	0.0114	KC179718
Haplotype 5	GATCCTGAAATGTTATGAGAGAAGT	0.648	KC179719
Haplotype 6	GGTTTTGAAATGTTATGAGAGAAGT	0.0114	KC179720
Haplotype 7	GATCCTGGAATGTTATGGGAGAAGT	0.159	KC179721
Haplotype 8	GATCCTGGAATGTTATGAGAGAAGT	0.0114	KC179722
Haplotype 9	GATCCTGGGATGTTACGAGAGAAGT	0.0455	KC179723
Haplotype 10	GATCCTGGGATGTTATGAGAGAAGT	0.0114	KC179724
Haplotype 11	GATCCTAGAATGTCGCGAGAGAGGT	0.0227	KC179725
Haplotype 12	GATCCTGGAATGTTGCGAGAGAAGT	0.0227	KC179726
Haplotype 13	GATCCTAGAACATTGTAAAGGAAAT	0.0114	KC179727

Analyses

Haplotype analysis (mtDNA)

A detailed justification and description of the approach summarised below is provided in Appendix VIII. The 88 sequences were confirmed as African elephant mtDNA sequences belonging to the genus *Loxodonta* by comparison with published sequences in the genetic sequence database GenBank (National Centre for Biotechnology Information). Sequences were then individually viewed and processed (CLC Bio Genomics workbench; CLC Bio, Cambridge, MA), and aligned (Webprank; Goldman Group Software, European Bioinformatics Intitute © 2011). The aligned sequences were then reformatted into a FASTA file (using Clustal X 2.0; Larkin et al., 2007). The FASTA file was converted into a haplotype data file and a Roehl data file (using DnaSP 5.10.01; Librado and Rozas, 2009). The FASTA file was used to draw a midpoint-rooted Neighbour Joining tree using the best-fit substitution model which was calculated by testing different substitution models of evolutionary change through Bayesian Information Criterion (BIC) and Akaike Information Criterion (AICc) comparisons (MEGA 5; Tamura et al., 2011). The haplotype data file was used to calculate relative haplotypic frequencies and test for neutrality through Tajima's test and Fu's Fs test of selective neutrality in Arlequin (Excoffier and Lischer, 2010), which is necessary to determine if the data can reliably be used in the ensuing population structure analysis. The Roehl data file was used to construct a Median-Joining network in Network 4.6.1.0 (Fluxux Technology Limited). I also mapped the geographic distribution of the haplotypes using the program ArcMap (©ESRI 2011, ArcGIS Desktop: Release 10, Redlands, CA: Environmental Systems Research Institute).

Population structure analysis (mtDNA and n DNA)

I used Geneland (Guillot et al., 2008) to determine how many sub-populations (K) of genetically similar individuals were present in my study area. Geneland refers to the sub-populations (K) as "clusters" of genetically similar individuals. Because Geneland uses geo-referenced multi-locus data as input information, I converted the sequence data to haploid data based on binary code, where all the polymorphic sites (n=25) were considered as individual loci, with the possibility of 4 alleles (ACTG) at each of the loci.

I used an Analysis of Molecular Variance (AMOVA) to calculate inter-population genetic variation (F_{st} – Fixation Index) for both nDNA and mtDNA, and also corrected for microsatellite allele length (R_{st}[14]) (Allendorf et al., 2013). The Fixation Index ranges from zero to one, where zero would mean that all sub-populations in the study region have equal allele frequencies, indicating that there is only one population and that this population is in panmixia (e.g. Allendorf et al., 2013). A Fixation Index of one means that all the sub-populations in the study area are fixed for different alleles, thus indicating that all the sub-populations are genetically differentiated from each other. The AMOVA for the mtDNA was done in Arlequin using the Kamura-2-Parameter (Excoffier and Lischer, 2010), while the AMOVA for the nDNA was done using GenAlEx (Peakall and Smouse, 2006).

I used a Mantel test (using the ade4 package (Chessel et al., 2004) for the free-ware statistical program R 2.13.1 (© 2011 The R Foundation for Statistical Computing) to test for Isolation By Distance (IBD) to establish if geographic distance (in the form of a geographic distance matrix) could explain genetic differentiation (in the form of a pairwise genetic distance matrix calculated

[14] Rst is an analog of Fst, but incorporates a step-wise mutational model to account for different microsatellite allele lengths.

in Arlequin) between sub-populations[15] (see Appendix IX for the distance matrices). Geographic distances were the distances between the centroid locations of each of the sub-populations (see Appendix X for the R-Scripts implemented in the mantel test and Appendix VIII for a full description of the centroid calculation and distance measurement done in ArcMap (©ESRI 2011, ArcGIS Desktop: Release 10, Redlands, CA: Environmental Systems Research Institute) and the Geospatial Modelling Environment –0.7.1.0 (© Hawthorne L. Beyer 2009-2012)).

For mtDNA and nDNA I projected the pairwise genetic distances between individuals on a geographic landscape using the "Single Species Genetic Divergence" tool in the Genetic Landscapes GIS Toolbox (Vandergast et al., 2011). This tool allows for the identification of genetic discontinuities across a landscape, and identifies areas in the landscape were gene flow is obstructed. The resulting genetic divergence landscapes were viewed in ArcMap 10 (©ESRI 2011, ArcGIS Desktop: Release 10, Redlands, CA: Environmental Systems Research Institute), where areas with genetic continuity (high genetic similarity between adjoining samples) were indicated in blue, while areas with genetic discontinuities (high genetic dissimilarity between adjoining samples) were indicated in red in figures 10 a and b.

RESULTS

Haplotype Results (mtDNA)

I found 13 haplotypes (Table 3), seven of which were unique (occurred only once), and one haplotype that featured in 65% of samples. Both the Tajima's test and the Fu's test indicated that the mtDNA marker used in this study is a selectively neutral marker (Tajima's D = -1.4, p = 0.054; Fu's Fs = -1.8, p = 0.28). The negative Fs value (Fu's test) suggests an excess of alleles

[15]Based on the Geneland analysis the microsatellite nDNA database partitioned into only two distinct populations, making it impossible to do a Mantel Test correlation for the nDNA populations. Therefore, I only tested mtDNA for IBD.

indicative of a recent population expansion, however, this was not significant ($p>0.2$). Most haplotypes were shared across the entire sampling region (for example, Haplotype 5 (red), Figure 5), with only a few haplotypes occurring in distinct locations (for example, Haplotype 1, 2 and 3 were closely grouped within Zimbabwe (Figure 5, haplotypes in shades of purple and pink).

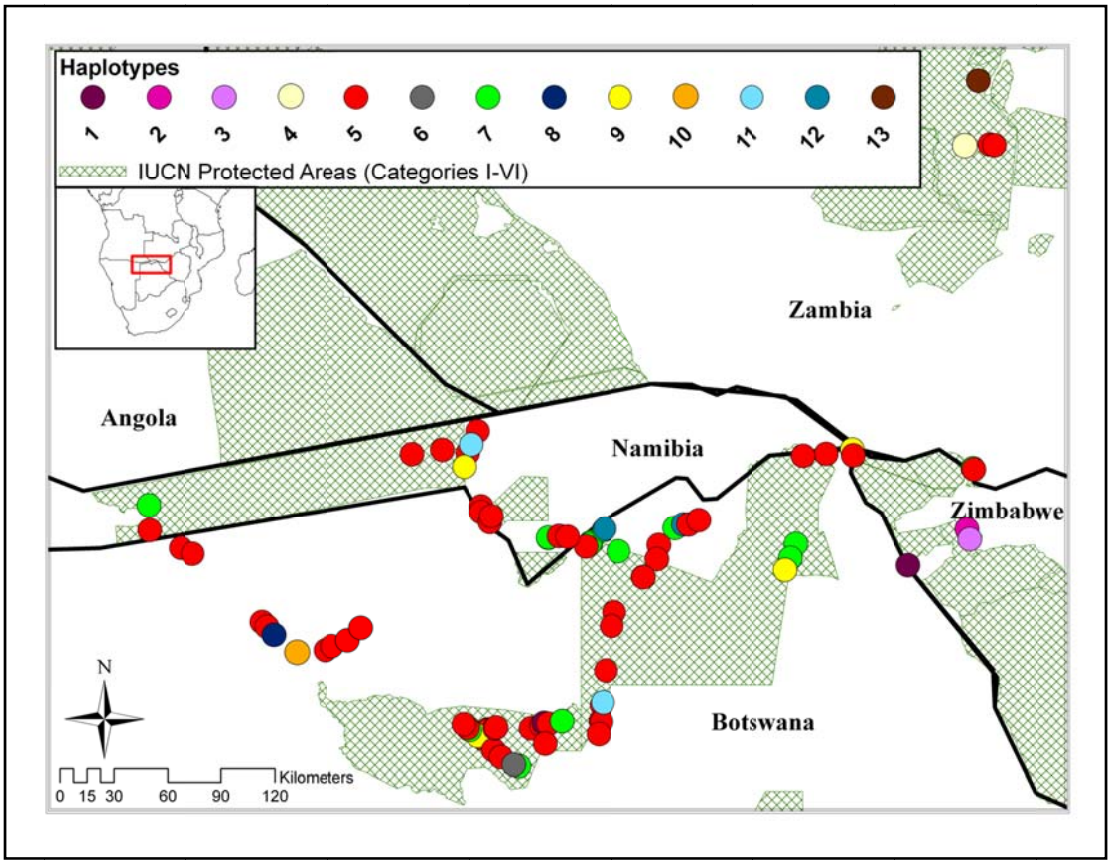


Figure 5. A geographical representation of the individual-based haplotypes found in the KAZA region. Colours annotate different haplotypes and green cross-hatched areas show the IUCN Category I-VI Protected Areas[16].

The Median-Joining Network (Figure 6) supported the geographic distribution of haplotypes. Most haplotypes were grouped together. Haplotypes 1, 2 and 3, however, formed a distinct group, as did Haplotype 13 (which was located at the extreme northern range of the sampling region in Zambia). The three haplotypes (1, 2 and 3) differ from each other by only one base-pair

[16] IUCN categories are defined in Appendix II

(one mutational event) while all other haplotypes differ from these three haplotypes by at least six base-pairs mutations. Haplotype 13 differed from all other haplotypes by at least 7 base-pair mutations.

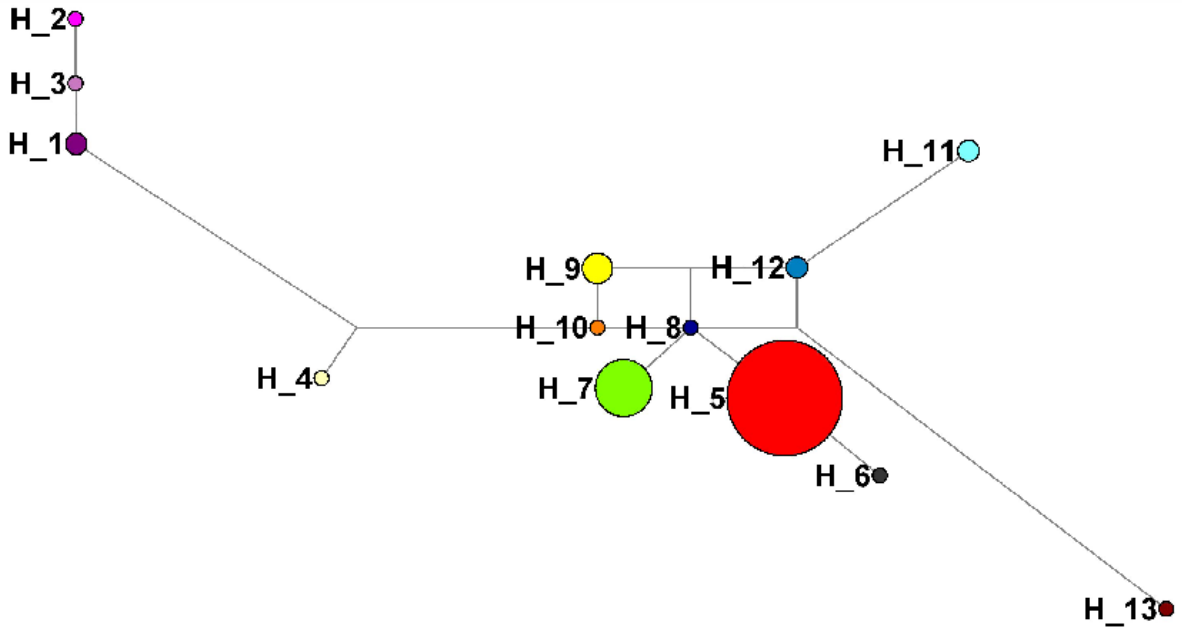


Figure 6. The Median-Joining Network showed evidence of separation between two groups of haplotypes, where most haplotypes were grouped together, with the exception of three haplotypes (Haplotype 1 (dark purple), Haplotype 2 (light purple) and Haplotype 3 (pink)) that were grouped separately. The haplotype symbol colours correspond to the assigned haplotype colours in Figure 1. The size of the circles corresponds to the number of individuals with that haplotype, and the length of the branches approximates the number of mutational events that separate the haplotypes.

The Neighbour-Joining Tree (Figure 7) supports both the geographic distribution of the haplotypes and the Median-Joining Network, with high support (99% bootstrap support) for the clustering of Haplotypes 1, 2, and 3 as a distinct clade, but little support for inter-haplotypic structuring throughout the rest of the proposed haplotype clades.

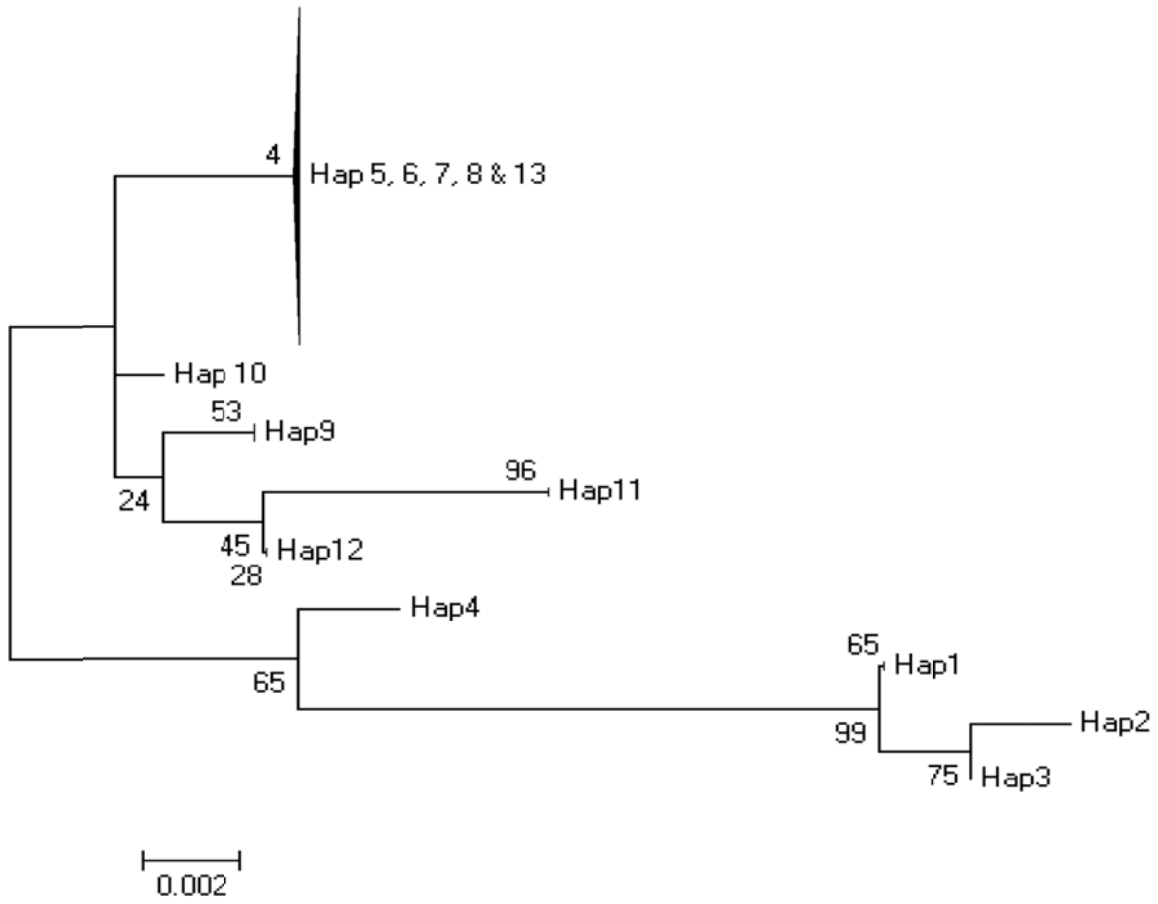


Figure 7. The midpoint-rooted Neighbour-Joining tree showed high support (99%) for the clustering of three haplotypes into a separated clade, with little support for inter-haplotypic structuring throughout the rest of the proposed haplotype clades. Numbers indicate a 1000 bootstrap repeat support for the indicated haplotypic clades.

Population Structure Results (mtDNA and nDNA)

After being subjected to the Geneland Bayesian Clustering Algorithms (BCAs) the mtDNA data partitioned into three genetically distinct sub-populations (Figures 8a, b and c), while the microsatellite nDNA data suggested two genetically distinct sub-populations (Figures 9a and b). The first sub-population (mtDNA Sub-population 1, Figure 8a) stretched across the study area, the second (mtDNA Sub-population 2, Figure 8b) suggested linkage between samples collected in the extreme northern limits of the study area and mtDNA Sub-population 1, while the third

(mtDNA Sub-population 3, Figure 8c), suggested linkage between samples collected in the extreme eastern limits of my study area and mtDNA Sub-population 1. However, Geneland did not indicate any linkage between mtDNA Sub-population 2 and 3.

Maps of the posterior probability to belong to a specific sub-population for the nDNA data suggested that there were only two distinct sub-populations, with nDNA Sub-population 1 (Figure 9a) extending across the entire study area while nDNA Sub-population 2 (Figure 9b) was limited to the extreme northern parts of the study area.

The non-significant correlation between the pairwise genetic distance and geographic distance (Appendix IX) amongst the mtDNA sub-populations ($r=-0.76$; $p>0.5$) suggested that IBD did not explain the genetic distance between sub-populations. The Analysis of Molecular Variance (AMOVA) suggested that most of the mtDNA based genetic variation could be ascribed to variation among sub-populations ($F_{st}=0.787$, Table 4), while the nDNA AMOVA ascribed such variation to differences between individuals within sub-populations ($F_{st}=-0.02$; $R_{st}=-0.045$, Table 5). The mtDNA sub-populations differed significantly, while nDNA sub-populations did not differ significantly for both measures of variation (Fst and Rst).

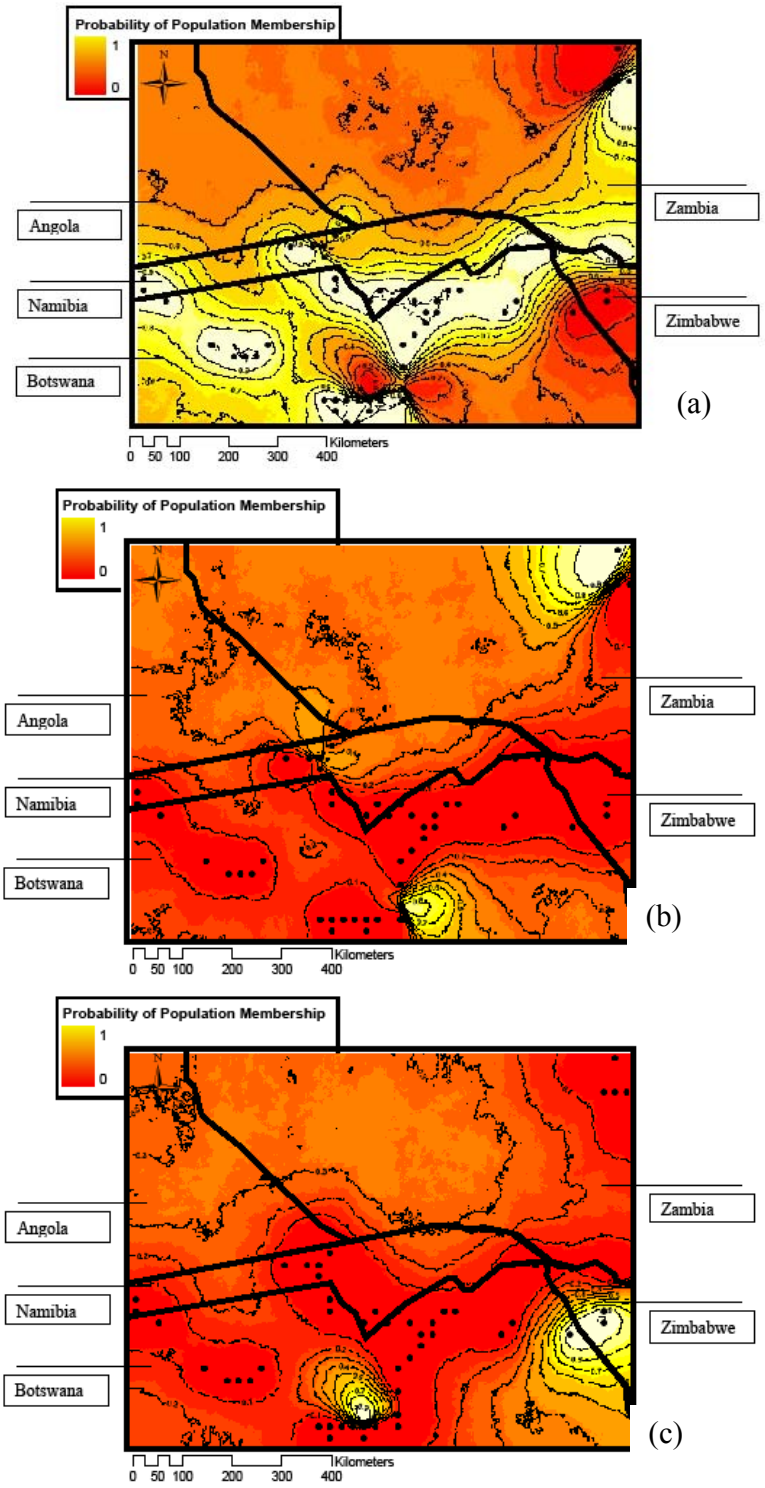


Figure 8a, b and c. Maps of the posterior probability to belong to Sub-population 1, 2 and 3 as calculated in Geneland using the mtDNA data. The black dots indicate the sampling locations (n=88).

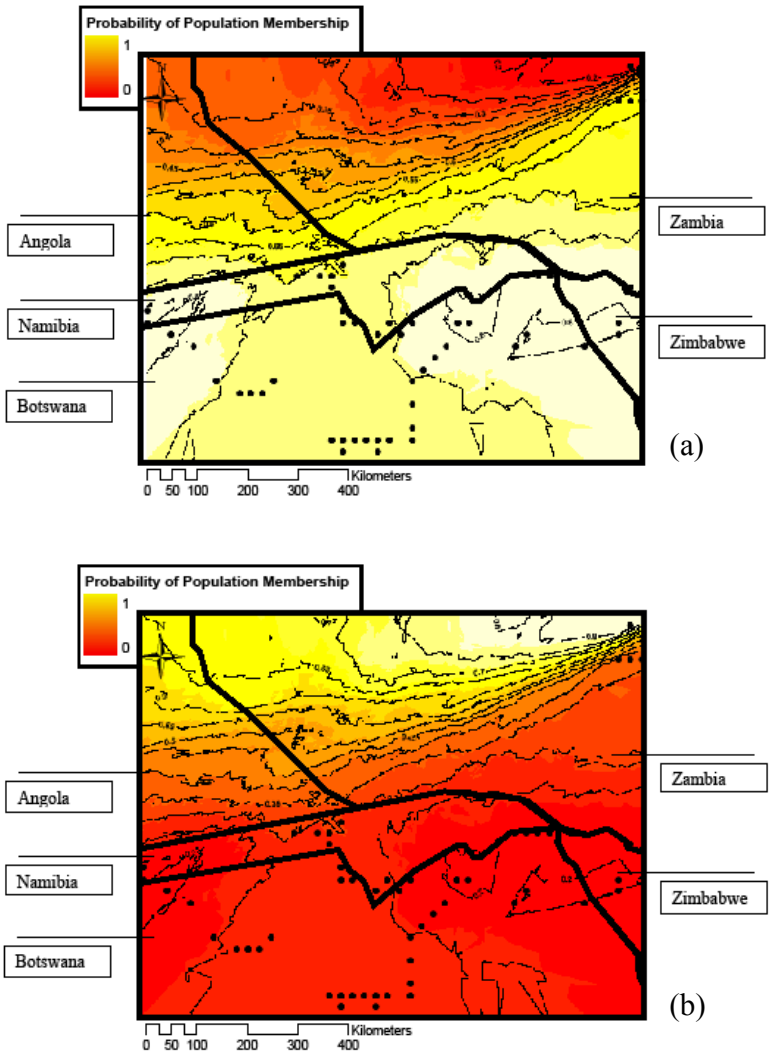


Figure 9a and b. Maps of the posterior probability to belong to Sub-population 1 and 2 as calculated in Geneland using the nDNA microsatellite data. The black dots indicate the sampling locations (n=100).

Table 4. An Analysis of Molecular Variation (AMOVA) indicated that most of the genetic variation found in the mtDNA data could be attributed to significant inter-population genetic differences ($p<0.01$).

Source of Variation	**d.f.**	**Sum of Squares**	**Variance components**	**Variation (%)**
Among Sub-populations	2	39.2	3.29	78.73
Within Sub-populations	85	75.48	0.89	21.27
Total	87	114.68	4.17	100
Fixations Index:	F_{st}	0.787		

Table 5. An Analysis of Molecular Variation (AMOVA) indicated that most of the genetic variation found in the nDNA data could be attributed to the genetic variation found within individuals, where the sub-populations do not show significant inter-population differentiation (Fst) ($p>0.5$), even when correcting for the Step-Wise Mutation Model (Rst).

Source of Variation	d.f.	Sum of Squares	Variance components	Variation (%)
Among Sub-populations	1	3.15	0.000	0
Within Sub-populations	98	417.71	0.566	15
Within Individuals	100	313	3.130	85
Total	199	733.87	3.696	100
Fixations Index:	F_{st}	-0.02		
	R_{st}	-0.045		

The mtDNA divergence map (Figure 10a) identified areas with high genetic discontinuities (indicated in red) and separated both the northern (mtDNA Sub-population 2) and the eastern sub-population (mtDNA Sub-population 3) from each other and from the central sub-population (mtDNA Sub-population 1). Conversely, the microsatellite nDNA divergence map (Figure 10b) showed high genetic continuity across the entire landscape, indicating that gene flow was high across the total study region, with no discernable area of high genetic discontinuity.

DISCUSSION

Although most mtDNA haplotypes in the KAZA-TFCA study region overlapped geographically, the Median-Joining Network, Neighbour-Joining Tree and AMOVA indicated that there were genetically and geographically distinct groups of mtDNA haplotypes. These haplotype groupings included one Zambian and three Zimbabwean haplotypes that were distinct from the haplotypes in Botswana. These distinct mtDNA haplotype groups may be the result of strong founder effects, which is supported by the lack of Isolation By Distance (IBD) and the absence of sustained historical or present-day genetic barriers in this region. Such founder-based population structuring differs from the genetic structuring described for elephant populations in Tanzania where structuring is ascribed to anthropogenic and landscape factors (Epps et al., 2013). The

distinct mtDNA haplotype groups may also be a result of persistent poaching in specific hotspots, which can subsequently lead to the extirpation of specific matrilineal lines, or the deterring of female movement between adjacent sub-populations.

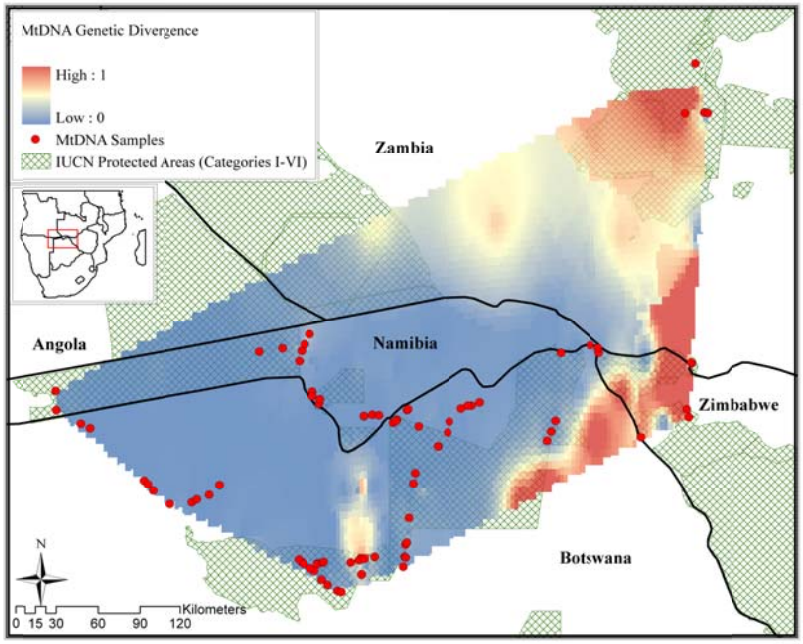


Figure 10a. A genetic divergence map based on the mtDNA data showed definite genetic discontinuities across the landscape (indicated in red), with samples in the north and east of my study area being separated from each other, and from the rest of the study region. The red dots indicate the sampling locations (n=88).

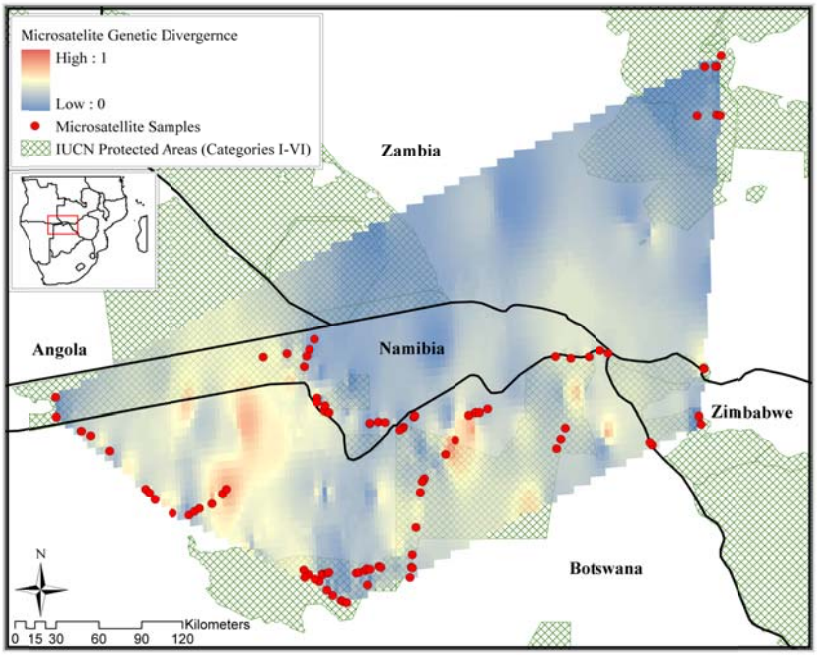


Figure 10b. A genetic divergence map based on the nDNA data showed little evidence for any genetic discontinuities, with high geneflow and genetic similarity (indicated in blue) across the landscape. The red dots indicate the sampling locations (n=88).

As expected, and in agreement with the studies of Archie et al. (2008) and Okello et al. (2008b) in Kenya and Nyakaana and Arctander (1999) in Uganda, I found a more defined structure for mtDNA- than nDNA-based sub-populations. The mtDNA sub-population structure analysis produced three genetically and geographically distinct sub-populations that showed significant genetic differentiation between sub-populations, while the nDNA sub-population structure analysis did not support significant population differentiation. Mitochondrial DNA Sub-population 1 spanned across the study area and contained 93% of the samples, but mtDNA Sub-population 2 and 3 provided support for genetic structuring, where these two sub-populations were linked with Mitochondrial DNA Sub-population 1, but not with each other. Conversely, the genetic differentiation between sub-populations as implied by the nDNA data was not significant. This lack of genetic population structuring suggests sustained gene flow across the

entire study region. The wide spanning mtDNA Sub-population 1 and the high nuclear gene flow also support trans-national elephant movements across the KAZA-TFCA (see Chapter 2 and references therein). These differences between mtDNA and nDNA sub-population structuring are also apparent in the genetic discontinuities across the landscape. The mtDNA sub-populations are clearly separated by areas of high genetic discontinuity, with high gene flow within sub-populations but obstructed gene flow between sub-populations, whereas there are no identifiable areas that appear to obstruct gene flow for nDNA.

Despite concerted efforts, no elephant movements have yet been recorded between the Chobe NP (Botswana) and the Kafue NP (Zambia) (see Chapter 2) (Jackson and Erasmus, 2005; Loarie et al., 2009), an observation that supports the distinct mtDNA sub-populations, but contradicts the high nuclear gene flow implied by the present study. By modelling resource selection functions for males and females, Roever et al. (2013a) identified potentially suitable habitat along multiple routes for elephant movements between the Chobe NP and Kafue NP. The results of my study, combined with these previous studies (Jackson and Erasmus, 2005; Loarie et al., 2009; Roever et al., 2013a), suggests that female elephants in Chobe NP and Kafue NP restrict their movements, while male elephants might roam between these parks.

The differences between mtDNA and nDNA genetic sub-population structures may be explained by differences in male and female social behaviour and spatial utilisation patterns. Elephant social behaviour and spatial use are not mutually exclusive (Charif et al., 2005; Wittemyer et al., 2007; Chiyo et al., 2011). The dynamic interaction between social behaviour and spatial use may explain gender differences in the genetic population structures found in the KAZA-TFCA (present study), as was the case in two studies done on elephant populations in Kenya (Archie et al., 2008; Okello et al., 2008b). My results also support the findings of others that ascribed

gender-biased population structuring to the restricted roaming behaviour of core social groups (Archie et al., 2006; Archie et al., 2008; Fishlock and Lee, 2012), and the predominantly male mediated gene flow typical of elephants (Nyakaana and Arctander, 1999; Archie et al., 2008).

CHAPTER 5
SYNTHESIS

Some believe that Africa's elephants were once continuously distributed across the continent (e.g. Barnes, 1999; Georgiadis et al., 1994). However, compared to the historical distribution and gene flow (see Epps et al., 2013), the current distribution of elephants across parts of Africa is patchy (van Aarde and Jackson, 2007) and enforced by landscape transformation (Hoare and du Toit, 1999; Osborn and Parker, 2003; Graham et al., 2009), the obstruction of dispersal opportunities (e.g. Loarie et al., 2009) and the extirpation of local populations by hunters (e.g. Whitehouse and Harley, 2001) and ivory poachers (Okello et al., 2008a). These forces may have reduced local population sizes and in synergy with the establishment of protected areas may have fragmented populations into discontinuous units. The compression of elephants into protected areas may also have contributed to isolation and fragmentation (Laws 1970; Melton, 1985). Several of these protected areas may not be large enough to sustain viable elephant populations (Armbruster and Lande, 1993). These factors have led to a much reduced and fragmented population and a species that is now listed as an Appendix I species by the Convention for the International Trade of Endangered Species (CITES) and as a vulnerable species in the red data list of the International Union for the Conservation of Nature (IUCN).

Population fragmentation has negative genetic consequences (Frankham et al., 2004; Allendorf and Luikart, 2013). Restricted or reduced gene flow between fragmented populations may accelerate inbreeding and reduce the ability of populations to adapt to changes in the environment, thus increasing the risk of extinction (Lande, 1998; Frankham et al., 2004, 2005, 2011; Allendorf and Luikart, 2013). Conservation incentives may counter the effects of

population fragmentation and isolation by restoring or maintaining landscape linkages between populations (e.g. through wildlife crossings (Foster and Humphrey, 1995; Donaldson, 2005; Beben, 2012) and movement corridors (Beier and Noss, 1998; Dixon et al., 2006)). Such landscape linkages may increase inter-population gene flow, which in turn may increase genetic diversity and gene pools of populations.

Transfrontier Conservation Areas (TFCAs) rely on landscape linkages to increase connectivity between isolated protected areas to create larger ecological networks (for instance, the Kavango-Zambesi Transfrontier Conservation Area includes 36 protected areas (van Aarde and Jackson, 2007) that form part of a larger network spanning 278 000km^2 (Metcalfe and Kepe, 2008)). However, the delineation of TFCA's in southern Africa is often driven by historical circumstances and socio-political realities, and therefore strives to create and implement conservation strategies that benefit the resident people (Hanks, 2001). Such delineations could benefit conservation if informed by the spatial and demographic requirements and dynamics of species populations (see van Aarde and Jackson, 2007). These requirements and dynamics may be reflected by the genetic structure of populations (Allendorf et al., 2013).

Establishing if populations are structured genetically despite living in a continuous landscape, such as the population in the KAZA-TFCA at the time of my study, is a case in point of time. Here, like elsewhere in Africa, elephants mostly live in protected areas (Douglas-Hamilton et al., 2005; Lee, 2013), some that are linked (e.g. Chobe National Park and Moremi Game Reserve), and others that are isolated from other formally protected areas, but surrounded by game management areas (e.g. Kafue National Park). At the time of my study elephants roamed extensively across the KAZA-TFCA and their movements on a regional scale were not limited by any barriers (Roever, 2013; Roever et al., 2013a). Locally, spatial structuring and roaming

limitations were driven by anthropogenic barriers (e.g. fences (Mbaiwa and Mbaiwa, 2006; Loarie et al., 2009) and human settlements (Hoare and du Toit, 1999; Roever et al. 2013a)), the distribution of water (de Beer and van Aarde, 2008; Loarie et al., 2009), and functional responses to habitat conditions (Young et al., 2009; Roever et al. 2013b). The spatial structuring of this population may furthermore be the consequence of the social structuring of elephants into family units (breeding herds), their matrilineal societies and female philopatry (described in Wittemyer et al. (200), Archie et al. (2006), Wittemyer and Getz (2007), and Gobush et al. (2009); and summarised in a review chapter by van Aarde et al. (2008)), where most females limit their land use to home ranges smaller than those of males that generally roam across the ranges of several breeding herds (Hall-Martin, 1987; Jackson and Erasmus, 2005).

Populations that are not fragmented by anthropogenic forces may be structured genetically and in response to extrinsic forces (natural landscape barriers of movement and gene flow (Rueness et al., 2003; Eriksson et al 2004)) and intrinsic forces (resource dependencies and social variables that limit movement). Because large-scale elephant movements in the KAZA-TFCA seem not to be hindered by anthropogenic forces, and because there are no historical or existing natural landscape barriers that may obstruct gene flow (extrinsic forces), I postulated that the genetic structure in the KAZA-TFCA elephant population may primarily be in response to intrinsic forces.

It is, however, important to keep in mind that elephant populations on the periphery of my study area have been subjected to intensive poaching, as evidenced by the largest recorded ivory seizure in 2002 (Wasser et al., 2007), where the confiscated ivory originated from elephants poached in the Kafue National Park which lies at the extreme northern part of my study area. Because this study investigated landscape scale genetic patterns, and because other studies have

shown that elephant poaching in the core of this region is less intense than in other African countries (Burns et al., 2010), I expected that the population structure of the elephants in the KAZA-TFCA would not be a function of poaching. I nevertheless take into account that persistent poaching hotspots may play an important role in the local structuring of elephant populations (Okello et al., 2008a). This would be especially evident if, for instance, the poaching hotspots cause the extirpation of entire matrilineal lines, or where poaching may deter female family groups from dispersing. To further investigate this, I would have to specifically look at female genotypes by extending this project to include a sex-determining marker system, which in turn would allow me to identify whether these matrilineal lines are directly affected through targeted poaching events, rather than historical founder effects and gender difference in space use and social behaviour.

In this thesis I aimed to determine i) if elephants in the KAZA-TFCA were structured into genetically distinct sub-populations, and ii) if these sub-populations, if present, had gender specific genetic structures. Earlier studies investigated the genetic structure of spatially fragmented or isolated elephant populations by comparing the genetic characteristics of pre-defined populations and sub-populations (Nyakaana and Arctander, 1999; Okello et al., 2008b; Munshi-South, 2011). Consequently, my study may be one of the few that investigated a naturally structured elephant population across a continuous landscape free of natural and apparent unnatural barriers.

I compared mitochondrial DNA (mtDNA – female inherited) to nuclear DNA microsatellite markers (nDNA – male and female inherited) of 120 faecal samples. In support of other African elephant studies (Archie et al., 2003; Okello et al., 2005; Ishengoma et al., 2008; Munshi-South,

2011), I found elephant faecal DNA to be an effective source for both mtDNA and nDNA ((73.3% and 83.3% successful extraction rate, respectively).

Similar to the findings of earlier studies (Nyakaana and Arctander, 1999; Okello et al., 2008b; Ishida et al., 2012), the results of this study showed that the KAZA-TFCA's elephant population was structured into genetically distinct mtDNA sub-populations. Although the regionally shared mtDNA haplotypes (e.g. Haplotype 5, Figure 5, Chapter 4) suggested some mtDNA gene flow, the significant genetic differentiation between the mtDNA sub-populations suggests that the existing gene flow between sub-populations was obstructed. The results suggests high mtDNA gene flow within each of the mtDNA sub-populations, moderate levels of gene flow between the Botswana sub-population and the Kafue and Zimbabwe sub-populations, but limited gene flow between the Kafue and Zimbabwe sub-populations. Nuclear DNA, on the other hand, showed high gene flow across the entire study region. In agreement with the findings of others (Nyakaana and Arctander, 1999; Nyakaana et al., 2002; Okello et al., 2008b; Ishida et al., 2012), the obstructed mtDNA gene flow compared to the high nuclear gene flow suggests a gender-difference in genetic structuring of the KAZA-TFCA elephant population. Similar to these earlier studies, I suggest that this gender-difference may most likely be explained by differences in the roaming activities of male and females, where most females limit their land use to home ranges smaller than those of males that generally roam across the ranges of several breeding herds, where gene-dispersion will therefore be predominantly male-mediated (Nyakaana and Arctander, 1999; Archie et al., 2008; Okello et al., 2008b).

The KAZA-TFCA elephant population is genetically structured at a regional spatial scale, as apparent country specific sub-populations in Zambia, Zimbabwe and Botswana. Earlier studies ascribe regional genetic variation in elephants to extrinsic (landscape barriers), intrinsic (social

behaviour) and anthropogenic forces (Nyakaana and Arctander, 1999; Okello et al., 2008a, b; Epps et al., 2013). The result of this study suggests that the KAZA-TFCA elephant population is most likely structured in response to intrinsic forces. Such natural, intrinsically driven genetic structuring in the absence of anthropogenic and extrinsic barriers distinguishes this study from others (Eggert et al., 2002; Nyakaana et al., 2002; Epps et al., 2013) in that it may discount man-made or landscape barriers as the sole factors that influence genetic structuring in elephant populations. I therefore suggest that future studies should consider intrinsic factors such as resource dependencies (de Beer and van Aarde, 2008; Loarie et al., 2009; Young et al., 2009) and social behaviour (Wittemyer et al., 2005; Archie et al., 2008; Okello et al., 2008b), in addition to anthropogenic and extrinsic factors, when investigating the genetic structure of elephant populations.

Given the realities of limited resources for conservation, it may be important to identify priority areas for conservation, for instance, by identifying historical or existing functional landscape linkages between populations that may restore or maintain population connectivity. This is especially important for TFCAs that strive to increase connectivity between isolated protected areas to create larger ecological networks. My results provide evidence for such a functional landscape linkage. The high nDNA gene flow in conjunction with the suggested suitable habitat movement routes (based on resource selection modelling) identified by Roever et al. (2013a) suggests a functional linkage between Chobe (Botswana) and Kafue (Zambia) National Parks. Despite concerted efforts, the present study is the first to provide support for inter-population movements between these two areas, where previously recorded spatial telemetry data has failed to illustrate inter-population movements (Loarie et al., 2009). The differentiated mtDNA genetic

structure compared to the high nDNA gene flow suggests that it is mainly bulls that roam between these National Parks.

The results of this study encourage conservation initiatives that strive to counter the effects of population fragmentation and isolation by restoring or maintaining landscape linkages between populations (e.g. through the creation of "megaparks" (van Aarde and Jackson, 2007) and TFCAs (Hanks, 2001, 2003)). Such initiatives may benefit spatially fragmented elephant populations by maintaining or restoring linkages that increase the chance of persistence, enlarge gene pools, increase genetic diversity, and offset the impacts of locally high elephant populations.

These landscape linkages may also ensure the maintenance of inter-population movements, and in so doing maintain the current genetic state of the KAZA-TFCA elephant population. The KAZA-TFCA population has a genetic diversity[17] higher than that found in other southern African populations (Essop, 1996; Whitehouse and Harley, 2001), suggesting that elephants in this region do not suffer inbreeding depression, genetic drift, or other genetically detrimental factors often associated with populations below a genetically viable population size (Fairbanks and Andersen, 1999).

This study may provide a baseline for future studies that seek to determine the genetic structure of naturally structured elephant populations and/or elephant populations that occur in continuous landscapes. Furthermore, by using a similar genetic approach as employed here and by integrating the results of Roever et al. (2013a), future researchers may identify ecologically

[17] I used the program Arlequin (Excoffier and Lischer, 2010) to calculate mtDNA diversity as the average pairwise differences (π), and nDNA diversity as the average number of alleles per locus and the Expected Heterozygocity averaged across all loci (He) and found mtDNA diversity to be 2.59, and nDNA diversity to be 7.5 and 0.71 respectively.

sensible linkages between elephant populations in other southern African countries. Identifying such landscape linkages may be crucial to the success of future conservation initiatives considering the current anthropocene era in which space for conservation is becoming an increasingly limited resource.

REFERENCES[18]

Ahlering, M.A., Maldonado, J.E., Fleischer, R.C., Western D. and Eggert, L.S. (2012) Fine-scale group structure and demography of African savanna elephants recolonizing lands outside protected areas. *Diversity and Distributions* **18**, 952-961

Allendorf, F.W. and Luikart, G. (2007) Conservation and the genetics of populations. (1st Edition) Blackwell Publishing, Oxford, UK

Allendorf, F.W., Luikart, G. and Aitken, S.N. (2013) Conservation and the genetics of populations. (2nd Edition) Blackwell Publishing, Wiley and Sons. LTD, West Sussex, UK

Anthony, N.M., Johnson-Bawe, M., Jeffery, K., Clifford, S.L., Abernethy, K.A., Tutin, C.E., Lahm, S.A., White, L.J.T., Utley, J.F., Wickings, E.J. and Bruford, M.W. (2007) The role of Pleostocene refugia and rivers in shaping gorilla genetic diversity in central Africa. *Proceedings of the National Academy of Sciences* **105**, 20432-20436

Archie, E.A., Maldonado, J.E., Hollister-Smith, J.A., Poole, J.H., Moss, C.J., Fleischer, R.C. and Alberts, S.C. (2008) Fine-scale population genetic structure in a fission-fusion society. *Molecular Ecology* **17**, 2666-2679

Archie, E.A., Moss, C.J. and Alberts, S.C. (2003) Characterization of tetranucleotide microsatellite loci in the African Savannah Elephant (*Loxodonta africana africana*). *Molecular Ecology: Primer Notes* **3**, 244-246

Archie, E.A., Moss, C.J. and Alberts, S.C. (2006) The ties that bind: genetic relatedness predicts the fission and fusion of social groups in wild African elephants. *Proceedings of the Royal Society: Biological sciences* **273**, 513-522

Archie, E.A. and Chiyo, P.I. (2012) Elephant behaviour and conservation: social relationships, the effects of poaching, and genetic tools for management. *Molecular Ecology* **21**, 765-778

[18] The references follow the style of Molecular Ecology

Armbruster, P. and Lande, R. (1993) A population viability analysis for African elephant (*Loxodonta africana*): How big should reserves be? *Conservation Biology* **7**, 602-610

Barnes, R.F.W. (1999) Is there a future for elephants in West Africa? *Mammal Review* **29**, 175-199

Ballard, J.W.O. and Whitlock, M.C. (2004) The incomplete natural history of mitochondria. *Molecular Ecology* **13**, 729-744

Beben, D. (2012) Crossings for animals – an effective method of wild fauna conservation. *Journal of Environmental Engineering and Landscape Management* **20**, 86-96

Beier, P. and Noss, R.F. (1998) Do habitat corridors provide connectivity? *Conservation Biology* **12**, 1241-1252

Bellemain, E., Swenson, J.E., Tallmon, D., Brunberg, S. and Taberlet, P. (2005) Estimating population size of elusive animals with DNA from Hunter-collected feces: four methods for brown bears. *Conservation Biology* **19**, 150-161

Birky, C.W. (Jr.) (2001) The inheritance of genes in mitochondria and chloroplasts: laws, mechanics, and models. *Annual Review of Genetics* **35**, 125-148

Bonin, A., Bellemain, E., Eidesen, P.B., Pompanon, F., Brochmann, C. and Taberlet., P. (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13**, 3261-3273

Branton, D., Deamer, D.W., Marziali, A., Bayley, H., Benner, S.A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A., Huang, X., Jovanovich, S., Krstic, P.S., Lindsay, S., Ling, X.S., Mastrangelo, C.H., Meller, A., Oliver, J.S., Pershin, Y.V., Ramsy, J.M., Riehn, R., Soni, G.V., Tavard-Cossa, V., Wanunu, M., Wiggin, M. and Schloss, J.A. (2008) The potential and challenges of nanopore sequencing. *Nature Biotechnology* **26**, 1146-1153

Brooks, H. (1994) The relationship between science and technology. *Research Policy* **23**, 447-486

Burn, R.W., Underwood, F.M., Blanc, J. (2010) Global trends and factors associated with the illegal killing of elephants: a hierarchical Bayesian analysis of carcass encounter data. *PloS one* **6**, e24165

Chamaillé-James, S., Fritz, H., Valeix, M., Murindagomo, F. and Clobert, J. (2008) Resource variability, aggregation and direct density dependence in an open context: the local regulation of an African elephant population. *Journal of Animal Ecology* **77**, 135-144

Charif, R.A., Ramey, R.R., Langbauer, W.R., Payne, K.B., Martin, R.B. and Brown, L.M. (2005) Spatial relationships and matrilineal kinship in African savanna elephant (*Loxodonta africana*) clans. *Behavioral Ecology and Sociobiology* **57**, 327-338

Chase, M. and Griffin, C.R. (2005) Ecology, population structure and movement of elephant populations in northern Botswana. *Final year-end report submitted to the Shikar Safari Club*

Chase, M.J. and Griffin, C.R. (2006) Elephant Distribution and Abundance in the Lower Kwando River Basin and Caprivi Strip. *US Fish and Wildlife Service, Washington, DC*

Chase, M. J. (2007) Home ranges, transboundary movements and harvest of elephants in northern Botswana and factors affecting elephant distribution and abundance in the Lower Kwando River Basin. (PhD Thesis). University of Massachusetts Amherst, Amherst, MA. *Retrieved from ProQuest (UMI Number: 3290051)*

Chase, M. J., and Griffin, C. R. (2009) Elephants caught in the middle: impacts of war, fences and people on elephant distribution and abundance in the Caprivi Strip, Namibia. *African Journal of Ecology* **47**, 223-233

Chase, M (2011) Dry season fixed-wing aerial survey of elephants and wildlife in northern Botswana. *Report by Elephants Without Borders, Kasane, Botswana, the Zoological Society of San Diego, California, USA, and the Department of Wildlife and National Parks, Gaborone, Botswana*

Chessel, D., Dufour, A.B. and Thiolouse, J. (2004) The ade4 package I: One table methods. *R News* **4**, 5-10

Chiyo, P.I., Archie, E.A., Hollister-Smith, J.A., Lee, P.C., Poole, J.H., Moss, C.J. and Alberts, S.C. (2011) Association patterns of African elephants in all-male groups: the role of age and genetic relatedness. *Animal Behaviour* **81**, 1093-1099

Comstock, K.E., Ostrander, E.A. and Wasser, S.K. (2003) Amplifying nuclear and mitochondrial DNA from African elephant ivory: a tool for monitoring the ivory trade. *Conservation Biology* **17**, 1840-1843

Comstock, K.E., Wasser, S.K. and Ostrander, E.A. (2000) Polymorphic microsatellite DNA loci identified in the African elephant (*Loxodonta africana). Molecular Ecology: Primer Notes* **9**, 993-1011

Comstock, K.E., Georgiadis, N., Pecon-Slattery, J., Roca, A.L., Ostander, E.A., O'Brien, S.J. and Wassers, S.K. (2002) Patterns of molecular genetic variation among African elephant populations. *Molecular Ecology* **11**, 2489-2498

Dankin, E.E. and Avise, J.C. (2004) Microsatellite null alleles in parentage analysis. *Heredity* **93**, 504-509

de Beer., Y. and van Aarde, R.J. (2008) Do landscape heterogeneity and water distribution explain aspects of elephant home range in southern Africa's arid savannas? *Journal of Arid Environments* **72**, 2017-2025

Dixon, J.D., Oli, M.K., Wooten, M.C., Eason, T.H., McCown, J.W. and Paetkau, D. (2006) Effectiveness of a regional corridor in connecting two Florida black bear populations. *Conservation Biology* **20**, 155-162

Donaldson, B.M. (2005) Use of highway underpasses by large mammals and other wildlife in Virginia and factors influencing their effectiveness. *Proceedings of the 2005 International Conference on Ecology and Transportation*, 433-441

Douglas-Hamilton, I. Krink, T. and Vollrath, F. (2005) Movement corridors of Africna elephants in relation to protected areas. *Naturwissenschaften* **92**, 158-163

Dudley, N. (Editor) (2008) Guidelines for Applying Protected Area Management Categories. *Gland, Switzerland, IUCN*

Eggert, L.S., Eggert, J.A. and Woodruff, D.S. (2003) Estimating population sizes for elusive animals: the forest elephants of Kakum National Park, Ghana. *Molecular Ecology* **12**, 1389-1402

Eggert, L.S., Ramakrishnan, U., Mundy, N.I. and Woodruff, D.S. (2000) Polymorphic microsatellite DNA markers in the African elephant (*Loxodonta africana)* and their use in the Asian elephant (*Elephas maximus*). *Molecular Ecology: Primer Notes* **9**, 2155-2234

Eggert, L.S., Rasner, C.A. and Woodruff, D.S. (2002) The evolution phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. *Proceedings of the Royal Society: London* **269**, 1993-2006

Eggert, L.S., Patterson, G. and Maldonado, J.E. (2007) The Knysna elephants: a population study conducted using faecal DNA. *African Journal of Ecology* **46**, 19-23

Epps, C.C., Wasser, S.K., Keim, J.L., Mutayoba, B.M and Brashares, J.S. (2013) Quantifying past and present connectivity illuminates a rapidly changing landscape for the African elephant. *Molecular Ecology* **22**, 1574-1588

Eriksson, J., Hohmann, G., Boesch, C. and Vigilant, L. (2004) Rivers influence the population genetic structure of bonobos (*Pan paniscus*). *Molecular Ecology* **13**, 3425-3435

Essop, M.F., Hall-Martin, A.J. and Harley, E.H. (1996) Mitochondrial DNA analysis of two southern African elephant populations. *Koedoe* **39**, 85-88

Evans, K.E. and Harris, S. (2008) Adolescence in male African elephants, *Loxodonta africana*, and the importance of sociality. *Animal Behaviour* **76**, 779-787

Excoffier, L. and Lischer, H.E. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analysis under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567

Fairbanks, D.J. and Andersen. W.R. (1999) Genetics: The Continuity of Life. Wadsworth and Brooks/Cole Publishing Companies, Pacific Grove, CA

Falush, D., Stephens, M. and Pritchard, J.K. (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7**, 574-578

Fernanado, P. and Lande, R. (2000) Molecular genetic and behavioral analysis of social organization in the Asian elephant (*Elephas maximus*). *Behavioral Ecology and Sociobiology* **48**, 84-91

Fernando, P., Pfrender, M.E., Encalada, S.E. and Lande, R. (2000) Mitochondrial DNA variation, phylogeography and the population structure of the Asian elephant. *Heredity* **84**, 362-372

Fernando, P., Vidya, T.N.C., Rajapakse, C., Dangolla, A. and Melnick, D.J. (2003) Reliable noninvasive genotyping: fantasy or reality? *Journal of Heredity* **94**, 115-123

Fishlock, V. and Lee, P.C. (2012) Forest elephants: fission-fusion and social arenas. *Animal Behaviour* **85**, 357-363

Fleischer, R.C., Perry, E.A., Muralidharan, K., Stevens, E.E. and Wemmer, C.M. (2001) Phylogeography of the Asian elephant (*Elephas maximus*) based on mitochondrial DNA. *Evolution* **55**, 1882-1892

Foster, M.L. and Humphrey, S.R. (1995) Use of highway underpasses by Florida panthers and other wildlife. *Wildlife Society Bulletin* **23**, 95-100

Frankham, R. (2005) Genetics and extinction. *Biological Conservation* **126**, 131-140

Frankham, R., Ballou., J.D. and Briscoe, D.A. (2004) A Primer of Conservation Genetics. Cambridge University Press, Cambridge, UK

Frankham, R., Ballou, J.D., Eldridge, M.D.B., Lacy, R.C., Ralls, K., Dudash, M.R. and Fenster, C.B. (2011) Predicting the probability of outbreeding depression. *Conservation Biology* **25**, 465-475

Frantzen, M.A.J., Silk, J.B., Ferguson, J.W.H., Wayne, R.K. and Kohn, M.H. (1998) Empirical evaluation of preservation methods for faecal DNA. *Molecular Ecology* **7**, 1423-1428

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

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

Georgiadis, N., Bischof, L., Templeton, A., Patton, J., Karesh, W. and Western, D. (1994) Structure and history of African elephant populations: I. Eastern and southern Africa. *Journal of Heredity* **85**,100-104

Gomes, I., Collins, A., Lonjou, C., Thomas, N.S., Wilkinson, J., Watson, M. and Morton, N. (1999) Hardy-Weinberg quality control. *Annals of Human Genetics* **63**, 535-538

Gerlach, G. and Musolf, K. (2000) Fragmentation of landscape as a cause for genetic subdivision in bank voles. *Conservation Biology* **14**, 1066-1074

Gobush, K.S., Mutayoba, B.M. and Wasser, S.K. (2008) Long-term impacts of poaching on relatedness, stress physiology, and reproductive output of adult female African elephants. *Conservation Biology* **22**, 1590-1599

Gobush, K., Kerr, B. and Wasser, S. (2009) Genetic relatedness and disrupted social structure in poached population of African elephants. *Molecular Ecology* **18**, 722-734

Graham, M.D., Douglas-Hamilton, I., Adams, W.M. and Lee, P.C. (2009) The movement of African elephants in a human-dominated land-use mosaic. *Animal Conservation* **12**, 445-455

Guillot, G., Santos, F. and Estoup, A. (2008) Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics: Application Note* **24**, 1406-1407

Hall-Martin, A.J. (1987) Role of musth in the reproductive strategy of the African elephant (*Loxodonta africana*). *Oryx* **83**, 616-620

Hanks, J. (2001) Conservation strategies for Africa's large mammals. *Reproduction, Fertility and Development* **13**, 459-468

Hanks, J. (2003) Transfrontier Conservation Areas (TFCA) in southern Africa: their role in conserving biodiversity, socioeconomic development and promoting a culture of peace. *Journal of Sustainable Forestry* **17**, 127-148

Harris, G.M., Russel, G.J., van Aarde, R.J. and Pimm, S.L. (2008) Rules of habitat use by elephants *Loxodonta africana* in southern Africa: insights for regional management. *Oryx* **42**, 66-75

Harrisson, K.A., Pavlova, A., Amos, J.N., Takeuchi, N., Lill, A., Radford, J.Q. and Sunnucks, P. (2012) Disrupted fine-scale population processes in fragmented landscapes despite large-scale genetic connectivity for a widespread and common cooperative breeder: the superb fairy-wren (*Malurus cyaneus*). *Journal of Animal Ecology* **82**, 322-333

Hoare, R.E. (1999) Determinants of human-elephant conflict in a land-use mosaic. *Journal of Applied Ecology* **39**, 689-700

Hoare, R.E. and du Toit, J.T. (1999) Coexistence between people and elephants in African savannas. *Conservation Biology* **13,** 633–639

Hoffman, J.I. and Amos, W. (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology* **14**, 599-612

Ishengoma, D.R.S., Shedlock, A.M., Foley, C.A.H., Foley, L.J., Wasser, S.K., Balthazary, S.T. and Mutayoba, B.M. (2008) Effects of poaching on bull mating success in a free ranging African elephant (*Loxodonta africana*) population in Tarangire National Park, Tanzania. *Conservation Genetics* **9**, 247-255

Ishida, Y., Georgiadis, N.J., Hondo, T. and Roca, A.L. (2012) Triangulating the provenance of African elephants using mitochondrial DNA. *Evolutionary Applications*, published online: ISSN 1752-4571

Jackson, T. P., and Erasmus, D. G. (2005) Assessment of seasonal home range use of elephants across southern Africa's seven elephant clusters. *Conservation Ecology Research Unit, University of Pretoria, South Africa.* Unpublished report for Peace Parks Foundation

Kohn, M., Knauer, F., Stoffella, A., Schröder, W. and Pääbo, S. (1995) Conservation genetics of the European brown bear – a study using excremental PCR of nuclear and mitochondrial sequences. *Molecular Ecology* **4**, 95-103

Lande, R. (1998) Anthropogenic, ecological and genetic factors in extinction and conservation. *Researches on Population Ecology* **40**, 259-269

Larkin,M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948

Laws, R.M. (1970) Elephants and habitats in north Bunyoro, Uganda. *East African Wildlife Journal* **8**, 163-180

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

Leal, S.M. (2005) Detection of genotyping errors and pseudo-SNPs via deviations from Hardy-Weinberg Equilibrium. *Genetic Epidemiology* **29**, 204-214

Lee, T. (2013) Determinants of protected area boundary crossings by savannah elephants, *Loxodonta africana*. (MSc Thesis). University of Pretoria, Pretoria, South Africa

Leggett, K.E.A. (2006) Home range and seasonal movement of elephants in the Kunene Region, northwestern Namibia. *African Zoology* **41**, 17-36

Loarie, S. R., Aarde, R. J. V., and Pimm, S. L. (2009) Fences and artificial water affect African savannah elephant movement patterns. *Biological Conservation* **142**, 3086-3098

Marrero, P., Fregel, R., Cabrera, V.M. and Nogales, M. (2009) Extraction of high-quality host DNA from feces and regurgitated seeds: a useful tool for vertebrate ecological studies *Biological Research* **42**, 147-151

Mbaiwa, J.E. and Mbaiwa, O.I. (2006) The effects of veterinary fences on wildlife populations in Okovango delta, Botswana. *International Journal of Wilderness* **12**, 17-41

McKelvey, K.S. and Schwartz, M.K. (2004) Genetic errors associated with population estimation using non-invasive molecular tagging: problems and new solutions. *The Journal of Wildlife Management* **68**, 439-448

McRae, B.H. and Beier, P. (2007) Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences* **104**, 19885-19890

Melton, D.A. (1985) The status of elephants in northern Botswana. *Biological Conservation* **31**, 317-333

Metcalfe, S. and Kepe, T. (2008) 'Your elephant on our land' The struggle to manage wildlife mobility an Zambian communal land in the Kavango-Zambezi Transfrontier conservation area. *The Journal of Environment & Development* **17**, 99-117

Monteiro, L., Bonnemaison, D., Vekris, A., Petry, K.G., Bonnet, J., Vidal, R., Cabrita, J. and Mégraud, F. (1997) Complex polysaccharides as PCR inhibitors in feces: Helicobacter pylori model. *Journal of Clinical Microbiology* **35**, 995-998

Moss, C.J. and Lee, P.C. (2011) Chapter 13: Female social dynamics: fidelity and flexibility, The Amboseli elephants: a long-term perspective on a long-lived mammal, The University of Chicago Press, Chicago, USA

Munshi-South, J. (2011) Relatedness and demography of African forest elephants: inferences from noninvasive fecal DNA analyses. *Journal of Heredity* **102**, 391-398

Murphy, M.A., Kendall, K.C., Robinson, A. and Waits, L.P. (2007) The impact of time and field conditions on brown bear (*Ursus arctos*) faecal DNA amplification. *Conservation Genetics* **8**, 1219-1224

Murphy, M.A., Waits, L.P., Kendall, K.C., Wasser, S.K., Higbee, J.A. and Bogden, R. (2002) An evaluation of long-term preservation methods for brown bear (*Usus arctos*) faecal DNA samples. *Conservation Genetics* **3**, 435-440

Navidi, W., Arnheim, N. and Waterman, M.S. (1992) A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. *American Journal of Human Genetics* **50**, 347-359

Nsubuga, A.M., Robbins, M.M., Roeder, A.D., Morin, P.A., Boesch, C. and Vigilant, L. (2004) Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. *Molecular Ecology* **13**, 2089-2094

Nyakaana, S. and Arctander, P. (1999) Population genetic structure of the African elephant in Uganda based on variation at mitochondrial and nuclear loci: evidence for male-biased gene flow. *Molecular Ecology* **8**, 1105-1115

Nyakaana, S., Arctander, P. and 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

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

Okello, J.B.A., Masembe, C., Rasmussen, H.B., Wittemyer, G., Omondi, P., Kahindi, O., Muwanika, V.B., Arctander, P., Douglas-Hamilton, I., Nyakaana, S. and Siegismund, H.R. (2008b) Population genetic structure of savannah elephants in Kenya: conservation and management implications. *Journal of Heredity* **99**, 443-452

Osborn, F. V. and Parker, G. E. (2003) Towards an integrated approach for reducing the conflict between elephants and people: a review of current research. *Oryx* **37**, 80-84

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L. and Hofreiter, M. (2004) Genetic analyses from ancient DNA. *Annual Review of Genetics* **38**, 645-679

Paetkau, D. (2003) An empirical exploration of data quality in DNA-based population inventories. *Molecular Ecology* **12**, 1375-1387

Panasci, M., Ballard, W.B., Breck, S., Rodrigues, D., Densmore III, L.D., Wester, D.B. and Baker, R.J. (2011) Evaluation of fecal DNA preservation techniques and effects of sample age and diet on genotyping success. *Journal of Wildlife Management* **75**, 1616-1624

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

Piggott, M.P. and Taylor, A.C. (2003) Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. *Wildlife Research* **30**, 1-13

Piggott, M.P., Banks, S.C., Stone, N., Banffy, C. and Taylor, A.C. (2006) Estimating population size of endangered brush-tailed rock-wallaby (*Petrogale penicillata*) colonies using faecal DNA. *Molecular Ecology* **15**, 81-91

Pritchard, J.K., Stephens, M. and Donelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959

Puechmaille, S.J. and Petit, E.J. (2007) Empirical evaluation of non-invasive capture-mark-recapture estimation of population size based on a single sampling session. *Journal of Applied Ecology* **44**, 843-852

Regnaut, S., Lucas, F.S. and Fumagalli, L. (2006) DNA degradation in avian faecal samples and feasibility of non-invasive genetic studies of threatened capercaillie populations. *Conservation Genetics* **7**, 449-453

Roever, C.L. (2013) Spatial determinants of habitat use, mortality and connectivity for elephant populations across southern Africa. (PhD Thesis). University of Pretoria, Pretoria, South Africa

Roever, C.L, van Aarde, R.J., Legget, K. (2013a) Functional connectivity within conservation networks: Delineating corridors for African elephants. *Biological Conservation* **157**, 128-135

Roever, C.L., van Aarde, R.J. and Chase, M.J. (2013b) Incorporating mortality into habitat selection to identify secure and risky habitats for savannah elephants. *Biological Conservation*, in press

Rueness, E.K., Stenseth, N.C., O'Donoghue, M., Boutin, S., Ellegren, H. and Jakobsen, K.S. (2003) Ecological and genetic spatial structuring in the Canadian lynx. *Nature* **425**, 69-72

Schmitt, T., Röber, S. and Seitz, A. (2005) Is the last glaciations the only relevant event for the present genetic population structure of the meadow brown butterfly *Maniola jurtina* (Lepidoptera: Nymphalidae)? *Biological Journal of the Linnean Society* **85**, 419-431

Shannon, G., Page, B.R., Duffy, K.J. and Slotow, R. (2006) The role of foraging behaviour in the sexual segregation of the African elephant. *Oecologia* **150**, 344-354

Shinde, D., Lai, Y., Sun., F. and Arnheim, N. (2003) *Taq* DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: $(CA/GT)_n$ and $(A/T)_n$ microsatellites. *Nucleic Acids Research* **31**, 974-980

Smouse, P.E. and Peakall, R. (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* **82**, 561-573

Stenglein, J.L., De Barba, M., Ausband, D.E. and Waits, L.P. (2010) Impacts of sampling location within a faeces on DNA quality in two carnivore species. *Molecular Ecology Resources* **10**, 109-114

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P. and Bouvet, J. (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194

Taberlet, P. and Luikart, G. (1999) Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society* **68**, 41-55

Taberlet, P., Waits, L.P. and Luikart, G. (1999) Noninvasive genetic sampling: look before you leap. *Trends in Ecology and Evolution* **14**, 323-327

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-595

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetic Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**, 2731-2739

The Elephant Database (2011) Population Survey: North East Namibia Landscape (C) 1995-2013 IUCN - The International Union for the Conservation of NatureAvailable at: http://www.elephantdatabase.org/population_submissions/199 (Last Updated 30 August 2012) (Accessed 09 April 2013)

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680

Transfrontier Conservation Consortium. (2006) Final report: Pre-feasibility study of the proposed Kavango-Zambezi Transfrontier Conservation Area (Vols. 1-3*). Cape Town, South Africa*

Vidya, T.N.C. and Sukumar, R. (2005) Social organization of the Asian elephant (*Elephas maximus*) in southern India inferred from microsatellite DNA. *Journal of Ethology* **23**, 205-210

Wasser, Sk., Joseph, C.W., Drori, O., Stephen, K.E., Mailand, C., Mutayoba, B., Stephens, M. (2008) Combating the illegal trade in African elephant ivory with DNA forensics. *Conservation Biology* **22**, 1065-1071

Wasser, S.K., Mailand, C., Booth, R., Mutayoba, B., Kisamo, E., Clark, B. and Stephens, M. (2007) Using DNA to track the origin of the largest ivory seizure since the 1989 trade ban. *Proceedings of the National Academy of Science* **104**, 4228-4233

Wasser, S.K., Shedlock, A.M., Comstock, K., Ostander, E.A., Mutayoba, B. and Stephens, M. (2004) Assigning African elephant DNA to geographic region of origin: applications to the ivory trade. *Proceedings of the National Academy of Sciences: Genetics* **101**, 14847-14852

Watson, J.D. and Crick, F.H.C. (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**, 737-738

Wehausen, J.D., Ramey, R.R. and Epps, C.W. (2004) Experiments in DNA extraction and PCR amplification from bighorn sheep feces: the importance of DNA extraction method. *Journal of Heredity* **95**, 503-509

Weir, B.S. (1996) Genetic data analysis II: methods for discrete population genetic analysis, Sinauer Associates, Sunderland, Massachusetts

Whitehouse, A. M. and Harley, E. H. (2001) Post-bottleneck genetic diversity of elephant populations in South Africa, revealed using microsatellite analysis. *Molecular Ecology* **10**, 2139-2149

Wittemyer, G., Douglas-Hamilton, I. and Getz, W.M. (2005) The socioecology of elephants: analysis of the processes creating multitiered social structures. *Animal Behaviour* **69**, 1357-1371

Wittemyer, G. and Getz, W.M. (2007) Hierarchical dominance structure and social organization in African elephants, *Loxodonta africana*. *Animal Behaviour* **73**, 671-681

Wittemyer, G., Getz., W.M., Vollrath, F. and Douglas-Hamilton, I. (2007) Social dominance, seasonal movements, and spatial segregation in African elephants: a contribution to conservation behaviour. *Behavioural Ecology and Sociobiology* **61**, 1919-1931

Wittemyer, G., Daballen, D. and Douglas-Hamilton, I. (2013) Comparative demography of an at-risk African elephant population. *PLoS ONE* **8**, Doi:10.1371/journal.pone.0053726

Wright, S. (1943) Isolation by Distance. *Genetics* **28**, 114-138

Vandergast, A.G., Perry, W.M., Lugo, R.V. and Hathaway, S.A. (2011) Genetic landscapes GIS Toolbox: tools to map patterns of genetic divergence and diversity. *Molecular Ecology Resources* **11**, 158-161

van Aarde, R.J. and Jackson, T.P. (2007) Megaparks for metapopulations: Addressing the causes of locally high elephant numbers in southern Africa. *Biological Conservation* **134**, 289-297

van Aarde, R.J., Jackson, T.P. and Ferreira, S.M. (2006) Conservation science and elephant management in southern Africa. *South African Journal of Science* **102**, 385-388

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. and Shipley, P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538

Xu, J., Turner, A., Little, J., Bleecker, E.R. and Meyers, D.A. (2002) Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? *Human Genetics* **111**, 573-574

Young, K. D., Ferreira, S. M. and Van Aarde, R. J. (2009) Elephant spatial use in wet and dry savannas of southern Africa. *Journal of Zoology* **278**, 189-205

Young, K.D. and van Aarde, R.J. (2010) Density as an explanatory variable of movements and calf survival in savanna elephants across southern Africa. *Journal of Animal Ecology* **79**, 662-673

APPENDICES

Appendix I. A summary of the findings of studies that investigated the genetic structure of elephant populations at three distinct scales – local, regional and continental.

Authors	Findings
ISOLATED populations (local scale)	
Archie et al. (2006) Kenya Markers: -mtDNA -11 microsatellites	Genetic relatedness predicted group fission. Associations between core social groups persist for decades after original maternal kin have died.
Archie et al. (2008) Kenya Markers: -11 microsatellites	Female matrilocality creates co-ancestry within core social groups. There was significant genetic differentiation between core social groups. There was male mediated gene glow – cohorts of similar-aged paternal relatives.

FRAGMENTED: country-specific populations (regional scale)	
Epps et al. (2013) Tanzania Markers: -16 microsatellite loci	They found differences between historical and contemporary connectivity between elephant populations. Historical connectivity was strongly influences by slope but not human settlement. Contemporary connectivity was influenced most by human settlements. Recent gene flow was strongly influenced by slope but was also correlated with contemporary resistance.
Nyakaana and Arctander (1999) Uganda Markers: -mtDNA -4 microsatellites	They found geographically localized haplotypes. There was significant population differentiation between the three populations when considering mtDNA. Nuclear DNA markers were analysed individually, where some markers showed significant differentiation between different populations while others did not. There were incongruent patterns of genetic variation within and between populations for the two genetic systems. Attributed the incongruent patterns to differences in the effective population sizes of the two genomes, and male-biased dispersal.
Okello et al. (2008a) Keyna Markers: -20 microsatellites	The peak of the poaching epidemic in the 1970 caused detectable temporary genetic impacts, with genetic diversity rebounding as juveniles became mature. Climatic history shaped the distribution and genetic history of elephants.

Okello et al. (2008b) Kenya Markers: -mtDNA CR (LafCR1,2) -20 microsatellites	Mitochondrial DNA showed significant differentiation among 3 geographic regions, with lower differentiation among populations within regions. They attributed the differences in mtDNA genetic differentiation to matrilineal social structure, female natal philopatry, and ancient vicariance. Nuclear DNA showed significant differentiation among pops within regions, but no regional significant differentiation. The attribute the nDNA genetic differentiation to male-mediated gene flow.
FRAGMENTED: international populations (Continental scale)	
Comstock et al. (2002) International Markers: -16 microsatellites	They found species level genetic differentiation. Savannah elephant populations showed modest levels of phylogeographic subdivision suggesting recent population isolation and restricted gene flow between locales. Savannah elephants showed significantly lower genetic diversity than forest elephants, reflecting a possible founder effect.
Eggert et al. (2002) International Markers: -mtDNA (Cytochrome *b*) -4 microsatellites	They found three main "groups" of elephants, namely, Central African forest elephants, west African forest and savannah elephants, central African east African and south African savannah elephants . This study suggests possible large scale structuring. They did not find adequate support for classifying African forest and savannah elephants as distinct species. The attribute the genetic patterns of variation in African elephants to historical climatic variation, allopatric divergence, and two distinct radiations leading to admixture and mtDNA haplotype divergence.

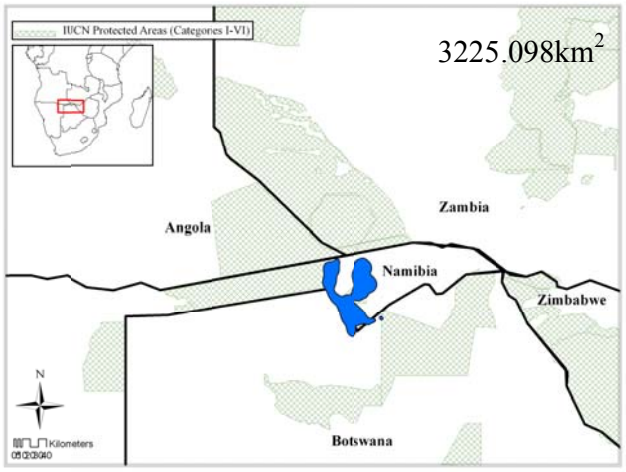
Georgiadis et al. (1994)	They found marked population subdivision at the continental level.
Eastern and southern Africa	There was evidence of Isolation by distance at the regional level.
Markers:	They found protracted gene flow across the content with one derived haplotype was found at all sampling locations.
-mtDNA	Mitochondrial DNA clades coexist at distant (<2000km) sampling locations.
	Their findings do not support ancestral isolation in allopatry, because there is no geographic barrier likely to have obstructed gene flow between regions for long enough.
	They suggest gene flow has maintained a sufficiently large effective population size for representatives of clades that diverged at least 4 million years ago to have persisted by chance within subdivided, but not isolated, populations.
Ishida et al. (2012)	They found regionally restricted mtDNA subclades.
International	Nuclear partitioning followed species boundaries and not mtDNA subclade boundaries.
Markers	Nuclear and mtDNA markers differed in phylogeography.
-mtDNA	Combining mtDNA and nuclear markers increased the accuracy of predicting provenance.
-compared to previous microsatellite data	
Nyakaana et al. (2002)	They found significant genetic differentiation between populations within regions and also between regions.
International	Mitochondrial DNA subdivision was higher than that of microsatellites.
Markers:	Attribute the genetic structure to Pleistocene refugia and allopatric differentiation, and recent admixture following a recent population expansion.
-mtDNA	Microsatellites showed less genetic differentiation between populations than mtDNA.
-4 microsatellites	

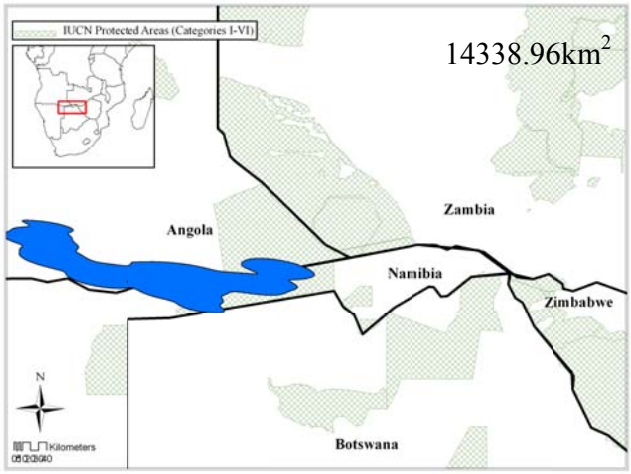
Appendix II. The categories of protected areas specified by the International Union for the Conservation of Nature (Dudley, 2008).

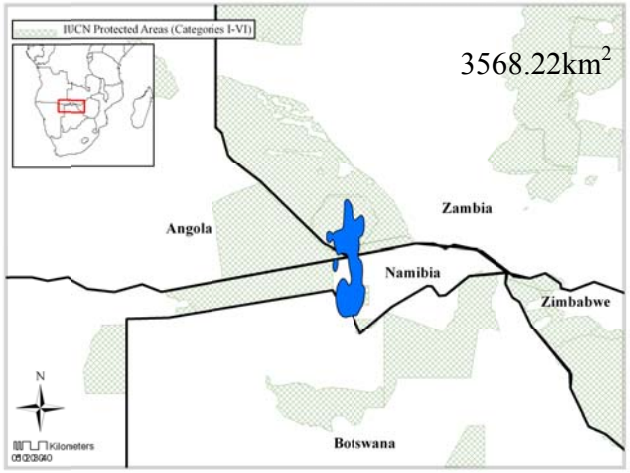
IUCN Category	Definition	Description
Category Ia &Ib	Strict Nature Reserves and Wilderness Areas	Nature Reserves and Wilderness areas with very limited human use.
Category II	National Parks	National parks that provide protection for functioning ecosystems, but are more lenient towards human use than Categories Ia & b.
Category III	Natural Monument of Feature	These areas are specifically allocated to protect natural monuments, which includes areas of high cultural or spiritual value.
Category IV	Habitat or Species Management Areas	Specific areas of conservation that are set aside for the continuous protection of a specific species or habitat.
Category V	Protected Landscape or Seascape	Includes entire bodies of land or ocean with the objective to safeguard regions with distinct ecological, biological, cultural or scenic value.
Category VI	Protected area with sustainable use of natural resources	Provides for the mutually beneficial relationship between nature conservation and the sustainable management of natural resources.

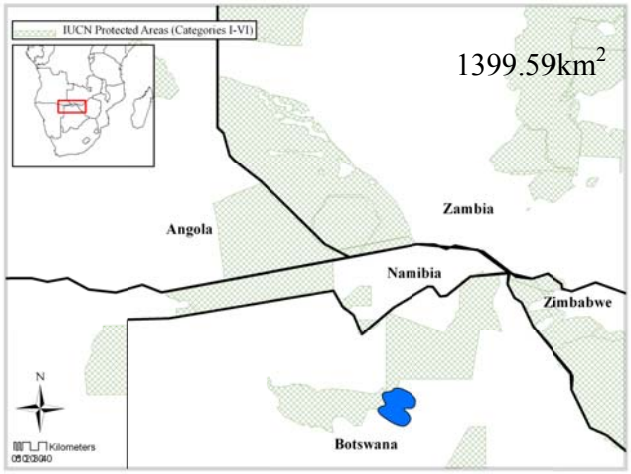
Appendix III. The annual 95% kernel home ranges of six female (red) and six male (blue) elephants in the KAZA-TFCA. Home range sizes are indicated in the top right corner of each map.

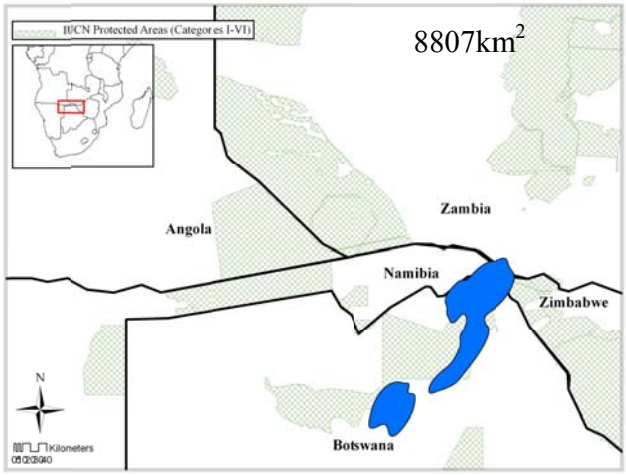
Appendix III. Continued

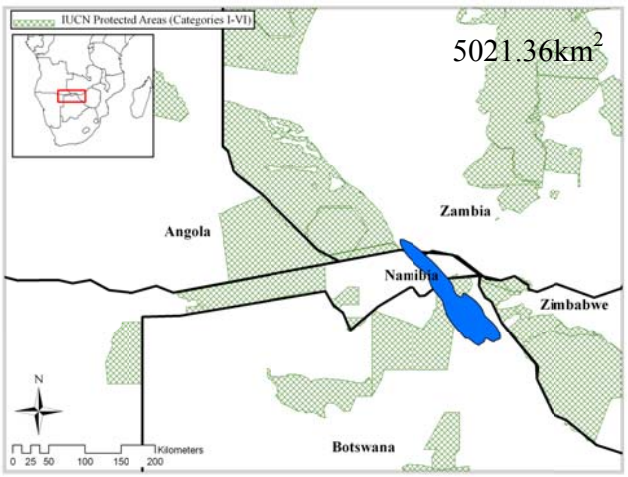








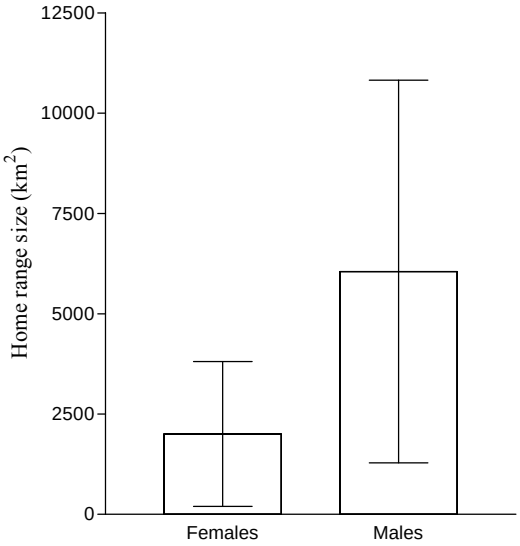




Appendix IV. A description of the spatial data and methods used to calculate and compare 95% kernel home ranges of 12 elephants that occur in the KAZA-TFCA.

I used spatial telemetry data for 12 elephants[19] (six males and six females) that occur in the KAZA region to calculate annual home ranges. The home ranges were delineated using kernel density isopleths that contained 95% of all the telemetry location points. Each of the isopleths were calculated using the kernel density estimates (kde) and isopleths extensions of the Geospatial Modelling Environment (Version 0.7.1.0 © Hawthorne L. Beyer 2009-2012) using the SCV smoothing parameter[20]. The telemetry data was collected using Africa Wildlife Tracking collars (model SM 2000E, Pretoria, S.A.) and Data Loggers, which used the Garmen GPS system to relay location data to a satellite from which it was downloaded to the CERU (Conservation Ecology Research Unit) spatial database. Data was filtered to include one location point per 24 hour period. These home ranges were then individually projected on a geographical map using ArcMap (©ESRI 2011, ArcGIS Desktop: Release 10, Redlands, CA: Environmental Systems Research Institute) (Appendix X2). The sizes of the annual home ranges were compared by implementing a Mann-Withney U test[21] in the program GraphPad Prism (Version 3, © 1999, GraphPad Software Incorporated, San Diego California, USA.).

[19] The telemetry data used in this thesis forms part of the CERU spatial database. Spatial data used to calculate home ranges for three of these elephants forms part of a data sharing agreement between CERU and Elephants Without Borders.
[20] Smoothing parameters were tested in a previous study by Roever (2013).
[21] Although a normality test showed the data to be normally distributed, I chose to use a Mann-Whitney U test instead of an unpaired T-test due to the small sample size.



Appendix V. Male elephants have significantly larger ($p<0.5$) annual home range sizes (n=6; mean annual home range = 6050 ± 1947km^2) than female elephants (n=6; mean annual home range = 2003 ± 738.2km^2).

Appendix VI. A modified protocol for nuclear and mitochondrial DNA isolation from degraded elephant faecal DNA. The asterisks indicate the steps of the original protocol (Qiagen DNA Stool Mini Kit - Southern Cross Biotechnology, Cape Town, SA) that were modified.

Step	Description
1*	Mix sample bottle contents till relatively homogenized. Centrifuge 2 ml of this mixture for 30 minutes. Discard supernatant. Repeat the process three times and use the remaining pellet as the starting material for DNA extraction.
2*	Add 1.7ml Buffer ASL to each sample and vortex continuously for 15 minutes.
3*	Centrifuge sample at full speed for 30 minutes to pellet stool particles
4	Pipet 1.4 ml of the supernatant into a new 2ml microcentrifuge tube and discard the pellet.
5*	Add 1 InhibitEX tablet to each sample and vortex immediately and continuously for 5 minutes. Incubate suspension for 5 minutes at room temperature to allow inhibitors to adsorb to InhibitEX matrix.
6*	Centrifuge sample at full speed to 6 minutes to pellet stool particles and inhibitors bound to InhibitEX matrix
7	Immediately after centrifuge stops, pipet all of the supernatant into a new 1.5ml microcentrifuge tube and discard the pellet. Centrifuge at full speed for 3 minutes.
8	Pipet 25 µl proteinase K into a new 2 ml microcentrifuge tube.
9	Pipet 600 µl supernatant from step 7 to the 2 ml microcentrifuge tube containing proteinase K.
10	Add 600 µl Buffer AL and vortex for 15 seconds.
11*	Incubate at 70°C for 30 minutes.
12	Add 600 µl of ethanol to the lysate ad mix by vortexing.
13	Label the lid of a new QIAamp spin column in a 2ml collection tube. Carefully apply 600 µl lysate from step 12 to the QIAamp spin column without moistening the rim. Close the cap and centrifuge at full speed for 1 minute. Place the QIAamp spin column in a new 2 ml collection tube, and discard the tube containing the filtrate.
14	Carefully open the QIAamp spin column, apply a second aliquot of 600 µl lysate and centrifuge at full speed for 1 minute. Place the QiaAmp spin column in a new 2 ml collection tube, and discard the tube containing the filtrate.
15	Repeat step 14 to load the third aliquot of the lysate onto the spin column.
16	Carefully open the QIAamp spin column and add 500 µl Buffer AW1. Close the cap and centrifuge at full speed for 1 minute. Place the QIAamp spin column in a new 2 ml collection

tube, and discard the collection tube containing the filtrate.

17 Carefully open the QIAamp spin column and add 500 µl Buffer AW2. Close the cap and centrifuge at full speed for 3 minutes. Discard the collection tube containing the filtrate.

18 Place the QIAamp spin column in a new 2 ml collection tube and discard the old collection tube with the filtrate. Centrifuge at full speed for 1 minute.

19* Transfer the QIAamp spin column into a new, labelled 1.5 ml microcentrifuge tube. Carefully open the QIAamp spin column and pipet 100 µl Buffer AE directly onto the QIAamp membrane. Close the cap and incubate for 10 minutes at room temperature, then centrifuge at full speed for 10 minutes to elute DNA.

Appendix VII. The respective reagents and protocols used in Cycle Sequencing and Sodium Acetate Precipitation of elephant faecal mtDNA using a D-loop control region primer set.

	Cycle Sequencing			**Sodium Acetate Precipitation**		
	Reagents[22]:	Big Dye	2 µl	Reagents:	Ice cold absolute ethanol	220.0 µl
		Primer (3.2 pmol/µl)	1 µl		SABAX water	44.0 µl
		5X Buffer	1 µl		3M Sodium Acetate	2.2 µl
		DNA	6 µl		71 % ethanol	200 µl
		SABAX water	(10 – 4 – x) µl		PCR product	10 µl
1	Combine all reagents			1	Make up a master mix containing the ethanol, water and NaAc	
2	Run on the following two stage thermocycle:			2	Add 250 µl of the master mix to a new 0.5 ml tube	
	Stage 1: run once (no replicates)			3	Add the PCR product and vortex well	
	96 °C for 1 minute			4	Centrifuge for 15 mins at 12 000 rpm	
	Stage 2: 25 replicates			5	Pipette off supernatant and spin again for 1s then remove all left over supernatant with a small pipette set to 10ul	
	96 °C for 10 seconds					
	57 °C (T_a) for 5 seconds			6	Add 190 µl 71 % ethanol and centrifuge for 10 mins at 12 000 rpm	
	60 °C for 4 minutes			7	Repeat steps 5 and 6	
	4 °C hold until precipitation			8	Pipette off supernatant and dry on a heat block for 3 min at 50 °C	
				9	Elute the purified, precipitated DNA in 30 µl SABAX water	

[22] Reagent quantity as necessary for a quarter volume reaction.

Appendix VIII. A justification of the analysis methods used in Chapter 4.

Sequence alignment, data file generation and haplotype analysis (mitochondrial DNA)

Sequences were confirmed as African elephant mtDNA sequences belonging to the genus *Loxodonta* by comparison with published sequences in the genetic sequence database GenBank (National Centre for Biotechnology Information). The sequences were then individually viewed and processed using the program CLC Bio Genomics workbench (CLC Bio, Cambridge, MA). Sequences were aligned using the online freeware program Webprank (Goldman Group Software ©, European Bioinformatics Institute, 2011). Aligned sequences were viewed and reformatted into a FASTA output file using ClustalX 2.0 (Larkin et al., 2007). The aligned FASTA file was used as an input file for the program DnaSP 5.10.01 (Librado and Rozas, 2009). DnaSP was used to generate a haplotype data file and a Roehl data file. The haplotype data file was used to calculate haplotypic frequencies (relative to the sample size, n= 88), and to test for neutrality using Tajima's and Fu's test (Tajima, 1989; Fu, 1997) of selective neutrality using the program Arlequin (Excoffier and Lischer, 2010). The Roehl data file (RDF) was used in the construction of a Median-Joining network in the freeware program Network 4.6.1.0 (Fluxus Technology Ltd ©).

The FASTA file was also used as an input dataset to construct a midpoint-rooted Neighbour-Joining tree in MEGA 5 (Tamura et al., 2011). The Neighbour-Joining tree incorporated the Maximum Composite Likelihood method for calculating evolutionary distances (based on the number of base substitutions per site), with a 1000 bootstrap repeat to test the reliability of the resulting lineages. I also calculated the Best-fit substitution model of evolutionary change for my data using MEGA 5. These models were tested using Bayesian Information Criterion (BIC) and

corrected Akaike Information Criterion (AICc). Of these models, the best fit model available for analysis in Arlequin and MEGA 5 was the Kamura-2-Parameter (with a discrete gamma distribution) model of evolutionary change (BIC = 3683.102; AICc = 2160.270) and was thus incorporated as a correctional method where possible.

Although most consider mtDNA markers as neutral markers because they do not undergo recombination (Birky, 2001), Ballard and Whitlock (2004) show that both direct and indirect selection influence mitochondria, and that it is not safe to assume *a priori* that mtDNA evolves as a strictly neutral marker. Because the ensuing population structure assumes linkage equilibrium, it was necessary to test whether the mtDNA marker I used in this study is a selectively neutral marker. I tested for neutrality by implementing two separate neutrality tests in the program Arlequin (Excoffier and Lischer, 2010). The first is the Tajima's test of selective neutrality (Tajima, 1989), which is based on the infinite-site model without recombination, and investigates the relationship between the number of segregating sites and the average number of nucleotide differences (test statistic D). Although also based on the infinite site model without recombination, I additionally implemented the Fu's F_s test of selective neutrality (Fu, 1997) because Fu's Fs test can be more sensitive to demographic changes, and also provide insights into the possible demographic factors that influence the marker neutrality tests. Fu's F_s tests the probability of observing a random neutral sample with a number of alleles similar or smaller than the observed allele value. A negative F_s (excess of alleles) indicates the possibility of a recent population expansion, whereas a positive F_s (allele deficiency) may be indicative of a recent population bottleneck. When using Fu's Fs, it is important to remember that only P values lower than 0.2 (not $p<0.5$) are considered significant (Fu and Li, 1993;Fu, 1997).

Finally, I used the program ArcMap (©ESRI 2011, ArcGIS Desktop: Release 10, Redlands, CA: Environmental Systems Research Institute) to create a visual representation of the geographic distribution of the individual-based haplotypes. I created a file containing the geographic coordinates of each individual (n=88) and assigned a haplotype to each individual. I then used this file as to create haplotype distribution map in ArcMap, and labelled each of the different haplotypes with a unique colour.

Population Structure Analysis (mitochondrial and nuclear DNA)

To determine how many sub-populations (K) of genetically similar individuals are present in my sampling region, I implemented the program Geneland (Guillot et al., 2008). Geneland is a program add-on package written for the free-ware statistical program R (ver 2.13.1 © 2011 The R Foundation for Statistical Computing). Geneland uses geo-referenced multi-locus genetic data to calculate systematic variation of allele frequencies. The allelic frequency variation is then used to infer a population structure (where Geneland refers to the number of populations or sub-populations (K) as the number of "clusters" of genetically similar individuals) in such a manner that the departure from Hardy-Weinberg and Linkage equilibrium is minimised. To create a multi-locus genotype file using DNA sequence data, I converted the sequence data for the polymorphic sites (n=25) to binary code, producing 25 loci with the possibility of 4 alleles (ACTG) at each of the sites. The binary sequence data was then imported into Geneland as haploid data file (this is to account for the singular data entry at each individual locus). The nuclear data consisted of genotype data for 100 individuals that were genotyped for 10 nuclear markers (as described in the Data section). The samples for each data type were geo-referenced and an individual label file was used to allow for post-analysis identification of individuals. Because of the possibility of having a continuous population, it is necessary to use population

structure analysis programs that can test for the possibility of having only one population. I thus selected Geneland, because it uses Markov-Chain Monte-Carlo simulations in combination with Bayesian Clustering Algorithms (BCAs) to determine the possibility of having one or many populations. Another program that is often used in population structure analysis is the program STRUCTURE (Prichard et al., 2000). Because STRUCTURE (Prichard et al., 2000) uses posterior-informed BCAs it is impossible to test for the possibility of only having a single population when combining multiple runs for each predicted number of populations (K). For interest sake, I did analyse both my datasets using STRUCTURE, but I abandoned this approach because both the mtDNA and nDNA analysis resulted in K=2, meaning that when using only posterior-informed BCAs (as compared to the geographically informed BCAs of Geneland) STRUCTURE may be unable to detect any population structure. In short, STUCTURE could possibly indicate no population structure, where all individuals are grouped together (K=1), but it is impossible to test this.

My run parameters for the Geneland analysis included using a spatial uncertainty coordinate of 0.01, the possibility of having 1 to 10 populations, 100 000 Markov-Chain Monte-Carlo (MCMC) iterations with a thinning factor of 100. For both data types, I ran the algorithms using both the Correlated Allele Frequency Model, and the Spatial Model, but not for the Null Allele Model (Guillot et al., 2008).

After I identified the number of sub-populations (K) for each of the data types that most accurately represent the genetic variability found in my sampling region (where allele frequency departure from Hardy-Weinberg and Linkage Equilibrium was the least), I determined if these sub-populations were significantly different in their genetic composition. I did an Analysis of Molecular Variance (AMOVA) by calculating inter-population genetic variation (F_{st}) for both

nDNA and mtDNA, and also corrected for microsatellite allele length by calculating inter-population genetic variation when incorporating the step-wise mutation model (R_{st}). The inter-population genetic variation (F_{st}) ranges from zero to one, where a F_{st} of zero would mean that all sub-populations in the study region have equal allele frequencies, indicating that the population is in complete panmixia. A F_{st} of one would meant that all the sub-populations in the study area are fixed for different alleles, thus indicating that all the sub-populations are genetically differentiated from each other. The range of inter-population variation is called the Fixation Index. The mtDNA analysis was done in Arlequin using the Kamura-2-Parameter (Excoffier and Lischer, 2010) while nDNA analysis was done using GenAlEx (Peakall and Smouse, 2006).

Because of the large extent of the sampling area (> 90 000km2)[23], it is highly possible that Isolation By Distance (IBD) (first proposed by Sewell Wright, 1943) may be an important factor influencing the population structure. The theory of IBD is based on the premises that geographic distance is positively correlated to genetic distance between individuals, where individuals that are further apart geographically will also differ more genetically. To find out if IBD can explain the genetic differentiation of elephants in my sampling area, I first calculated the geographic distance between sub-populations by measuring the distances (in kilometres) between the centroids of each of the sub-populations. I calculated the centroid of a sub-population using a combination of ArcGIS (ArcMap 10) (©ESRI 2011, ArcGIS Desktop: Release 10, Redlands, CA: Environmental Systems Research Institute) and the Geospatial Modeling Environment (GME) (Version 0.7.1.0 © Hawthorne L. Beyer 2009-2012). I started by creating a shape file for all the location points in a sub-population, after which I changed the spatial reference system (WGS1984) of each of the original shape files by projecting them into the Albert's Equal Area

[23] A Minimum Convex Polygon calculated for all the samples collected in this study showed that the total area included in our sampling range was more than 90 000km^2.

Conic projection for Africa using ArcMap 10 (©ESRI 2011, ArcGIS Desktop: Release 10, Redlands, CA: Environmental Systems Research Institute). This projection does not only produce more accurate spatial features, but is also a necessity for the use of files in GME. I calculated the Minimum Convex Polygon (MCP) for each of the shape files containing the location points using the "genmcp" function in GME. To calculate the centroid of each MCP, I used the "genpointinpol" function in GME, specifying that I want one point per MCP polygon, and that that point should be the centroid. I re-opened the centroid shape files in ArcMap and measured the geographic distances between all centroids. I then used the distances between sub-populations to construct a geographic distance matrix between sub-populations. I used pairwise genetic distances between sub-populations (calculated in Arlequin (Excoffier and Lischer, 2010) to create a genetic distance matrix. I finally tested if IBD could explain the sub-population differentiation by comparing the two matrices in the free-ware statistical program R (ver 2.13.1 © 2011 The R Foundation for Statistical Computing) using the ade4 package (Chessel et al., 2004) (See R-Scripts in Supplementary Material), to see if there was any correlation between the geographic and genetic distances between the sub-populations.[24]

The final analysis for this chapter included finding a way to project genetic distances between individuals on a geographic landscape. The Genetic Landscapes GIS Toolbox (Vandergast et al., 2011) provides tools to create genetic divergence and diversity landscapes in ArcGIS. The "Single Species Genetic Divergence" tool provided in the Landscape GIS Toolbox creates a genetic divergence raster surface from pairwise sub-population genetic distances for a single species or marker/locus. This tool functions by using a spatial interpolation algorithm to generate a landscape surface from the mapped genetic distance values between individuals in a study area.

[24] Geneland found the microsatellite nDNA database to partition into only two distinct populations, thus making it impossible to do a Mantel Test Correlation. I thus used only the mtDNA to test for IBD.

The genetic divergence values for the produced maps were scaled to values ranging from 0-1. My input files contained pairwise genetic distances between individuals for both mtDNA (using the nucleotide substitution model based on the Maximum Composite Likelihood in MEGA 5 (Tamura et al., 2011) and nDNA (co-dominant genotypic distance method as described in Smouse and Peakall, 1999, in the program GenAlEx: Genetic Analysis in Excel, Version 6.4, Peakall and Smouse, 2006), and a MCP (Maximum Convex Polygon) polygon file created for all location points for each of the data types respectively (mtDNA=88 points, nDNA=100points) (calculated in GME) to specify the extent of the sampling region. The resulting genetic divergence landscape was viewed in ArcMap 10 (©ESRI 2011, ArcGIS Desktop: Release 10, Redlands, CA: Environmental Systems Research Institute), where areas with genetic continuity (high genetic similarity between adjoining samples) where indicated in blue, while areas with genetic discontinuities (high genetic dissimilarity between adjoining samples) where indicated in red.

Appendix IX. The geographic distance (bottom tri-matrix) and the pairwise genetic distance (top tri-matix) between the three sub-populations identified in Geneland for the mtDNA data (n=88)

	mtDNA Sub-population 1	**mtDNA Sub-population 2**	**mtDNA Sub-population 3**
mtDNA Sub-population 1	*	0.476	0.882
mtDNA Sub-population 2	131.2km	*	0.444
mtDNA Sub-population 3	122.4km	153.9km	*

Appendix X. The R-scripts for a mantel test comparison of a geographic distance matrix (km) and a genetic distance matrix (pairwise genetic distance) in the freeware statistical program R (ver 2.13.1 © 2011 The R Foundation for Statistical Computing) using the ade4 package (Chessel et al., 2004).

a) Install and load ade4

b) Input geographic distance (kilometres between centroids of sub-populations) matrix

>geo<-matrix (c(X1,X2,X3,X4,X5,X6,X7,X8,X9),ncol=3)

Where X1-X9 annotates distance values:

	A	B	C
A	X1	X2	X3
B	X4	X5	X6
C	X7	X8	X9

c) Input genetic distance (pairwise distance /Fst) matrix (similar input order as geographic distance matrix)

>gen<-matrix(c(X1,X2,X3,X4,X5,X6,X7,X8,X9),ncol=3)

d) Convert matrices to comparable distance matrices

>geodist<-as.dist(geo)

>gendist<-as.dist(gen)

e) Implement the Mantel Test

>mantel.rtest(geodist, gendist, nrepeat=XXX)

(XXX is the number of permutations)