UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Conservation genetics of the suni antelope *Nesotragus moschatus* von Dueben, 1846

by

Monique Swanepoel

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae

in the Faculty of Natural Agricultural Sciences

Department of Biochemistry, Genetics and Microbiology

University of Pretoria

Pretoria

South Africa

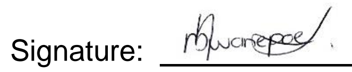
12 August 2019

Supervisor: Prof Paulette Bloomer

Co-supervisor: Mr Arrie Klopper

Declaration

I, Monique Swanepoel, declare that this dissertation, which I hereby submit for the degree Magister Scientiae at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature:  Date: 08/08/2019

Ethics statement

I, Monique Swanepoel, have obtained, for the research performed in this study, the applicable research ethics approval from the Animal Ethics Committee of the University of Pretoria (code: EC040-16).

Conservation genetics of the suni antelope *Nesotragus moschatus* von Dueben, 1846

by

Monique Swanepoel

Supervisors: Prof Paulette Bloomer

Molecular Ecology and Evolution Programme

Department of Biochemistry, Genetics and Microbiology

University of Pretoria

Mr Arrie Klopper

Molecular Ecology and Evolution Programme

Department of Biochemistry, Genetics and Microbiology

University of Pretoria

Department of Biochemistry, Genetics and Microbiology

Magister Scientiae

Dissertation summary

The suni (*Nesotragus moschatus*) is a very small and secretive animal distributed in the eastern parts of Africa, from Kenya south to St. Lucia in South Africa. This antelope can also be found on small islands off the Tanzanian coast. The species is a habitat specialist associated with dense woodland, including the fragmented sand forests of southern Africa. To date six different subspecies have been described based on phenotypical features as well as distribution; since no genetic evidence was considered, there is uncertainty regarding the validity of these taxa.

From phylogenetic studies performed to date, *Nesotragus batesi* was identified as the closest relative to the suni and they appear to be distantly related to the impala (*Aepyceros melampus*). A karyological study performed on captive suni individuals found cryptic variations and several cytotypes. Hybridization between the subspecies (*N. m. akeleyi* and *N. m. zuluensis*) could have

contributed to the variety of chromosome numbers, as well as the decrease in fertility and increase in perinatal mortality. In addition to taxonomic uncertainties, the status and behavioural ecology from different parts of the species' range need to be investigated.

In my MSc I focussed on assessing the phylogenetic relationship between some of the suni subspecies as well as the degree of connectivity between fragmented woodland and forest patches (with emphasis on South Africa and southern Mozambique). A total of 64 samples (dung and tissue) from South Africa (Tembe Elephant Park (TEM), Ndumo Game Reserve (NDU), Phinda Private Game Reserve (PHI) and Tshanini Nature Reserve (TSH)), Mozambique and East Africa (Mnemba Island) was analysed. With the use of phylogenetic analysis of the mitochondrial cytochrome *b* gene and population analyses based on five microsatellite markers, I was able to identify two Evolutionarily Significant Units in southern and eastern Africa, as well as two possible Management Units in southern Africa. In addition to this, some degree of structuring was identified for the South African localities, however, finer scale resolution is needed.

Within this dissertation I allocated one chapter solely for genetic non-invasive sampling, in which I have shown that DNA extraction, PCR amplification and genotyping of highly degraded dung samples are possible. To this end, I made use of 95 degraded suni dung samples to compare two extraction kits (NucleoSpin DNA Stool kit, Macherey-Nagel and QIAamp DNA mini stool kit, Qiagen), of which the NucleoSpin DNA Stool kit performed the best. With this I was able to develop guidelines for optimizing DNA extraction and PCR amplification specifically for highly degraded small dung samples. In addition, I identified and optimized 16 microsatellite markers for future use in finer ecological-scale suni research.

Preface

This dissertation presents the results obtained in my MSc which was a population genetics study focussed on the suni, *Nesotragus moschatus*, found in the eastern parts of Africa. In this dissertation I made use of non-invasive sampling (dung) as well as population genetics approaches to investigate the genetic diversity and population structure of suni in South Africa, Mozambique and Mnemba Island. This dissertation includes a literature review, two research chapters and a short concluding chapter summarising the main findings in this research as well as some future research questions generated by these findings. The three main chapters were written as stand-alone manuscripts and contain some repetition. The referencing style used in this dissertation follows the journal *Conservation Genetics* and a single reference list is presented at the end of the document.

Chapter 1 represents a literature review which focusses on various topics in conservation genetics and genomics, species, conservation units and population processes. I furthermore discuss some current background available on the suni including its taxonomy, subspecies, distribution, habitat and conservation status. I also highlight some of the different studies performed to date on this species and mention some of the topics which need further investigation. Thereafter I briefly discuss the approaches used in the two research chapters and provide some background pertaining to non-invasive sampling. Lastly, I introduce the research questions addressed in the present study.

Chapter 2 In this chapter I investigated the use of genetic non-invasive sampling with the focus on highly degraded, small dung samples. I made use of two dung extraction kits, the NucleoSpin DNA Stool kit (Macherey-Nagel) and the QIAamp DNA mini stool kit (Qiagen), to compare the success of DNA extraction and subsequent molecular analyses. With the use of suni dung samples, I was able to identify the best extraction kit (NucleoSpin DNA Stool kit, Macherey-Nagel). Besides this I adapted its protocol and obtained better quality DNA. Thereafter, microsatellite markers for the suni antelope were meticulously selected and optimised. I was thus able to develop guidelines for optimizing PCR amplification specifically for highly degraded dung samples from a small mammal.

Chapter 3 reported the first population genetics study performed on the secretive suni found in the eastern parts of Africa, including areas such as South Africa, Mozambique and Mnemba Island. I made use of genetic non-invasive (dung) sampling methods along with conservation genetics approaches to study some of the suni subspecies. With the use of the mitochondrial cytochrome *b* gene sequences and microsatellite genotypes, phylogenetic relationships, haplotypes, connectivity and population structure were assessed. I was able to identify two Evolutionarily Significant Units in eastern and southern Africa, as well as two possible Management Units in southern Africa.

In **Chapter 4** I highlight some of the main findings of my MSc study. I also discuss some of the key future questions that were raised from this study and how they can contribute to future conservation.

Acknowledgements

First, I would like to give the greatest of thanks to the Lord, without whom this project would have not been possible.

I would like to thank my supervisors, Prof Paulette Bloomer and Mr Arrie Klopper for all their support and guidance throughout the years. I am absolutely privileged to have had the opportunity to learn from you. You have broadened my knowledge and contributed to the scientist I am today. Thank you. I would also like to thank my committee members Mr Jonathan Bloomer (Centurion Academy) and Dr Irene Barnes (FABI and BGM) for advice on my project proposal and Mr. Bloomer for commenting on a draft of the dissertation. I wish to acknowledge the logistical, co-funding and conceptualisation contributions of Mr. Bloomer and Mr Danie van Eeden of Centurion Academy during the pilot phases of suni research that preceded my study.

For funding, I would like to thank the National Research Foundation (NRF), the University of Pretoria, my supervisors and Mr Brian Randall for financially supporting me.

I would also like to thank every person that was involved in the field work of this project. To each and every nature reserve that was part of this study, thank you for welcoming and allowing us to do our research in your reserve. To Dr Catharine Hanekom and Mr Leonard Muller from Tembe Elephant Park, thank you for sacrificing your time and expertise during the field work. If not for you it wouldn't have been possible; you literally kept us alive. Thank you.

To the Veterinary Genetic Laboratory (VGL) at Onderstepoort, thank you for allowing me to use your facility in the first few testing stages of the microsatellite genotyping.

I would also like to thank the following people for offering up their time and expertise during my MSc: Axel Hunnicutt for sample collection and sharing his knowledge and expertise on the suni; Martin van Rooyen for helping with field work and motivating me to achieve my best; Deon de Jager for helping with field work as well as assisting me with data analysis and interpretation; Anton Klingenberg and Claire Lenahan for teaching me how to use Geneland; Renate and Gladys from the sequencing facility at UP, thank you for always being willing to help, as well as supporting and encouraging me when I felt down, you always gave me hope.

To my fellow scientist friends Diego Castillo, Ané Pieters, Martin van Rooyen and Samantha Mynhardt, thank you for always being there for me. You were some of my biggest supporters and without your encouragement and advice I would have never achieved this. Thank you for listening to all my complaints, I know it couldn't have been easy.

To my mom, dad, sisters and brother-in-law thank you for supporting me for the past few years. Your unconditional love and support meant the world to me. Thank you for all your hugs, kisses, messages and phone calls, those were the most cherished. This achievement is as much yours as it is mine.

To Ivy and Yolandi thank you for all the laughs and giggles you gave me when I was most stressed, it was much needed, thank you.

Lastly, to Brendan, thank you for motivating me when I was down, believing in me when I didn’t and pushing me when I couldn’t. Your tough love and support is what kept me going. Thank you for always listening and never giving up on me, if it weren’t for you, I would have given up a long time ago.

Table of contents

Declaration ii
Ethics statement iii
Dissertation summary iv
Preface vi
Acknowledgements vii
Table of contents ix
List of figures xii
List of tables xv
Chapter 1 Background and biology of the suni (*Nesotragus moschatus*) 1
Abstract 2
Introduction 2
Conservation genetics and genomics 3
Species and conservation units 6
What is a species? 6
Conservation management units 6
Genetic diversity and its importance in conservation genetics 8
Population processes 10
Connectivity 12
Population structure 13
Background on suni (*Nesotragus moschatus*) 14
Phylogeny 14
Subspecies 16
Distribution 17
Description 17
Suni habitat and niche 18
Age of South African and Mnemba Island suni 20
Suni diet 21
Population size and conservation status 21
Behavioural ecology 22
Investigative tools 23
Microsatellite markers 23
Genetic non-invasive sampling and its applications 24
Analytical approaches and software 27
Research questions and objectives 27
Chapter 2 Optimization of markers for DNA sequencing and microsatellite genotyping of suni DNA extracted from dung, blood and tissue samples 28
Abstract 29
Introduction 29

Materials and methods 31
Sample collection 31
Sample preparation 33
DNA extraction from dung samples 34
DNA extraction for tissue samples 37
Species identification using mtDNA 37
Microsatellite marker selection 38
Microsatellite- and sex marker optimization 39
Results and discussion 43
Sample quality 43
DNA extractions 45
Species identification 47
Marker selection 48
Marker optimisation 49
Conclusion 52
Supplementary material 53
Chapter 3 Conservation genetics of the suni antelope *Nesotragus moschatus* von Dueben, 1846 62
Abstract 63
Introduction 63
Materials and methods 65
Sample collection and selection 65
DNA extractions and species identification 66
Phylogenetic relationships and haplotype network 67
PCR optimization and amplification of microsatellites and one sex marker 69
Scoring, binning and comparing microsatellites 69
Marker selection and power 71
Genetic diversity and statistics 71
Genetic structure and connectivity 71
Identifying ESUs and MUs 72
Results 72
DNA extractions and species identification 72
Cytochrome *b* analysis 73
Microsatellite marker optimization and amplification 75
Microsatellite validation, genotyping and comparison 75
Marker selection and power analysis 76
Genetic diversity and structuring of suni samples over all populations 78
Spatial analysis and genetic structure of the South African localities 80
Discussion 82
Implications of using dung DNA analysis 82

Microsatellite and sexing markers 83
Genetic diversity of the suni 84
ESUs and MUs in the suni antelope 84
Degree of connectivity across South African suni localities 86
Conclusion 87
Supplementary data 88
Chapter 4 Concluding remarks 99
References 102

List of figures

Fig. 1.1 Diagram illustrating the different relationships considered in the a) conservation genetics (white) and b) conservation genomics approach (grey). Adapted from Ouborg et al. (2010) 4
Fig. 1.2 Flow diagram representing interacting factors which can be found in the conservation of natural populations. White blocks indicate interacting factors which can be identified by traditional methods and the grey blocks indicate factors determined with conservation genomic approaches. Adapted from Allendorf et al. (2010).. 5
Fig. 1.3 Diagram illustrating a theoretical allele distribution among four populations indicating both the allele frequencies and phylogenetic relationships present relative to the described definitions of Evolutionarily Significant Units (ESU) and Management Units (MU). Within this tree populations 1 and 2 do have similar allele frequencies and form part of the same MU. Whereas with the 3rd and 4th population fixed differences are present when compared to populations 1 and 2. Nevertheless only population 4 shows phylogenetic separation. Therefore, population 1, 2 and 3 forms part of the same ESU, whereas population 4 can be considered a separate ESU. Reprinted from Moritz (1994a) with permission. 7
Fig. 1.4 Population structure along a continuum ranging from random mating to complete independence between sub-populations. Each group of green patches represents a group of sub-populations with different degrees of connectivity: a) Highly structured populations; b) populations with modest connectivity; c) substantial connectivity; and d) panmixia (random mating). The rectangle indicates the hypothesized population structure of suni between forest patches. Adapted from Waples and Gaggiotti (2006) 13
Fig. 1.5 Parsimony tree based on the mtDNA cyt *b* gene of 38 bovid taxa. Suni is clustered as a distant sister taxon of the impala. Bootstrap values are indicated at nodes. Reprinted from Matthee and Robinson (1999) with permission. 15
Fig. 1.6 Combined phylogenetic trees based on nuclear genes and mtDNA of 34 bovid taxa. It is suggested that the suni is a unique lineage and cannot be clustered within any of the subfamilies. Bootstrap values are presented at the nodes, with numbers above the diagonal branches representing parsimony values; those below the branches indicate neighbor-joining and maximum likelihood support. Subfamilies are represented as follow: black square = Bovinae; # = Hippotraginae; black circles = Alcelaphinae; white squares = Caprinae; * = Reduncinae; O = Antilopinae; diamond = Cephalophinae; ? = uncertain status. Reprinted from Matthee and Davis (2001) with permission. 15
Fig. 1.7 Distribution map of the six described suni (*Nesotragus moschatus*) subspecies (Ansell 1971; Skinner and Chimimba 2005; Groves and Grubb 2011; Frost and Carnaby 2015). 16
Fig. 1.8 Image of a suni antelope (*Nesotragus moschatus*) (Megan Loftie-Eaton, http://vmus.adu.org.za/) 18
Fig. 1.9 The distribution of different biomes across East and Southern Africa. Colours correspond to the different biomes. The dark blue line represents the boundaries of the suni distribution (Olson et al. 2001)... 19
Fig. 1.10 a) Sand forest in Tshanini Nature Reserve (JP Bloomer) and b) grassland borders found surrounding the forest patches in Tembe Elephant Park (M Swanepoel) 20
Fig. 2.1 Distribution map of suni subspecies and sampling localities. The subspecies shown in bold are the four currently recognised subspecies (Ansell 1971; Skinner and Chimimba 2005; Groves and Grubb 2011; Frost and Carnaby 2015)... 32
Fig. 2.2 Comparison of different dung samples, a) suni dung pellets and b) red duiker dung pellets 33
Fig. 2.3 Extraction protocol using the QIAamp DNA mini stool kit from Qiagen. Adapted from Costa et al. (2017)... 36
Fig. 2.4 Designed multiplex panel for 16 dinucleotide microsatellite markers and one sexing marker (Amelogenin) using Multiplex Manager v.1.2... 41
Fig. 2.5 Agarose gels representing DNA extractions of samples from three different nature reserves: a) PHI b) TEM and c) MNE Island using the Nucleospin® DNA stool kit (Macherey-

Nagel). *- lanes with 100 bp DNA ladder for both the top and bottom sections of each separate gel 44
Fig. 2.6 Agarose gels representing amplicons of the cyt *b* gene from three different nature reserves: TEM (lanes a2-7), PHI (lanes a8-10 and a14) and MNE (lanes b2-6 and b10-16). 100 bp DNA ladder = lanes a1, a13, b1 and b9. Positive control = lanes a11 and b7. Negative controls = lanes a12 and b8. 44
Fig. 2.7 Agarose gel images of DNA extractions and cyt *b* PCR amplifications illustrating a comparison between the NucleoSpin DNA Stool kit (Macherey-Nagel) and the QIAamp DNA mini stool kit (Qiagen). a) DNA extracts from the Qiagen kit, b) DNA extracts from Macherey-Nagel kit, c) PCR amplicons from DNA extracted with Qiagen kit, d) PCR amplicons from DNA extracted with Macherey-Nagel kit. *- lanes with 100 bp DNA ladder, positive (+) and negative (-) control. 46
Fig. 2.8 Scatter plots of PCR success and DNA concentration determined by a) Nanodrop and b) Qubit. 47
Fig. 2.9 a) Agarose gel comparing the PCR amplification of three microsatellite markers performed on tissue samples. Lane 1: 100 bp DNA ladder, lane 2 and 3: BM415 samples, lane 4: BM415 positive control (blood sample), lane 5: BM415 negative control, lane 6 and 7: SRCRSP24 samples, lane 8: SRCRSP24 positive control, lane 9: SRCRSP24 negative control, lane 10: BM757 sample, lane 11: BM757 positive control and lane 12: BM757 negative control. b) an electropherogram of the sample in lane 11. 49
Fig. 2.10 Comparison of PCR amplification of the microsatellite marker, BM2113, in tissue and dung samples. Lanes 1 and 13: 100 bp DNA ladder, lanes 2-7, 10-12 and 14-20: dung samples, lanes 8-9: tissue samples, lane 21: positive control and lane 22: negative control. 50
Fig. 2.11 Electropherograms of the marker INRA6 (two blue peaks between 120-150 bp) in a) a dung and b) blood sample 51
Fig. 2.12 Gradient PCR in dung samples for marker OARCP26 with normal *Taq*. *- lanes with 100 bp DNA ladder, positive (+) and negative (-) control. 51
Fig. 2.13 PCR amplification for markers I206 (top) and INRA6 (bottom) with the use of multiplex mix in dung samples. * - 100 bp DNA ladder, positive (+) and negative (-) control. 52
Fig. 3.1 Google earth images indicating the absence of connectivity between the different sampling localities (red pins) in a) southern Africa and b) South Africa (KZN) 66
Fig. 3.2 Illustration of the two primer pairs used to amplify two separate cyt *b* gene fragments, separated by a non-overlapping region of approximately 259 bp. Primer positions are labelled according to the human mtDNA sequences (Kocher et al. 1989). 68
Fig. 3.3 Phylogenetic tree based on part of the mtDNA cyt *b* gene (1034 bp) in suni individuals from five locations (including NCBI sequences). 73
Fig. 3.4 Haplotype network of cyt *b* sequences in six localities in eastern and southern Africa. The white numbers inside each circle/ block corresponds to the haplotype number shown in Table 3.4, with the size of each circle proportional to the number of individuals displaying that specific haplotype. White block/circles = size scale with black numbers indicating number of individuals. Rectangle = haplotype with the highest frequency in southern Africa. Blue = MOZ, Pink = TEM, Red = TSH, Yellow = NDU, Purple = PHI and Green = MNE. 74
Fig. 3.5 Statistical power comparison of five microsatellite loci to detect a predefined level of genetic differentiation (F_{ST}). Dashed line indicates a 95% probability of detecting genetic structure with approximate F_{ST} values of 0.02254 and 0.0334 respectively when using the same or different sample sizes. Dif = Different sample sizes, Same = Same sample sizes, Chi =Chi-square approach and Fisher = Fisher's exact test 77
Fig. 3.6 Principal Coordinates Analysis (PCoA) plot clustering 64 suni individuals from five different localities based on genetic distances between genotypes at five microsatellite loci 79
Fig. 3.7 STRUCTURE results depicting individual-based clustering of 64 suni individuals from five localities based on genotypes at five microsatellite loci for $K = 2$, $K = 3$ and $K = 4$ clusters. Each vertical bar along the X-axis represents an individual and the Y-axis is the probability of assignment (q) to K clusters. Location was not specified as prior in the analysis. Localities are arranged from north to south. 79

Fig. 3.8 Principal Coordinates Analysis (PCoA) plot clustering 34 suni individuals from three different localities in South Africa based on genetic distances between genotypes at five microsatellite loci 80
Fig. 3.9 STRUCTURE results depicting individual-based clustering of 34 suni individuals from three localities in South Africa based on genotypes at five microsatellite loci for $K = 2$ and $K = 3$ clusters. Each vertical bar along the X-axis represents an individual and the Y-axis is the probability of assignment (*q*) to *K* clusters. Location was not specified as prior in the analysis. Localities are arranged from north to south 81
Fig. 3.10 Maps of posterior probabilities of population membership at $K = 3$, based on 34 sunis' microsatellite genotypes at five loci from three South African localities. The black dots represent the different sampling localities while the white-yellow colour indicates a high probability of population membership and the orange-red colour indicates a low probability of population membership. Contour lines indicate posterior probability. 81
Fig. 3.11 Spatial autocorrelation analysis in suni populations from South Africa. The correlogram indicates the correlation coefficient (*r*) at various distance classes (km). Dashed lines U and L indicate the upper and lower confidence limits. 82

List of tables

Table 1.1 Evolutionarily Significant Unit (ESU) criteria since Ryder (1986), adapted from Fraser and Bernatchez (2001)........ 8
Table 1.2 A representation of different population definitions, adapted from Waples and Gaggiotti (2006) 11
Table 2.1 Tissue and dung samples collected from different locations used during extraction protocol testing 32
Table 2.2 Primers used for mtDNA cytochrome *b* gene amplification of *Nesotragus moschatus*... 37
Table 2.3 Final set of dinucleotide microsatellite markers and one sexing marker (Amelogenin) used in this study 40
Table 2.4 Comparison of genotyping success over different nature reserves, sample types and freshness using eight microsatellite loci. 44
Table 2.5 Comparison of Nanodrop values, Qubit values and PCR amplification success........ 47
Table 2.6 Comparison of amplification success in tissue samples for markers amplified in single reactions as well as multiplex reactions. A total of 14 microsatellite markers were evaluated. 49
Table 3.1 Tissue and dung samples used for sequencing (mtDNA cytochrome *b*) and microsatellite genotyping over six different localities in eastern and southern Africa........ 66
Table 3.2 Two primer pairs used for the amplification of separate cyt *b* fragments in suni 68
Table 3.3 Final set of 16 dinucleotide microsatellite markers and one sexing marker used in this study. Grey = Optimised and remaining markers 70
Table 3.4 Different cyt *b* haplotypes present in six localities in southern and eastern Africa. *n* = Number of individuals with a specific haplotype. 74
Table 3.5 Diversity indices generated by Cervus for 16 selected individuals from all different localities........ 76
Table 3.6 Comparison of number of alleles found in suni (this study) and combined data of other previously studied ungulates (Miller et al. 2016; Van Wyk et al. 2018). * = not available, Imp = Impala, Roa = Roan, Buf = Buffalo, Sab = Sable, Gnu = Wildebeest, Ela = Eland and Gem = Gemsbok. 76
Table 3.7 Null allele frequencies generated for all genotyped individuals with the Oosterhout test in Micro-checker. (NS = Not Significant, ND = Not Determined, * P<0.01, ** P<0.001) 77
Table 3.8 Pairwise F_{ST} and corresponding *Nm* values between different suni localities based on five microsatellite loci. F_{ST} and *Nm* values are shown above and below the diagonal, respectively. *Nm* values above 1 are shown in bold........ 78
Table 3.9 Summary statistics of five localities based on five microsatellite loci generated in Genepop, GenALex and HP-RARE 78

Chapter 1 Background and biology of the suni (*Nesotragus moschatus*)

Abstract

The suni is a very small and secretive antelope found in the eastern parts of Africa. Many different subspecies have been described however, currently only four are formally recognised. In this literature review I discuss several topics covering conservation genetics, genomics, species and conservation units. Additionally, I discuss details regarding genetic diversity and how it could enable one to study connectivity between populations and identify population structure. Subsequently, I focus on the background of the suni, including its phylogeny, conservation status, subspecies, distribution and previous studies performed. Finally, I dedicated a small section discussing the different approaches (software, genetic markers and material) that were used in this study, including a detailed discussion on non-invasive sampling (dung) as well as its advantages and disadvantages.

Introduction

Extinction rates are on the rise and the impact of human activities is a major contributor (Festa-Bianchet and Apollonio 2003). Species are overexploited and in some cases their habitats are being destroyed and decreased thus having an overall negative impact on their future persistence (Ehrlich 1991; Caughley 1994). Conservation biology has been a growing field since the 1980s (Festa-Bianchet and Apollonio 2003). Although the goals for conservation are very broad, these include the preservation of genetic diversity in species to ensure they maintain the ability to adapt and persist if environmental changes do occur (Festa-Bianchet and Apollonio 2003). Over the past decade climate change has been a prominent field of interest, amongst others raising concern over coastal regions (McMichael et al. 2006). With the predicted rise in temperature, sea levels could increase by more than a metre, impacting many coastal communities and habitats (McMichael et al. 2006; Collier et al. 2008). Coastal eastern Africa is also a concern, as this area contains highly diverse and threatened forest ecoregions (WWF). Maputaland, in the northeast of KwaZulu-Natal (KZN) in South Africa is part of this coastal belt; its unique sand forests are young but highly biodiverse (Kirkwood and Midgley 1999). One of its often-overlooked mammalian habitat specialists, is the suni (*Nesotragus moschatus* (von Dueben, 1846)).

The suni is a very shy, secretive animal found in enclosed dense woodlands of East Africa and unique sand forests of southern Africa (Skinner and Chimimba 2005; Frost and Carnaby 2015). This species is spread over the eastern part of Africa and more specifically distributed in a long stretch from the eastern part of KZN north to the south-eastern part of Kenya (Skinner and Chimimba 2005; Frost and Carnaby 2015). Several subspecies along with their distributions have been described however, all were based on phenotypical features with no genetic information taken into account (Ansell 1971; Kingdon and Largen 1997; Frost and Carnaby 2015). In addition to this the current status of this species has been listed as of Least Concern on the IUCN Red List (IUCN SSC Antelope Specialist Group 2016). Nevertheless, the Red List of Mammals of South Africa, Lesotho and Swaziland has listed the species as Endangered (Hunnicutt et al. 2016), highlighting the need for more local research on the species.

Coupled with this, current genetic information available for this species is very limited. The few studies performed include karyotypic differences of two of the subspecies and phylogenetic relationships of the species (Kingswood et al. 1998; Matthee and Robinson 1999). According to phylogenetic studies unexpected information came to light. These studies showed that *N. batesi* was *N. moschatus'* closest relative however, both of them were clustered as distant sister taxa of the impala (*Aepyceros melampus*), separate from other *Neotragini* (Matthee and Robinson 1999; Bärmann and Schikora 2014). The latter authors suggested that *N. batesi* and *N. moschatus* should be recognized in the genus *Nesotragus*, as they are not monophyletic within *Neotragus*. The study by Kingswood et al. (1998) was performed on captive individuals in which cryptic genetic variations between subspecies were identified (Kingswood et al. 1998). This study revealed a variety of diploid numbers, 2n = 52-56 in the suni (Kingswood et al. 1998). They showed that the diploid numbers 2n = 52 and 2n = 56 belonged to *N. m. akeleyi* and *N. m. zuluensis* respectively (Kingswood et al. 1998). The intermediate karyotype numbers were likely due to hybridization occurring between these subspecies. Gathering additional conservation genetics data on suni, including assessing genetic diversity across different spatial scales, would add to our understanding of the relationships within the species and connectivity between isolated habitat patches. These data would be valuable in formulation of conservation plans for the species.

In this literature review I cover various topics such as conservation genetics and genomics, species and conservation units. In addition to this I discuss the role of genetic diversity and how it could be used to identify connectivity and population structure. Thereafter, I focus on the background of the species, covering phylogenetic relationships, subspecies, distribution, habitat and previous studies conducted on the species. Finally, I discuss the investigative tools (software, genetic markers and material) used in this study, including a thorough discussion on non-invasive sampling (dung), the advantages as well as the problems which could be experienced.

Conservation genetics and genomics

Conservation genetics as a discipline developed as part of the growing interest in conservation biology since the 1970s (Frankham 2010a, b). The main aim of this field was and remains to understand and determine the underlying processes that cause the extinction of species (Hedrick and Hurt 2012). Conservation genetics contributes to both the evolutionary and ecological scale within conservation biology and thus adds to the development of more effective management plans (Hedrick and Hurt 2012). In the 1990s the IUCN and the Convention on Biological Diversity (CBD, 1992) formally recognized the importance of conservation of biological diversity and suggested that biodiversity should be considered at three different levels: genes, species and ecosystems (McNeely et al. 1990). Since these early periods, the field has contributed to the conservation of many species (captive, endangered and wild populations), as well as highlighted several key issues such as taxonomy, conservation units and inbreeding.

Even though conservation genetics is still relevant and useful in studies, there are some limitations to this approach, one of which is the number of neutral markers that can be used (Ouborg et al. 2010). This limitation has led to unresolved questions, e.g. what are the mechanisms responsible for the link between neutral and adaptive genetic variation? Or, what is the correlation between population size and the level on non-neutral genetic variation (Frankham 2010b; Ouborg et al. 2010)? Due to the advent of whole genome sequencing and genotyping-by-sequencing (Ouborg et al. 2010) conservation genetics has now transitioned to conservation genomics. This field of study is still expensive and not necessary or appropriate in all cases, but it is believed that it will become applicable for answering future research questions since more markers can be used and sequencing costs are decreasing (Allendorf et al. 2010; Frankham 2010a).

In the paragraphs above, brief descriptions were given regarding conservation genetics and genomics. As mentioned earlier there are two different approaches towards addressing genetic questions in conservation. Conservation genetics mainly consists of three major goals: The first of which is to ensure that populations' genetic diversity remains high and that inbreeding occurs at very low levels (Woodruff 2001; Frankham et al. 2002). This takes into account that genetic drift and inbreeding are potential threats in small populations. Therefore the conservation genetics approach (Fig. 1.1a) focuses on the relationships between population size and neutral variation, keeping in mind that the relationship with fitness is based on the assumption that neutral marker variation is indicative of selectively important variation (Ouborg et al. 2010). Secondly, to investigate any uncertainties in the taxonomy of particular species and in such a manner identify appropriate conservation management units; this is discussed below (Frankham et al. 2002; Frankham 2010a). Thirdly using molecular genetic techniques and approaches to understand species biology (Frankham et al. 2002; Frankham 2010a).

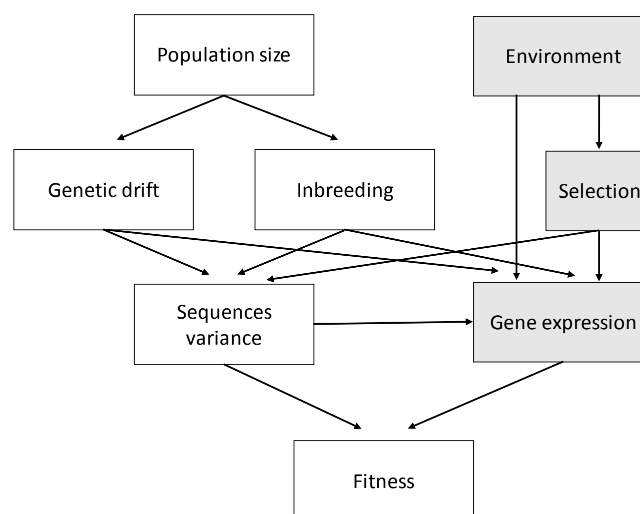


Fig. 1.1 Diagram illustrating the different relationships considered in the a) conservation genetics (white) and b) conservation genomics approach (grey). Adapted from Ouborg et al. (2010)

The conservation genomics approach (Fig. 1.1b) on the other hand evaluates the relationships between population size and both neutral and selective variation (Ouborg et al. 2010) making use of sequence variation and how this affects expression, and thereby incorporating potential selection effects that could occur (Allendorf et al. 2010; Ouborg et al. 2010). Additionally conservation genomics allows the assessment of genotype-phenotype-environment interactions (Allendorf et al. 2010; Ouborg et al. 2010). One of the advantages of conservation genomics is the fact that more markers can be analysed at the same time allowing coverage of the entire genome (Allendorf et al. 2010; Ouborg et al. 2010). In doing so more parameters (such as effective population size, gene flow, introgression) can be estimated with precision and less bias, making the estimates more reliable (Frankham 2010a, b; Ouborg et al. 2010).

Other advantages of conservation genomics include opportunities to study multi-locus effects, such as selective sweeps and distinguishing between neutral and non-neutral markers (Frankham 2010a, b; Ouborg et al. 2010). Loci subjected to natural selection and involved in speciation can also be identified (Frankham 2010a, b). Coupled with this, genes which are functionally important such as those responsible for environmental and genetic stress, as well as genes and pathways involved in adaption and inbreeding can also be identified with conservation genomics approaches (Allendorf et al. 2010; Ouborg et al. 2010). The techniques such as marker-based genotyping, reduced representation sequencing and whole genome sequencing make conservation genomics more applicable than conservation genetics (Allendorf et al. 2010). Figure 1.2 depicts a summary of the different factors that can be addressed between the two approaches (Allendorf et al. 2010).

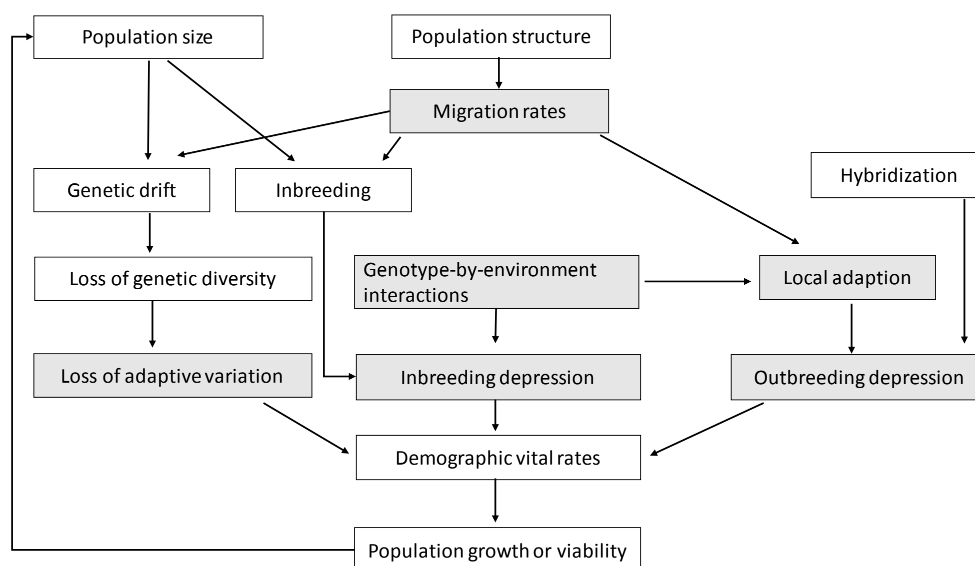


Fig. 1.2 Flow diagram representing interacting factors which can be found in the conservation of natural populations. White blocks indicate interacting factors which can be identified by traditional methods and the grey blocks indicate factors determined with conservation genomic approaches. Adapted from Allendorf et al. (2010)

By knowing the current status of a species it can enable conservation managers to use the information gathered and design management strategies to reduce extinction risk (Frankham 2010a,

b). In the event where no investigation has been done regarding a species' population status, concerning genetic variation, there is no way in saying whether or not a species is a viable population or that it will exist in future generations. Thus, by studying conservation genetics we can predict (albeit not with complete accuracy) future outcomes of populations and species (Woodruff 2001; DeSalle and Amato 2004; Frankham 2010b).

Species and conservation units

What is a species?

According to Woodruff species can be described as fundamental units of evolution and taxonomic classification (Woodruff 2001). Even though it might seem like a very simple concept this is not the case. For many years there has been a big debate regarding how to define a species, which in the end resulted in many different species concepts arising (De Queiroz 2007). Many authors opt to use a combination of principles underlying different species concepts, to define species (De Queiroz 2007). In addition to this, species can now also be distinguished by using a multi-gene approach i.e. using multiple loci or different genes, which has increased species delimitation success (Dupuis et al. 2012). However, due to using a single mitochondrial DNA (mtDNA) gene, for this study I adopted the Phylogenetic Species Concept (PSC) (De Queiroz 2007). The PSC defines a species as an irreducible group of members which is monophyletic, contains exclusive coalescence alleles and who all possess shared qualitative, fixed differences (De Queiroz 2007).

Conservation management units

At a meeting of the Zoological Society of Philadelphia in 1985 problems with subspecies definitions in the context of conservation were highlighted. Subsequently, Ryder (1986) provided an overview of the key questions raised and introduced the concept of Evolutionarily Significant Units (ESUs) as a more appropriate unit for conservation below the species level. The questions which were raised were: Which gene pools should be preserved?; Which species should become the focal point of concern and which could be neglected?; How many individuals are needed in order for the population to maintain enough genetic diversity over a period of time?; Which subspecies represented populations which possessed genetic attributes important for current and future generations (Ryder 1986)?

Ryder described ESUs as "subsets of the more inclusive entity, which possess genetic attributes significant for the present and future generations of the species" (Ryder 1986). Moritz subsequently advocated that conservation was about how to measure genetic diversity, identify evolutionarily divergent units and to assess conservation value of populations or areas (Moritz 1994a), hence emphasizing historical population structure and mtDNA phylogeny (Moritz 1994b). However, prioritising management just based on mtDNA diversity was inappropriate (Moritz 1994a). Consequently Moritz suggested quantitative criteria for ESUs, in that they needed to show reciprocal monophyly for mtDNA haplotypes, as well as significant allele frequency differences at nuclear loci (Moritz 1994a). Crandall et al. (2000) on the other hand were completely against conservation units

and mentioned that the ESU concept should rather be abandoned(Crandall et al. 2000). Instead they suggested that a more holistic concept of species, consisting of populations with varying levels of gene flow due to genetic drift and natural selection should rather be used (Crandall et al. 2000).

In addition to ESUs Moritz suggested adding another classification unit called a Management Unit (MU), below ESUs (Fig. 1.3). This was done so that populations that were demographically distinct could be identified and managed in such a manner so that it could ensure the viability of a larger ESU. All of which were based on the fact that significant differences could be present in allele frequencies, in both nuclear and mitochondrial loci, irrespective of what phylogenetic analysis showed (Moritz 2002).

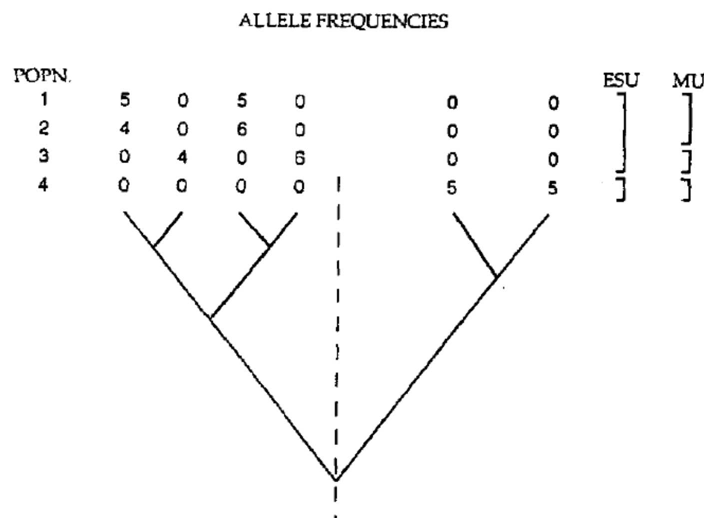


Fig. 1.3 Diagram illustrating a theoretical allele distribution among four populations indicating both the allele frequencies and phylogenetic relationships present relative to the described definitions of Evolutionarily Significant Units (ESU) and Management Units (MU). Within this tree populations 1 and 2 do have similar allele frequencies and form part of the same MU. Whereas with the 3rd and 4th population fixed differences are present when compared to populations 1 and 2. Nevertheless only population 4 shows phylogenetic separation. Therefore, population 1, 2 and 3 forms part of the same ESU, whereas population 4 can be considered a separate ESU. Reprinted from Moritz (1994a) with permission.

Although these definitions created a large debate for the past three decades, all the authors who created these definitions (Table 1.1) were trying to achieve the same goal: identify sections of species whose divergence can be determined by placing different emphasis on the evolutionary forces that played a role at different times (Fraser and Bernatchez 2001). In order to reach common ground between all the new definitions regarding the ESU concept, Fraser and Bernatchez (2001) came to the conclusion that every definition that was created (Table 1.1) possessed strengths which are counterbalanced to some extent by weaknesses. Thus, it is important to note that the approaches may work in varying situations much better than others. Therefore ESUs should be evaluated on a case-by-case basis, rather than attempting to use one universal definition. In their opinion it is better to combine these different definitions than to create more and not get to the end goal which is conserving the species (Fraser and Bernatchez 2001). Even though I concur with the latter authors,

for the purposes of assessing diversity in the suni, I adopted Moritz's quantitative criteria as a precautionary first step towards identifying conservation units in the species.

Table 1.1 Evolutionarily Significant Unit (ESU) criteria since Ryder (1986), adapted from Fraser and Bernatchez (2001).

Author(s)	ESU criteria
(Ryder 1986)	"Subsets of the more inclusive entity species, which possess genetic attributes significant for the present and future generations of the species in question"
(Waples 1991)	"A population or group of populations that: (i) is substantially reproductively isolated from other conspecific population units; and (ii) represents an important component of the evolutionary legacy of the species"
(Dizon et al. 1992)	"Populations or groups of populations demonstrating significant divergence in allele frequencies"
(Avise and Moritz 1994)	"Sets of populations derived from consistently congruent gene phylogenies"
(Moritz 1994b)	"Populations that: (i) are reciprocally monophyletic for mtDNA alleles; and (ii) demonstrate significant divergence of allele frequencies at nuclear loci"
(Vogler and Desalle 1994)	"Groups that are diagnosed by characters which cluster individuals or populations to the exclusion of other such clusters"
(Crandall et al. 2000)	"Abandon term ESU for more holistic concept of species, consisting of populations with varying levels of gene flow evolving through drift and selection"
(Fraser and Bernatchez 2001)	"A lineage demonstrating highly restricted gene flow from other such lineages within the higher organizational level (lineage) of the species"

Genetic diversity and its importance in conservation genetics

One of the factors that allows us to answer so many questions regarding species is genetic diversity/variation. To elaborate, genetic variation could for example provide information regarding the health of a species or in some cases help delineate different species or lineages. According to Lacy (1997), genetic diversity can be classified at two different levels: individual and population level. At individual level genetic diversity can be determined by the percentage of loci which are heterozygous (Lacy 1997). Whereas at the population level genetic diversity is defined by either the

number of distinct alleles per locus or by the percentage of loci that are polymorphic (Lacy 1997). Due to high throughput sequencing and genotyping more information on genetic variation can now be gathered in less time (Allendorf et al. 2010). According to Schlötterer (2004) there are several methods that can be used to evaluate variation but the ones that stand out from the rest are DNA sequencing, single nucleotide polymorphisms (SNPs) and microsatellite genotyping. Highly variable markers such as microsatellites and SNPs are preferred due to their utility in assessing individual and population level variation (Bruford and Wayne 1993; Brumfield et al. 2003; Zhang and Hewitt 2003; Selkoe and Toonen 2006).

Ever since genetic diversity was identified as a core factor in conservation, it has been utilised a lot more, like in cases where populations had to be evaluated (Woodruff 2001; Frankham 2010a). In addition to this, a positive correlation between genetic variation and population viability have also been identified (Lacy 1997; Woodruff 2001; Keller and Waller 2002). This connection between genetic diversity and viability is really important for conservation, the reason being that if viability of a population/species is very low it could be used as an indicator showing that the population/species is close to extinction (Lacy 1997). The question can then be raised: how does genetic variation influence the viability of a population?

The way in which genetic variation contributes to the viability of a population is through genetic erosion especially in small populations (Lacy 1997; Woodruff 2001; Ouborg et al. 2010). Genetic erosion refers to the decrease in genetic variation due to processes such as genetic drift and inbreeding (Woodruff 2001). Genetic drift can be defined as the random fluctuations in allele frequencies over time which can lead to the loss of adaptive alleles and the fixation of deleterious alleles (Lacy 1997; Ouborg et al. 2010). Inbreeding on the other hand refers to the process in which close relatives breed with each other causing an increase in the homozygotes of the population which in the end reduces the overall genetic diversity (Pusey and Wolf 1996; Keller and Waller 2002; Ouborg et al. 2010).

Moreover, these two processes are quite common and can be found in both captive and wild animals. With all things considered there is no mammal that have not shown inbreeding occurring (Lacy 1997). In many cases of inbreeding it has been found that low variation can cause the fitness of individuals to decrease hence, affecting the viability of the population (Lacy 1997; Frankham et al. 2001; Keller and Waller 2002). These events are observed more often in small populations especially when they are isolated with no gene flow causing weak alleles to become fixed through genetic drift (Keller and Waller 2002). In addition, inbreeding is associated with a reduction in resistance to pathogens and parasites (Lacy 1997). When genetic diversity decreases the ability of the population to adapt to environmental change also decreases (Lacy 1997; Keller and Waller 2002). The reason being that if genetic diversity decreased due to inbreeding, it would have meant that most of the alleles were lost, creating the opportunity for the remaining alleles to become fixed (Lacy 1997; Woodruff 2001). Once this occurs adaption might not be possible, especially if the allele which was core for adaptation were lost during inbreeding. Not only will the decrease in genetic diversity affect

adaptability, it will also decrease the animals growth rate and fertility (Lacy 1997; Woodruff 2001). More importantly, the longer inbreeding continuously takes place, the more homozygotes will increase causing deleterious alleles to be continuously fixed, resulting in inbreeding depression which manifests as a reduction in viability and fecundity of the population (Pusey and Wolf 1996; Woodruff 2001).

For this reason, determination of inbreeding in populations is essential for conservation genetics since it can help in the management plans of captive, as well as wild animals' persistence. In order to reduce inbreeding managers can make use of purging (Hedrick 1994; García-Dorado 2012) and genetic rescue (Tallmon et al. 2004). With use of genetic rescue new beneficial gene variations are introduced from outside populations for genes which have undesirable variations (Frankham et al. 2011; Hedrick and Garcia-Dorado 2016). Outbreeding on the other hand can be described as the event in which two unrelated individuals are mated together (Woodruff 2001). In an attempt to increase genetic diversity this was applied to different animal species and the results were quite contradicting (Woodruff 2001; Frankham 2010b; Frankham et al. 2011). Some unsuccessful breeding rescue attempts have been the result of outbreeding depression. If two distantly related individuals of the same species are mated, the male and female genomes may be too divergent, causing the offspring to have genetic disorders or being infertile (Woodruff 2001; Frankham et al. 2011). Moreover, it has been proposed that chromosomal differences, adaptive differences present between the two individuals mated, bottlenecks and genetic drift were some of the mechanisms responsible for outbreeding depression occurring (Frankham et al. 2011).

In all the work done pertaining to conservation genetics, creating effective management plans that maintain genetic diversity are vital (Woodruff 2001; Frankham 2010b). By doing this it could increase a population's chance of survival and persistence in the future which is a main concern in biodiversity conservation. However, it is important to note that even though genetic variation is a good indicator of a populations' status, other factors such as environment and social interactions can also play a role and should be taken into consideration.

Population processes

In order to understand what population processes are and how they play a role in conservation genetics it is important to first know what a population is. Over the years many different definitions regarding populations have been described (Table 1.2). Although they may differ in some way they do have some similarities in that they emphasize cohesive processes which unites individuals within a population (Waples and Gaggiotti 2006). Biological definitions of a population can be grouped into two different paradigms: ecological and evolutionary (Andrewartha and Birch 1986). In the evolutionary paradigm the cohesive forces are mainly genetic and focusses on reproduction hence, describing a population as a group of individuals (same species) which co-occurs in space and time having the opportunity to interact with one another (Waples and Gaggiotti 2006). Whereas in the

Table 1.2 A representation of different population definitions, adapted from Waples and Gaggiotti (2006)

Population definitions	Reference
Ecological paradigm	
"A group of organisms of the same species occupying a particular space at a particular time"	(Krebs 1994; Roughgarden et al. 2014)
"A group of individuals of the same species that live together in an area of sufficient size that all requirements for reproduction, survival and migration can be met"	(Huffaker et al. 1984)
"A group of organisms occupying a specific geographical area or biome"	(Lapedes 1978)
"A set of individuals that live in the same habitat patch and therefore interact with each other"	(Hanski 1998)
"A group of individuals sufficiently isolated that immigration does not substantially affect the population dynamics or extinction risk over a 100-year time frame"	(McElhany et al. 2000)
Evolutionary paradigm	
"A community of individuals of a sexually reproducing species within which matings take place"	(Dobzhansky 1970)
"A major part of the environment in which selection takes place"	(Williams 1966)
"A group of interbreeding individuals that exist together in time and space"	(Hedrick 2000)
"A group of conspecific organisms that occupy a more or less well-defined geographical region and exhibit reproductive continuity from generation to generation"	(Futuyma 1998)
"A group of individuals of the same species living close enough together than any member of the group can potentially mate with any other member"	(Hartl and Clark 1989)
Statistical paradigm	
"An aggregate about which we want to draw inference by sampling"	(Snedecor and Cochran 1967)
"The totality of individual observations about which inferences are to be made, existing within a specified sampling area limited in space and time"	(Sokal and Rohlf 1969)

ecological paradigm cohesive forces are largely demographic, placing the emphasis on the co-occurrence of individuals allowing them to interact demographically. Therefore, defining a population as a group of individuals (same species) that lives in a close proximity to one another so that any member of the group can potentially mate with another member (Waples and Gaggiotti 2006).

Scientists study populations by evaluating the interaction between individuals, as well as with the environment (Tarsi and Tuff 2012). This includes demography (Lebreton et al. 1992) which is the study of populations making use of statistical measures and demographic parameters (Tarsi and Tuff 2012). This allows population ecologists to evaluate a populations' characteristics and how these change over time (Lebreton et al. 1992). These parameters include aspects such as birth rates, population size and density, sex ratios, as well as factors which can influence population growth (Dodge 2006).

Given these definitions of a population, population processes are considered as those that can influence a population's genetic variation and viability, including inbreeding, outbreeding, genetic drift, gene flow, mating systems, natural selection and adaptation (Woodruff 2001; Frankham 2010b; Ouborg et al. 2010). Many of these processes are interlinked. In conservation genetics both within and between population processes are important. Processes within a population largely deal with current and historical demography (e.g. changes in effective population size), as well as local adaptation, whereas processes between populations relate back to population structure which is determined by the extent of connectivity and gene flow among sub-populations.

Connectivity

According to Kool et al. (2013) connectivity is the outcome of the relationships or dependence among or between different populations that can be separated spatially or temporally. Depending on the population structure of a species, natural connectivity or changes in connectivity can affect the persistence of the sub-populations. Connectivity can also be defined as the process in which material (e.g. genes, gametes, individuals) are exchanged between populations, affecting the demography of the population(s) (Cowen and Sponaugle 2009). The study of connectivity enables linking a variety of topics in ecology and evolution (Kool et al. 2013), such as the development of population genetic structure (Kool et al. 2011), dispersal and migration (Lowe and Allendorf 2010) and source-sink dynamics (Figueira and Crowder 2006). In addition, information on connectivity can add value especially when there are conservation decisions that need to be made regarding reserve network design (Cerdeira et al. 2010) and restoration (Raeymaekers et al. 2008). These aspects are important since there is a need for reliable information regarding the distribution and connectivity of populations that affect species persistence (Cushman et al. 2013). In order to ensure future existence of a species it is important to maintain and maximize the connectivity between protected areas located in fragmented landscapes (Wegmann et al. 2014).

Even if gene flow is low and does not contribute significantly in a demographic context, it can still play a vital role in the maintenance of genetic diversity (Woodruff 2001; Lowe and Allendorf 2010).

The determination of gene flow between populations is thus important. Moreover, if population sizes are very small and gene flow between populations is absent this can cause highly structured isolated populations which may in the end experience inbreeding depression causing a reduction in genetic variation (Woodruff 2001; Ouborg et al. 2010). Whereas, if populations that are genetically very distinct exchange genes, outbreeding depression may occur (Frankham et al. 2011).

Population structure

Population structure can be considered along a continuum ranging from panmixia to highly structured populations (Fig. 1.4) (Waples and Gaggiotti 2006). Panmixia refers to random mating in populations (Forbes and Boyd 1997). This can only take place if there is a high degree of connectivity between sub-populations to ensure high levels of gene flow (Woodruff 2001). If animals are highly mobile this enables them to disperse over long distances, causing high levels of gene flow between sub-populations, resulting in low levels of differentiation between the sub-populations (Boulet et al. 2007). Highly structured populations on the other hand are sub-populations which are entirely isolated with no or very limited gene flow (Waples and Gaggiotti 2006).

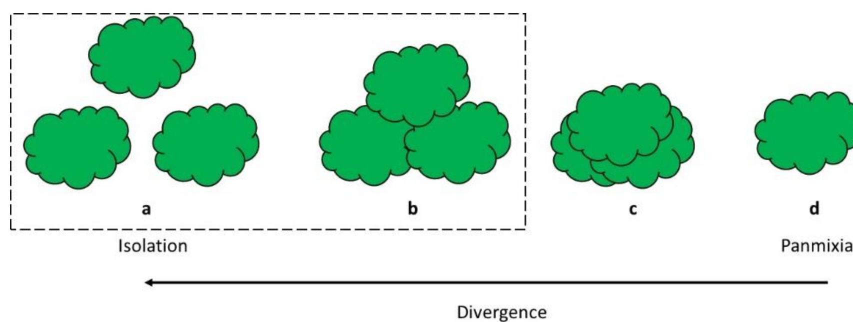


Fig. 1.4 Population structure along a continuum ranging from random mating to complete independence between sub-populations. Each group of green patches represents a group of sub-populations with different degrees of connectivity: a) Highly structured populations; b) populations with modest connectivity; c) substantial connectivity; and d) panmixia (random mating). The rectangle indicates the hypothesized population structure of suni between forest patches. Adapted from Waples and Gaggiotti (2006)

Along the continuum (from high to low gene flow) several other forms of population structure have been identified, such as isolation-by-distance, stepping-stone patterns, metapopulations, or mainland-island models (Kessler and Shnerb 2015; Hernández et al. 2016). A metapopulation can be defined as a group of sub-populations of the same species, variously connected occurring in the same geographical area, allowing the exchange of individuals through dispersal and migration (Woodruff 2001; Allendorf et al. 2010). Therefore, a metapopulation typically functions as a single population sustaining extinction and recolonization over time. The stepping-stone model allows in each generation an individual population to exchange migrants with many other populations in close proximity but the effective population size essentially remains the same (Kimura and Weiss 1964). The mainland-island model refers to a mainland patch or population which is surrounded by multiple small populations within the dispersal distance (Harrison and Taylor 1997). This allows for source-sink dynamics in which the mainland population is the source providing the emigrants and the small

surrounding populations are the sinks or net recipients of immigrants (Harrison and Taylor 1997; Lowe and Allendorf 2010). This creates a rescue effect, which allows the small populations to gain genetic diversity and persist in the future (Harrison and Taylor 1997).

Through inferring parameters such as effective population size, population splits and migration rates we are able to identify spatial structure, as well as gene flow between populations and thus the underlying population structure using some of the best available software (Woodruff 2001; Kool et al. 2013; Greenbaum et al. 2016; Petkova et al. 2016; Bradburd et al. 2018). This further contributes to the identification of ESUs and MUs which can guide decisions regarding translocations and management (Woodruff 2001; Coates et al. 2018).

Background on suni (*Nesotragus moschatus*)

Phylogeny

The suni is a member of the Bovidae family which comprises of both buffalo and antelope species (Fernández and Vrba 2005). Some of the oldest bovid fossils have been discovered in France and sub-Saharan Africa, with roughly the same age, suggesting a likely origin of approximately 23 MYA (Matthee and Davis 2001). These fossils create uncertainty with regards to where this family originated from however, other studies based on allozyme divergence suggests that the Bovidae originated from Africa (Georgiadis et al. 1990).

The Bovidae is a highly diverse family which consists of a few tribes and subfamilies. Molecular markers and morphological features have been used to identify the relationships amongst the tribes and subfamilies (Fig. 1.5 and 1.6). Interestingly, the subfamily Antilopinae was thought to include the tribes Antilopini and Neotragini (Gatesy et al. 1997; Matthee and Robinson 1999), of which the Neotragini tribe consisted of the small antelopes (genera *Ourebia*, *Oreotragus*, *Madoqua*, *Raphicerus*, *Neotragus, Nesotragus* and *Dorcatragus)*. However, the relationships among the members of the Neotragini are not well resolved since some phylogenies clustered the current genus *Nesotragus* (previously *Neotragus*) with the impala (which is classified in its own subfamily) (Matthee and Robinson 1999; Matthee and Davis 2001). In addition to this a more recent study have been performed on all three *Neotragus* species (*N. batesi, N. moschatus and N. pygmaeus*) (Bärmann and Schikora 2014). The results from this study indicated that *N. moschatus* and *N. batesi* were closely related but that *N. pygmaeus* had no close relationship with either one of them and the authors thus suggested that the genus name *Nesotragus* should be resurrected for both *Nesotragus moschatus* and *Nesotragus batesi* which was previously given by von Dueben in 1846.

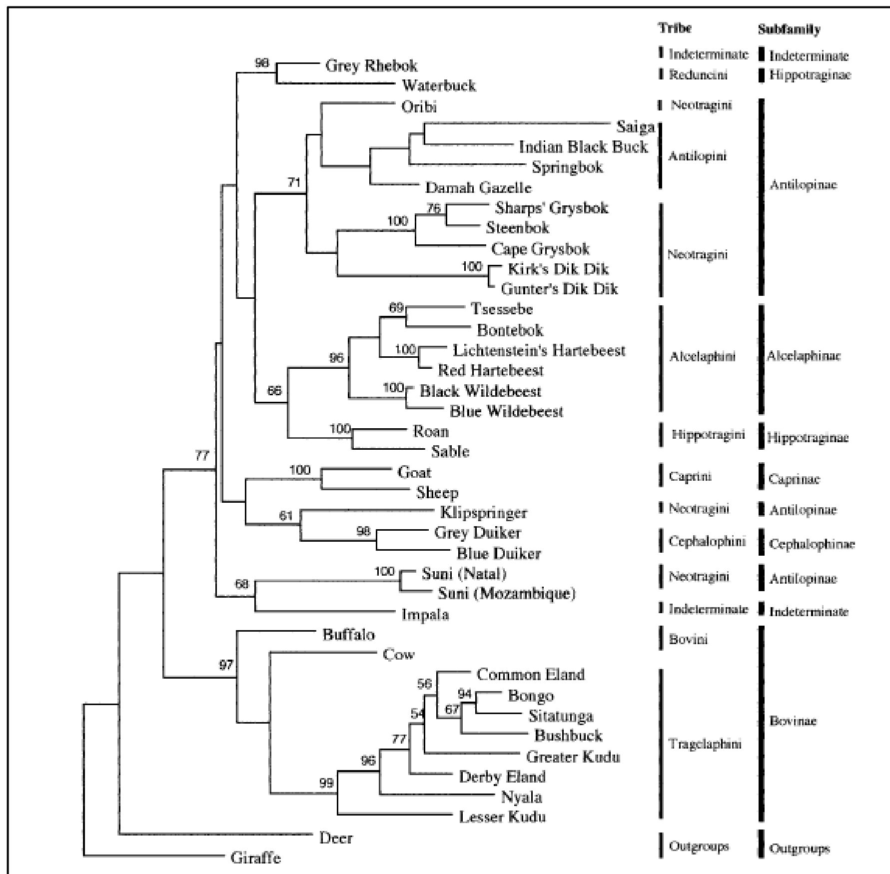


Fig. 1.5 Parsimony tree based on the mtDNA cyt *b* gene of 38 bovid taxa. Suni is clustered as a distant sister taxon of the impala. Bootstrap values are indicated at nodes. Reprinted from Matthee and Robinson (1999) with permission.

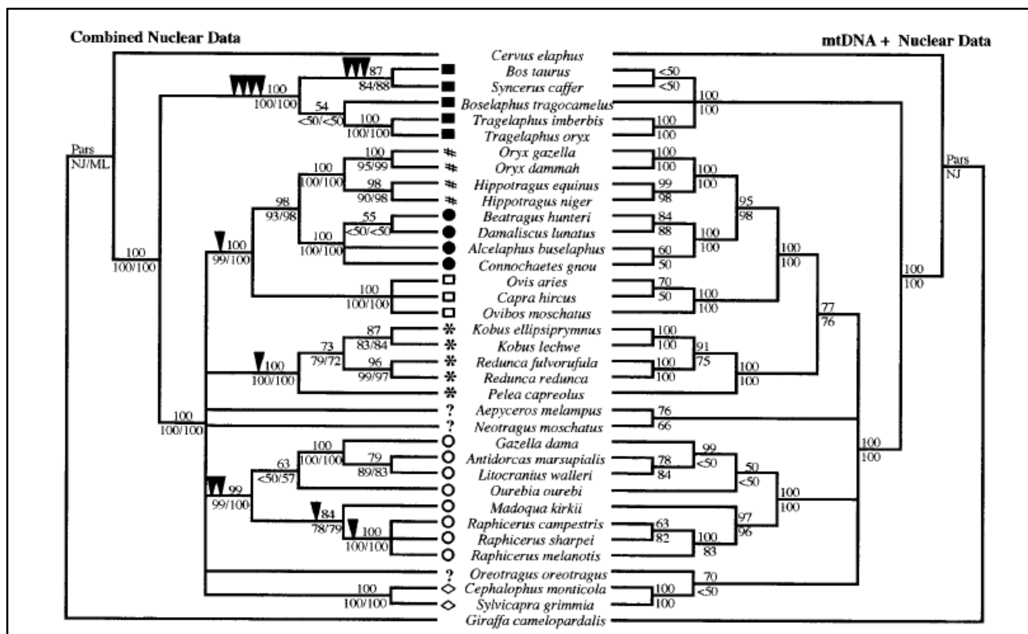


Fig. 1.6 Combined phylogenetic trees based on nuclear genes and mtDNA of 34 bovid taxa. It is suggested that the suni is a unique lineage and cannot be clustered within any of the subfamilies. Bootstrap values are presented at the nodes, with numbers above the diagonal branches representing parsimony values; those below the branches indicate neighbor-joining and maximum likelihood support. Subfamilies are represented as follow: black square = Bovinae; # = Hippotraginae; black circles = Alcelaphinae; white squares = Caprinae; * = Reduncinae; O = Antilopinae; diamond = Cephalophinae; ? = uncertain status. Reprinted from Matthee and Davis (2001) with permission.

Subspecies

In total six different subspecies (*N. m. moschatus* (Zanzibar)), *N. m. livingstonianus, N. m. akeleyi, N. m. deserticola, N. m. kirchenpaueri* and *N. m. zuluensis*) have been described over the years, based on phenotypical features and geographic distribution (Fig. 1.7) (Ansell 1971; Kingdon and Largen 1997; Frost and Carnaby 2015). However, since the boundaries between these subspecies are not clearly delineated, no subspecies were assessed in the most recent IUCN Red List (IUCN SSC Antelope Specialist Group 2016). Nevertheless, the regional Red List of Mammals of South Africa, Lesotho and Swaziland does acknowledge the presence of two subspecies (*N. m. livingstonianus* and *N. m. zuluensis*) in southern Africa (Hunnicutt et al. 2016).

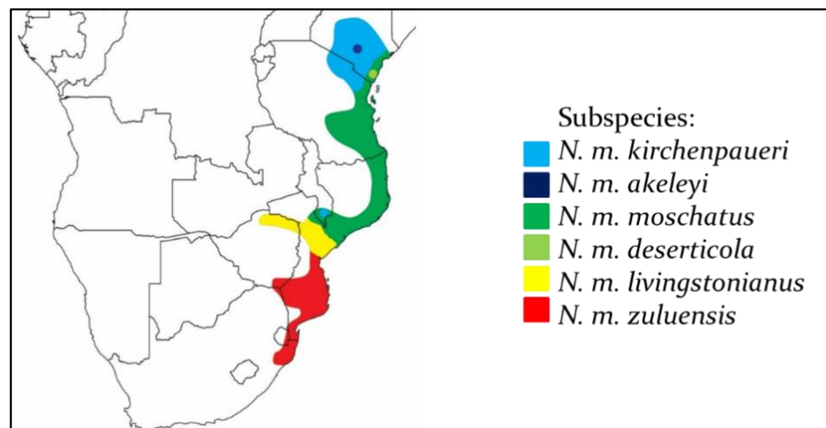


Fig. 1.7 Distribution map of the six described suni (*Nesotragus moschatus*) subspecies (Ansell 1971; Skinner and Chimimba 2005; Groves and Grubb 2011; Frost and Carnaby 2015).

The majority of subspecies described are phenotypically very similar with small differences present in their coat colour. These differences may not be significant since environmental factors could cause changes in the physical appearance of animals (Lillycrop and Burdge 2014). One study on captive individuals have identified some cryptic genetic variations between subspecies. A karyological analysis performed on 38 captive-born suni in North American Zoos, revealed a variety of diploid numbers, 2n = 52-56 (Kingswood et al. 1998). These variations between the cytotypes are due to two centric fusions of chromosomes 4;19 and 18;22 (Kingswood et al. 1998).

With the use of pedigree records, additional karyotyping of the captive individuals and previous results (Seluja et al. 1985; Robinson et al. 1997), the study was able to show that the diploid numbers 2n = 52 and 2n = 56 belonged to *N. m. akeleyi* and *N. m. zuluensis* respectively (Kingswood et al. 1998). According to these results, as well as the karyological analysis of parent and offspring comparisons, it was hypothesized that the cytotype variations 2n = 53, 54 and 55 were all due to hybridization between the two subspecies (Kingswood et al. 1998). Yet, there are still some uncertainties whether or not these variations are indeed due to hybridization, or if there are fixed chromosomal differences that occur in the wild (Kingswood et al. 1998). Previous studies on antelopes and cattle have revealed that centric fusions can cause an increase in perinatal mortality as well as a decrease in fertility within heterozygous offspring (Dyrendahl and Gustavsson 1979;

Benirschke et al. 1984). This type of research regarding reproduction has not yet been done in suni. Hence, with all of above taken into account the study hypothesized that centric fusions might have been the reason for the decrease in fitness, but caution had to be taken since factors such a inbreeding, which could also increase mortality rates, were not taken into account (Kingswood et al. 1998).

Due to the lack of knowledge concerning suni reproduction, uncertainties regarding causes of cytotype variations and the fact that the study was based on captive-born individuals, this necessitate further investigation especially since the results of the study (Kingswood et al. 1998) indicated that reproductive breakdown was possibly taking place in captivity. Moreover these cryptic genetic variations are important and should be taken into consideration during practices such as captive breeding, translocations and reintroductions because if ignored this species could be affected negatively in the future (Kingswood et al. 1998; Gibson and Reed 2008). Due to this and the interest of the suni subspecies it would be advised that *N. m. akeleyi* and *N. m. zuluensis* be managed as two separate units until new information has come to light which could help in the management and conservation of the species (Kingswood et al. 1998).

Distribution

The suni is widely spread over the eastern parts of Africa; the currently understood distribution is depicted in Fig. 1.7 (Skinner and Chimimba 2005). More specifically these animals are distributed in a long stretch from the eastern part of KZN north to the south-eastern part of Kenya, hence covering regions such a Malawi, Zimbabwe, Tanzania, Mozambique and South Africa (Ansell 1971; Skinner and Chimimba 2005; Frost and Carnaby 2015). Due to the secretive nature of the species and rarity of sightings in parts of the distribution, the full extent of the species' distribution and abundance is not currently known. As mentioned earlier all subspecies are phenotypically very similar leading to uncertainty in the status, boundaries and distributions of each subspecies. It has been suggested that *N. m. kirchenpaueri* can be found in Tanzania and on Mount Kenya along with *N. m. moschatus* and *N. m. akeleyi* respectively (Ansell 1971; Groves and Grubb 2011). In addition to this *N. m. zuluensis* is said to be distributed in north-eastern KZN, southern Mozambique and the south-eastern parts of Zimbabwe (Ansell 1971; Skinner and Chimimba 2005). *N. m. moschatus* on the other hand can be observed in East African areas such as the forests on Kilimanjaro, the Aberdares, Mount Kenya and a coastal region in Tanzania and Zanzibar (Ansell 1971; Frost and Carnaby 2015). Lastly in Malawi, Zimbabwe, Mozambique and north-eastern Africa the subspecies *N. m. livingstonianus* can be found (Ansell 1971; Frost and Carnaby 2015).

Description

These animals are very small with an average height of 36.2 cm and approximate weight of 5.3 kg (Fig. 1.8) (Skinner and Chimimba 2005; Groves and Grubb 2011). Their bodies are compact with slender legs, a short neck and broad head (Skinner and Chimimba 2005; Frost and Carnaby 2015). In addition their coat colour ranges from a light to a rufous brown, which appears to be darker on the

back than on the sides and legs of the animal (Skinner and Chimimba 2005; Frost and Carnaby 2015). Their belly, chin, throat and inside of their legs are white with black rings around the eyes and above their hooves (Skinner and Chimimba 2005; Frost and Carnaby 2015). They also have a broad black band on their faces which stretches from their nose to a point between their eyes (Skinner and Chimimba 2005) and a noticeable continuously flickering short tail (Frost and Carnaby 2015). Prominent periorbital glands can also be observed which are responsible for producing a sticky product that are used by males to mark their territories (Skinner and Chimimba 2005; Frost and Carnaby 2015). Coupled with this only the males can be observed having horns which on average are 8 cm long and ridged for the majority length of the horn (Frost and Carnaby 2015).

Fig. 1.8 Image of a suni antelope (*Nesotragus moschatus*) (Megan Loftie-Eaton, http://vmus.adu.org.za/)

Suni habitat and niche

In order to understand the niche of the suni, it is important to first understand the habitat that the species occurs in, its conservation status and how these factors are linked. Africa is very rich in biodiversity and consist of various biomes, for example South Africa alone consists of seven biomes (Rutherford and Westfall 1994; Mucina and Rutherford 2006). The biomes in relation to the distribution of the suni are shown in Fig. 1.9. The species has been recorded from four different biomes, namely the tropical and subtropical moist broadleaf forest, the tropical and subtropical grasslands, savannas and shrublands, the flooded grasslands and savannas and lastly the deserts and xeric shrublands. However, it is important to note that the outline indicating the suni's distribution (Fig. 1.9) should not be interpreted as a continuous distribution, but rather as the possible borders of the species discontinuous distribution.

As shown in Fig. 1.7 and Fig. 1.9 the suni's distribution expand over several countries and different biomes. In Kenya e.g. suni can be found in dry forest, woodlands, grasslands and riverine forests (Hoppe 1977; Hoppe et al. 1983; Rovero et al. 2005). Whereas in Tanzania they inhabited coastal savanna and forests (Treydte et al. 2005), in Malawi they were found in short grasslands, miombo woodlands and afromontane forests (Byrne et al. 2019) and lastly in the savanna and sand forests of Mozambique (Prins et al. 2006). However suni can also be found on Mnemba Island which is 4.5 km off the northeast tip of Unguja Island (one of the main islands of Zanzibar archipelago) (Fiske 2011) which consists of *Caussuarina* forest associated with sand dunes and sandy beaches (Fiske 2011). In South Africa, the focal point of this study, suni can be seen in habitats such as sand forests, woodlands and lowveld riverine forests however, suni show preference towards the sand

forest over woodlands in South Africa as well as in Mozambique (Prins et al. 2006; Belton et al. 2008; Ramesh et al. 2016).

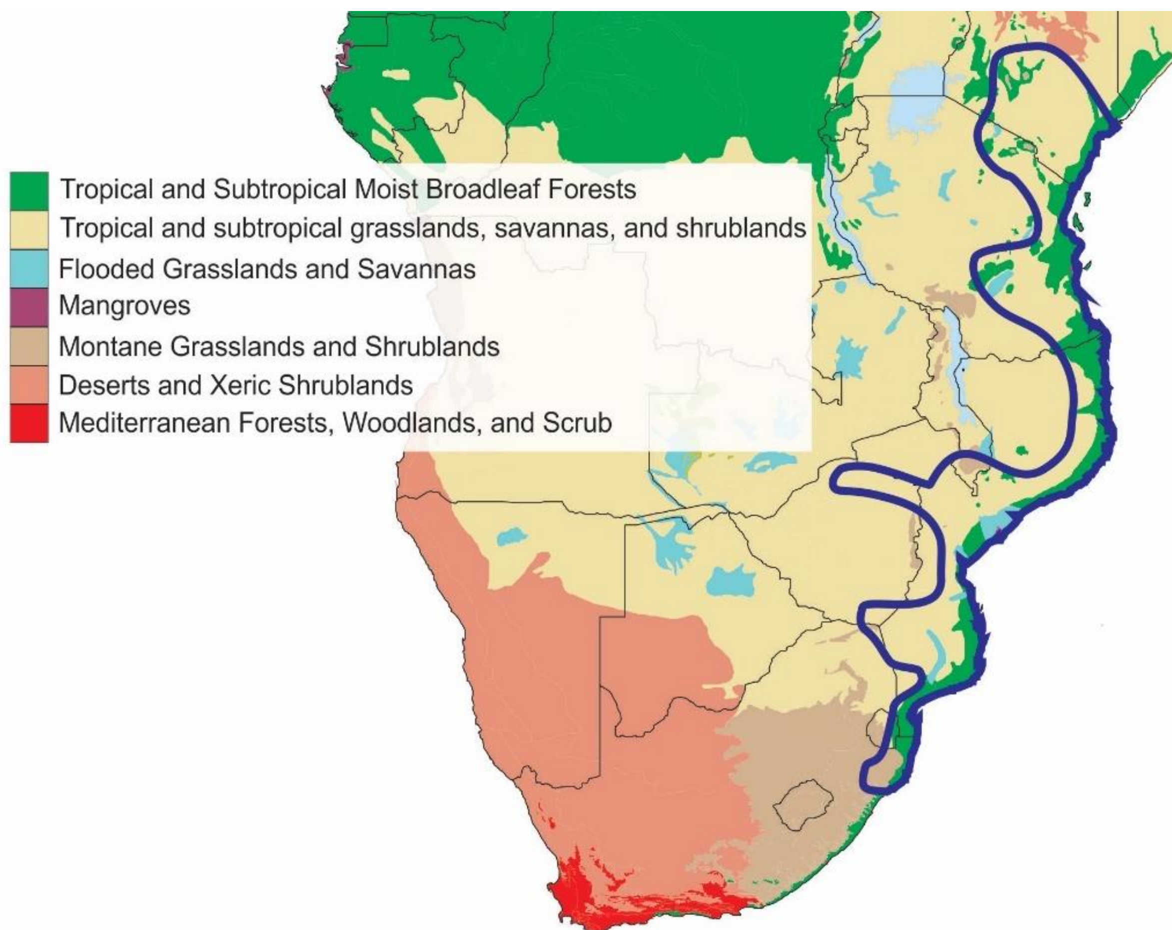


Fig. 1.9 The distribution of different biomes across East and Southern Africa. Colours correspond to the different biomes. The dark blue line represents the boundaries of the suni distribution (Olson et al. 2001).

Sand forests in southern Africa are mainly found in KZN and Mozambique, distributed in a broad and highly fragmented belt (Chytrý et al. 2008), with Maputaland particularly rich in forest dwelling species (Smith et al. 2008). It consists largely of dense evergreen thickets, which could grow up to 15 m, creating tall enclosed canopies as well as a well-developed shrub-layer with high density stems (Matthews et al. 2001; Gaugris et al. 2004; Chytrý et al. 2008). The ground on the contrary is sparsely covered with a few mixed grasses and herbs (Matthews et al. 2001; Gaugris et al. 2004; Skinner and Chimimba 2005). In addition to this sand forests are made up of patches containing the above mentioned, with each patch surrounded by grassland communities creating borders for both plant and animal species (Fig. 1.10) (Matthews et al. 2001).

Although these sand forests are said to be dense, animals such as nyala (*Tragelaphus angasii*) and elephant (*Loxodonta africana*) have been shown to decrease the density of these forests (Belton et al. 2008; Ramesh et al. 2016). In addition to this other factors such as uncontrolled extraction of wood for woodcraft and fuel have also added to the decline in habitat density (Chytrý et al. 2008). Due to the above mentioned factors the sand forest have become more fragmented and has a conservation status of critically endangered (Mucina and Rutherford 2006; Chytrý et al. 2008). This

plays an important role, especially since suni is dependent on the habitat and if the habitat becomes more fragmented, it could cause a decline in the population which could ultimately lead to local extinction (Brooks et al. 1999; Laurance et al. 2002; O'Connor and Kuyler 2009). Moreover, fragmentation could cause a decrease in the habitat area along with an increase of distance between patches (Brooks et al. 1999; Laurance et al. 2002), which in the end could prevent dispersal from occurring (Renjifo 1999; Laurance et al. 2002; Jewitt et al. 2015).

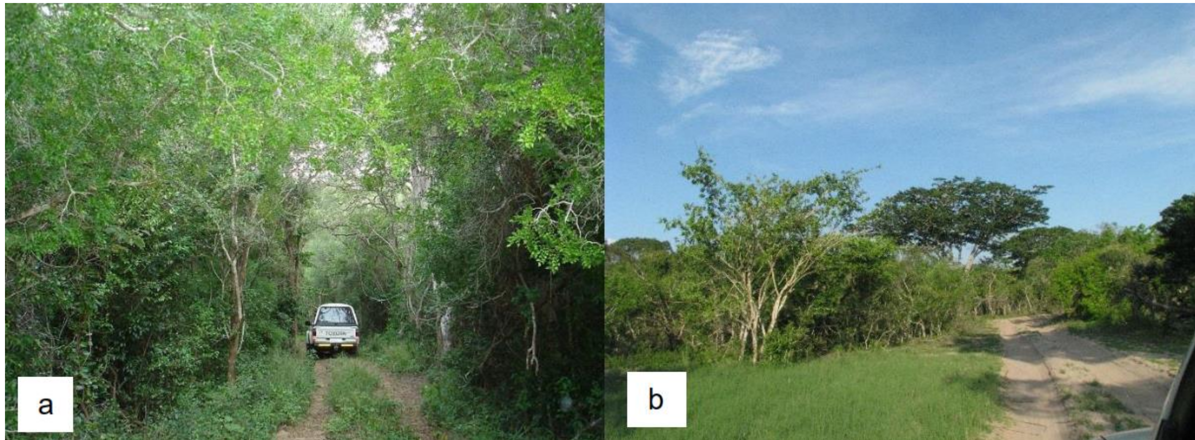


Fig. 1.10 a) Sand forest in Tshanini Nature Reserve (JP Bloomer) and b) grassland borders found surrounding the forest patches in Tembe Elephant Park (M Swanepoel)

Since sand forests are highly threatened due to fragmentation (more than in its natural fragmented state) it is critical for more research to be performed especially since studies have shown that habitat loss and land cover are some of the biggest drivers of biodiversity loss (Sarukhán et al. 2005; Jetz et al. 2007). In addition to this it has also been shown that habitat specialists are affected more by habitat fragmentation than are generalists (Manne and Pimm 2001). Through studying the genetic diversity of suni over different spatial and temporal scales, more insight can be gained into their association with different habitat types and the impact of changes in these habitats.

Age of South African and Mnemba Island suni

Since the suni is an understudied species, very little information is currently available regarding their history/age in South Africa. Due to this, other information for example age of habitats can be used to estimate the age of the suni species in South Africa. Moreover, as discussed earlier suni are habitat specialists and are found more associated with the sand forest habitat in South Africa. Due to this I will making inferences on the age of the species in South Africa with means of the information gained from the sand forest. According to Matthews (2005), Maputaland coastal plains consists of three dune cordons and one dune field. Cordon three can be found distinctly throughout most of Maputaland covering areas such as NDU and TEM dating back to 3 million to 125 000 BP (Matthews 2005). Therefore, allowing me to hypothesize that the maximum time for the suni species to have been present in South Africa to be approximately 3 million to 125 000 BP.

The suni found on Mnemba Island on the other hand is a very recent population. In the year 1991 three suni were first introduced to the island and have grown dramatically in numbers ever since (Fiske 2011). Since then four translocation events to Jozani forest have taken place with the most recent known translocation, of 96 individuals, in 2011 (Fiske 2011). These translocations were deemed necessary in order to reduce the risk of overgrazing especially since suni reproduces very quickly (Fiske 2011).

Suni diet

With regards to food, suni utilizes a variety of food sources and have been recorded to utilise 70 different items in Mozambique (Prins et al. 2006). However suni have also shown some preference towards dicotyledonous leaves in Kenya (Hoppe et al. 1983) and in most cases a high preference for fresh fallen leaves (Lawson 1986; Lawson 1989). In addition to this they also make use of fruits such as figs, flowers and growing tips of shoots however, they are not dependent on these items (Lawson 1989; Skinner and Chimimba 2005). In other instances, these animals have also been observed eating mushrooms in KZN (Heinichen 1972; Skinner and Chimimba 2005). All of these foods provide soluble plant sugars as well as other substances which are necessary for their nourishment (Lawson 1989; Skinner and Chimimba 2005). They are rarely seen drinking water due to the majority of their water intake coming from moisture of their food (Lawson 1986; Lawson 1989). With regards to their diet on Mnemba Island it has been found that suni does consume *Guettarda speciosa* (Mkungu wa pwani) and *Tetraceara boivaniana* (Pumba za paka) however this was only based on a three-day observation study period (Fiske 2011).

Population size and conservation status

Determining population estimates for suni have not been easy due to their patchy occurrence and shy nature (Hunnicutt et al. 2016). In addition to this, studies have also shown that the sand forests, which they favour in South Africa, are currently highly fragmented and could cause the populations to become more isolated (Jewitt et al. 2015; Hunnicutt et al. 2016). According to the IUCN Red List the suni is classified as Least Concern with an approximate population size of 365,000 individuals (IUCN SSC Antelope Specialist Group 2016). However, a recent study using dung abundance counts has shown a dramatic decline in suni populations in both Tembe Elephant Park and Mkhuze Game Reserve (Hunnicutt et al. 2016). Furthermore, a decline of > 50% in population size was observed over more than three generations in a previously identified sub-population, which was the largest in their local assessment area (Hunnicutt et al. 2016). Additionally, the authors estimated that in South Africa populations can consist of 440-4890 individuals and sub-populations of 62-750 mature individuals. Due to these observations it was decided to list the suni as regionally Endangered based on their decline, restricted area of occurrence and the lower limits of population and sub-population size ranges (Hunnicutt et al. 2016).

The observed decline in the South African suni population can be due to a variety of threats. Different authors have emphasized the roles of poaching and predators in decreasing suni numbers

(Lawson 1986; Skinner and Chimimba 2005). These predators include martial eagles (*Polemaetus bellicosus*), crowned eagles (*Stephanoaetus coronatus*), pythons (*Python sebae*), leopard (*Panthera pardus*), jackal (*Canis aureus*) and spotted hyena (*Crocuta crocuta*) (Skinner and Chimimba 2005; Frost and Carnaby 2015). In addition to this, it has also been suggested that human impact could have been a contributing factor to habitat reduction and fragmentation hence, making it likely one of the biggest reasons for the recent species decline (Lawson 1986). Furthermore an increase in habitat utilization by big animals such as nyala and elephant also tend to reduce the density of the shrubs making suni more vulnerable to predation (Matthews et al. 2001; Skinner and Chimimba 2005; Belton et al. 2008; Hunnicutt et al. 2016; Ramesh et al. 2016). This overgrazing can lead to permanent changes in the sand forest which is not ideal for local suni persistence (Gaugris et al. 2004; Gaugris and Van Rooyen 2007).

In order to promote the persistence of suni, different measures have been put in place, with conservation areas playing a significant role in the species' conservation. In most parts of the suni's distribution, conservation parks have been created, for example Aberdares National Park (Kenya), Lengwe National Park (Malawi) and Maputo Game Reserve (Mozambique) (Gaugris et al. 2004; Frost and Carnaby 2015). The largest suni population in South Africa can be observed in Tembe Elephant Park (Gaugris et al. 2004; Hunnicutt et al. 2016). This park has a very high woodland density which is advantageous for the suni's existence. Unfortunately the vegetation density have been decreasing due to larger animals (Gaugris et al. 2004; Hunnicutt et al. 2016).

Even though conservation parks do contribute to the protection of the suni population other methods, such as conservation genetics studies, can also help increase their future chances of persistence (Allendorf et al. 2010; Frankham 2010b). With the use of genetic studies scientists are able to determine conservation units, population status, connectivity, genetic variability and outbreeding depression (Allendorf et al. 2010) that will add value to future management plans.

Behavioural ecology

Behavioural ecology can be defined as the study of how natural selection shapes behaviour (Raven and Johnson 1986). This is an important factor to be studied since adaption in behaviour can increase the survival of an organism and in such a manner increase their fitness (Tuomainen and Candolin 2011). During environmental changes, traits can evolve through genetic drift, gene flow and natural selection. Some of the questions asked regarding behavioural ecology pertain to the physiology behind the behaviour observed, the function of the behaviour and whether or not it is advantageous (Raven and Johnson 1986).

Although knowledge of suni behaviour is limited, a few studies have revealed habits that are relevant in studying the conservation genetics of the species (Lawson 1986; Lawson 1989; Somers et al. 1990; Frost and Carnaby 2015). It was observed that 77% of suni can be found individually, 12% in pairs and 11% in families (consisting of a male, female and offspring) but, due to the density of shrubs these values might be incorrect since observation is quite difficult (Lawson 1986). The

species has been shown to use middens, which are communal defecation and urinal spots. With regards to the pairs, if a female used a midden, the male will follow and defecate on the exact spot, again adding to marking his territory (Skinner and Chimimba 2005).

Future fine-scale genetic studies utilizing the markers developed in the present study, would add valuable insight into suni population density (by being able to identify individual suni), relationships between suni in local habitat patches (relatedness and social structure), as well as connectivity over small to medium spatial scales (determining what represents a suni demographic unit/population).

Investigative tools

Microsatellite markers

From many articles, it has been shown that genetic markers can be used to answer different conservation genetics questions (Boulet et al. 2007; Allendorf et al. 2010; Ouborg et al. 2010; Allendorf 2017). Genetic markers can be defined as short DNA sequences with a known physical location on chromosomes (Benavides and Guénet 2012). These markers then act as points of variation which could help one identify individuals, populations and in some cases even cells (Bruford and Wayne 1993; Jarne and Lagoda 1996; Ramón-Laca et al. 2018). These genetic markers also create the opportunity to determine genetic diversity for example in antelope (Pinto et al. 2015). In most studies the choice of marker depends on the characteristics of the marker, as well as the study species' characteristics. There are multiple genetic markers today that can be used for molecular studies including allozymes, microsatellites, simple sequence length polymorphisms (SSLP), single nucleotide polymorphisms (SNPs) and many more (Selkoe and Toonen 2006; Al-Samarai and Al-Kazaz 2015; Grover and Sharma 2016). However, this study only made use of microsatellites and mtDNA. The reasoning behind the use of microsatellites is discussed in the section below.

Microsatellites, also known as short tandem repeats (STRs), are short DNA sequences which can consist of one to six nucleotides which are repeated multiple times (Selkoe and Toonen 2006). The number of nucleotides present in the repeat (repeat unit) determines the different classes, for example mono-, di-, tri- and tetranucleotides, which can be found. Microsatellites are widely spread in eukaryotic genomes and are highly polymorphic since the number of repeats can vary between individuals (Bruford and Wayne 1993). These sequences/repeat units can be repeated between five to forty times due to various aspects such as mutation rate (Selkoe and Toonen 2006). The main mechanism by which repeat variation is generated is replication slippage, leading to the gain or loss of repeat units (Ellegren 2004). The flanking regions are typically conserved sequences within species and are unique for each locus (Schlötterer 2000), enabling the design of locus-specific primers flanking the repeats. However, insertions and deletions may also occur in the flanking regions altering the allele sizes by different increments, leading to homoplasy (Grimaldi and Crouau-Roy 1997).

Microsatellite markers have been one of the most popular marker to use in population studies during the years 2000-2015 (Morin et al. 2004; Chistiakov et al. 2006; Narum et al. 2008; Väli et al.

2010) due to their polymorphic characteristics and informative nature (Payseur and Cutter 2006) yet, ever since the discovery of Single Nucleotide Polymorphisms (SNPs) the use of microsatellites have declined. This could be explained by the fact that SNPs are less expensive, have a high throughput during analysis, has a lower mutation rate and a simple underlying mutation model (Morin et al. 2004). In some articles it has also been briefly mentioned that SNPs might become the chosen marker instead of microsatellites (Al-Samarai and Al-Kazaz 2015; Vieira et al. 2016). However these authors (Al-Samarai and Al-Kazaz 2015; Vieira et al. 2016) also believe that even though the latter might occur, microsatellite markers will still be used since they have other advantages.

To elaborate microsatellites are highly rich in genetic variation compared to other markers (Schlötterer 2004; Payseur and Cutter 2006; Selkoe and Toonen 2006; Vieira et al. 2016) and can be amplified from low quality DNA such as in the present study (lower than what is necessary for SNPs) (Grover and Sharma 2016). In addition microsatellite markers are also highly informative since they are inherited in a codominant Mendelian fashion (Abdul-Muneer 2014) allowing these markers to have a high resolution for individual identification, relatedness and parentage assignment (Morin et al. 2004; Harrison et al. 2014). Furthermore evolutionary history as well relationships (Vieira et al. 2016) can be studied by determining estimates such as bottlenecks, kinship and migration rates with the data obtained from microsatellites (Selkoe and Toonen 2006; Du et al. 2016; Miller et al. 2016; Vieira et al. 2016; Radhika et al. 2018; Shi et al. 2019). To date no microsatellite markers have been developed specifically for the suni. However, a study has shown that microsatellite markers of other bovids can be used on suni (Engel et al. 1996). Hence, the present study made use of many microsatellite markers developed in other bovids (Miller et al. 2016), as well as ones previously tested in a pilot study focussed on suni (Reitmann 2008).

Genetic non-invasive sampling and its applications

DNA recovery from dung samples

Non-invasive sampling has become an attractive method to use in answering questions about species in the wild (Chiou and Bergey 2018). This is possible since many different types of material could be used, for example dung, urine, or hair from mammals (Kohn and Wayne 1997). However dung samples have become the most used source overall (Fernando et al. 2003; Bourgeois et al. 2019), following the first successful PCR on dung DNA (Höss et al. 1992). This method utilises the epithelial cells present on the outer layer of the dung samples to extract DNA (Flagstad et al. 1999).

This type of sampling has also become the most used since so many obstacles are experienced during more invasive sampling such as tissue and blood collection. Moreover obtaining samples from wild animals could be difficult especially in the event of large animals which need to be tranquilised/restrained or in small and elusive animals which are hard to find (Beja-Pereira et al. 2009; Bourgeois et al. 2019). By making use of the less invasive method, cost of handling the animals could be reduced as well as enable scientists to obtain samples without hurting or encountering animals in the field (De Bondi et al. 2010; Ramón-Laca et al. 2018). Another advantage of this

method is that researchers are now able to answer more questions about animals in the wild and make valuable contributions to the field of conservation genetics of wild populations (Schwartz et al. 2007; Allendorf and Luikart 2009; Gray et al. 2014; Caragiulo et al. 2015). Questions regarding connectivity, migration rates, genetic diversity and relatedness can now be answered for species which were hard to study previously (Höss et al. 1992; Morin et al. 1994; Kohn and Wayne 1997; Schwartz et al. 2008; Beja-Pereira et al. 2009; Ferreira et al. 2018; Proença-Ferreira et al. 2019). In addition to this a study have also shown that making use of dung abundance counts are more accurate to estimate population numbers compared to other methods such as camera trapping (Guschanski et al. 2009; Mondol et al. 2009; Hedges et al. 2013; Ferreira et al. 2018).

Due to this, interest have grown with regards to non-invasive sampling and although being used extensively in the study of large animals, the approach has not been adopted as a general method in studying small and elusive animals (Ferreira et al. 2018; Bourgeois et al. 2019; Proença-Ferreira et al. 2019). Instead this method has been mostly used in larger mammals (Ferreira et al. 2018) such as elephants (Fernando et al. 2003; Bourgeois et al. 2019), big horn sheep (Wehausen et al. 2004), reindeer (Flagstad et al. 1999), wild ass (Costa et al. 2017), brown bears (Murphy et al. 2003), Bengal tigers (Bhagavatula and Singh 2006), gaur, banteng (Rivière-Dobigny et al. 2009), baboons (Chiou and Bergey 2017; Chiou and Bergey 2018), goats and lynx (Ramón-Laca et al. 2015). A few studies have been performed on small animals such as abbotts duiker (Bowkett et al. 2009), foxes (Piggott 2004), bush tailed rock wallabies (Piggott 2004), otters (Hájková et al. 2006; Lampa et al. 2007) and cabrera's vole (Ferreira et al. 2018; Proença-Ferreira et al. 2019). The more limited application of this approach for smaller mammals might be due to the ease experienced during the search of bigger animals compared to smaller animals.

Despite the fact that DNA can now be successfully extracted from dung samples, non-invasive sampling still has some problems with regards to the quantity and quality DNA extracted (Ramón-Laca et al. 2015; Chiou and Bergey 2018; Bourgeois et al. 2019). Environmental factors such as UV-rays and rain could cause degradation of the sample if it was exposed, thus resulting in low quality DNA yield (Fernando et al. 2003; Kovach et al. 2003; Piggott 2004; Hájková et al. 2006; Jeffery et al. 2007; Wultsch et al. 2015). Therefore, it is better to collect some of the freshest samples possible, reducing the chances of exposure. Aside from degraded samples being a drawback, samples could also be influenced further down during PCR. Since dung of bovids consist of plant material containing pigments, RNA and polysaccharides these could all act as PCR inhibitors during amplification (Litvaitis and Litvaitis 1996).

With all of the above taken into account it could be said that all of these factors can cause complications ranging from PCR amplification difficulty, allelic drop out and a high risk of genotyping errors (Fernando et al. 2003; Lampa et al. 2007; Chiou and Bergey 2018; Ramón-Laca et al. 2018; Bourgeois et al. 2019). Due to this, multiple different studies have been performed in order to improve protocols of existing extraction kits (Ramón-Laca et al. 2015; Costa et al. 2017), sampling methods used (Ramón-Laca et al. 2015; Bourgeois et al. 2019) and reducing cost regarding this type of work

(Fernando et al. 2003; Ferreira et al. 2018; Bourgeois et al. 2019). However, very few studies have been done where specific extraction and PCR protocols have been designed for highly degraded and exposed samples from small mammals.

Even though different challenges could be experienced, many different extraction kits are currently available which could reduce difficulty, making genetic non-invasive sampling much more appealing (Wehausen et al. 2004). All of these kits have been specifically designed to extract high quality DNA from dung samples however one kit (QIAamp DNA mini stool kit from Qiagen) has become the favourite amongst them all (Ramón-Laca et al. 2015; Chiou and Bergey 2017; Chiou and Bergey 2018; Bourgeois et al. 2019). This kit is very well-known for its good results and have proven to be successful in many different species (Kovach et al. 2003; Wehausen et al. 2004; Costa et al. 2017). Hence, supporting its use for DNA purification in suni dung samples in preparation for downstream genetic analyses.

PCR amplification and genotyping

As mentioned earlier successful DNA extractions from dung samples can be performed but factors such as sample degradation as well as PCR inhibitors can affect PCR amplification (Fernando et al. 2003; Lampa et al. 2007; Chiou and Bergey 2018; Ramón-Laca et al. 2018; Bourgeois et al. 2019). Both these issues can influence the success rate of amplification and as a result cause genotyping errors such as allelic drop out and false alleles (Ramón-Laca et al. 2018; Bourgeois et al. 2019). Thus, it has been advised to genotype each sample multiple times in order to reduce genotyping error, as well as perform analyses to ensure that the null allele frequency are low (Pompanon et al. 2005).

During the sampling process many different samples are collected and due to this the chances are highly likely that a different species' dung could also be collected, for example red duiker (*Cephalophus natalensis*) instead of suni. In a recent study it was shown that most of the identification errors that took place where due to similar sized species that co-occurred (Spitzer et al. 2019). Thus, it is critical to ensure that the samples used in further down analyses are in fact of the correct species. A previous study has shown that amplification and sequencing of mtDNA (12S) can be used to distinguish between dung of ungulate species (Van Vliet et al. 2008). The mitochondrial cytochrome *b* (cyt *b*) gene is an alternative, more variable gene that can be used since a large number of bovid cyt *b* sequences are available (Matthee and Robinson 1999; Ntie et al. 2010; Gaubert et al. 2015) and has been proven successful in a pilot study on suni (Reitmann 2008).

With the use of PCR amplification of the cyt *b* gene the fragments obtained could be sequenced and finally compared to the data available on NCBI (https://www.ncbi.nlm.nih.gov) by making use of a BLAST search (Altschul et al. 1990). The advantage of using a cyt *b* barcode is that the sequences generated for suni can also be used in downstream analyses of haplotype and nucleotide diversity as well as the phylogeographic patterns between sampling sites. In addition to this, collecting multiple samples could also cause resampling of the same individuals' dung. Due to this it is

important to make use of a highly variable marker such as microsatellites to help identify different individuals (Jarne and Lagoda 1996; Bhagavatula and Singh 2006; Ramón-Laca et al. 2018).

Analytical approaches and software

The present study made use of DNA sequences for species identification and downstream phylogenetic and gene genealogy analyses making use of software such as MEGA (Kumar et al. 1994; Tamura et al. 2011) and TCS (Clement et al. 2000). The mtDNA sequence data together with microsatellite genotypes, were used to determine how distinct the sampled subspecies are. Many studies have applied this approach and have been able to construct phylogenies based on the sequence differences detected (Holder and Lewis 2003; Tamura et al. 2011).

In addition to this the data generated through microsatellite amplification and genotyping were analysed using multiple different software programs (Montgelard et al. 2014). Parameters included linkage disequilibrium (LD), genetic differentiation (F_{ST}), null allele frequencies, number of alleles (N_a), observed and expected heterozygosity (H_O and H_E), private alleles (P_A), allelic richness (A_R) and inbreeding coefficients (F_{IS}). Genetic distances between subspecies as well as variation within and between sampling sites were determined in order to identify the genetic population structure of the species (Boulet et al. 2007; Ouborg et al. 2010; Kool et al. 2013; Eastwood et al. 2016).

Research questions

From this review it was clear that the suni has been understudied and that a number of aspects could inform the conservation genetic management of the species. In this dissertation, I addressed the following questions:

(1) How can currently available DNA extraction kits be optimized for extraction from highly degraded, small dung pellets for subsequent amplification of mitochondrial DNA and microsatellite loci?;

(2) What are the phylogenetic relationships between suni from sampling locations thought to represent *N. m. moschatus, N. m. livingstonianus* and *N. m. zuluensis*?;

(3) What is the degree of connectivity between the different suni populations in southern Africa across small to medium scales?

The overarching aim was to provide a foundation for future in-depth studies and to make a contribution to the identification of conservation units for management of the species.

Chapter 2 Optimization of markers for DNA sequencing and microsatellite genotyping of suni DNA extracted from dung, blood and tissue samples

Note that aspects of this chapter will be published - Swanepoel M, Klopper AW and Bloomer P. Optimization of dung DNA extraction for sequencing and microsatellite genotyping in an elusive small forest antelope. Target journal: *Conservation Genetics Resources*.

Abstract

Many different studies have been performed on dung samples, all of which have shown DNA extraction, PCR amplification and genotyping to be possible. However, many past studies investigated large mammals and few targeted smaller species, such as forest antelope. In this study I made use of two extraction kits, the NucleoSpin DNA Stool kit (Macherey-Nagel) and the QIAamp DNA mini stool kit (Qiagen), to compare the success of DNA extraction and subsequent molecular analyses. With the use of 95, degraded suni (*Nesotragus moschatus*) dung samples I found that the NucleoSpin DNA Stool kit (Macherey-Nagel) performed the best. In addition to this I also improved its protocol and obtained better quality DNA. In downstream optimization (i) a short cytochrome *b* barcode could successfully distinguish suni dung from co-distributed red duiker dung, (ii) molecular sexing was optimized and (iii) multiple microsatellite markers from other bovid species tested on the suni, with 16 loci identified for potential genotyping of suni. The observations from the present study compliments other non-invasive genetic studies, making use of highly degraded, small dung samples as DNA source material.

Introduction

The field of conservation genetics makes valuable contributions to the study of animal populations in the wild (Schwartz et al. 2007; Allendorf and Luikart 2009; Gray et al. 2014; Caragiulo et al. 2015). However, obtaining material from wild animals for genetic analyses can be challenging, such as in instances where study species are large and have to be physically restrained/tranquilized, or when species are elusive and rarely observed (Beja-Pereira et al. 2009; Bourgeois et al. 2019). Nevertheless, non-invasive sampling enables scientists to overcome these challenges and the cost of animal handling, in addition to lessening the disturbance to the animals in their natural habitats (Beja-Pereira et al. 2009; De Bondi et al. 2010; Ramón-Laca et al. 2018). Despite these advances, non-invasive sampling still has some drawbacks with respect to the quality and quantity of DNA obtained, requiring extensive optimization to ensure accurate genotyping (Chiou and Bergey 2018; Bourgeois et al. 2019). Several studies have aimed to improve protocols for existing extraction kits (Ramón-Laca et al. 2015; Costa et al. 2017). However, few studies have been able to optimize extractions and amplification of mitochondrial and nuclear DNA on degraded samples. In the present study the focus was on optimizing protocols for DNA extraction and subsequent DNA sequencing and microsatellite genotyping from small, often highly degraded dung samples.

Non-invasive studies utilize many different sample types, such as hair, dung and feathers (Kohn and Wayne 1997). Dung is a commonly used source and have become a good alternative (Fernando et al. 2003; Bourgeois et al. 2019) ever since the first successful PCR on dung samples was demonstrated (Höss et al. 1992). The approach has mostly been used on larger mammals (Ferreira et al. 2018). Since smaller species are elusive and infrequently observed, obtaining their fresh dung samples is also challenging.

Genetic non-invasive sampling (gNIS) methods have advanced (Chiou and Bergey 2018) and offer many benefits compared with invasive methods. Samples are obtained without disturbing or causing stress to the animal (Beja-Pereira et al. 2009; De Bondi et al. 2010). Aspects such as relatedness, genetic diversity, kinship and dispersal can now also be answered for species that are challenging to study (Höss et al. 1992; Morin et al. 1994; Kohn and Wayne 1997; Schwartz et al. 2008; Beja-Pereira et al. 2009; Ferreira et al. 2018; Proença-Ferreira et al. 2019). Population censuses can be done more accurately when using this method instead of invasive methods e.g. live-trapping (Guschanski et al. 2009; Mondol et al. 2009; Hedges et al. 2013; Ferreira et al. 2018).

Coupled with all the above benefits, DNA extraction kits optimized for dung samples have made the gNIS method more attractive (Wehausen et al. 2004). Performing extractions from dung material have always been a challenge but due to the newest kits on the market, dung extraction has become less of a struggle. From the variety of kits currently available for these types of extractions only one extraction kit appears to be preferred (Ramón-Laca et al. 2015; Chiou and Bergey 2017; Chiou and Bergey 2018; Bourgeois et al. 2019). The QIAamp DNA mini stool kit from Qiagen is particularly well-known for its capabilities and has proven to be successful in many mammal species, such as lagomorphs (Kovach et al. 2003), bighorn sheep (Wehausen et al. 2004) and equids (Costa et al. 2017) . Even though it is possible to extract DNA from dung, multiple studies have pointed out difficulties that can be expected in downstream methods when working with dung extracted DNA (extracted with no specific kit) (Taberlet et al. 1999; Morin et al. 2001; Vigilant 2002).

These difficulties are the main reasons why this methodology has not yet been widely adopted (Fernando et al. 2003; Ramón-Laca et al. 2015; Bourgeois et al. 2019). One of the biggest challenges is obtaining sufficient concentrations of high-quality DNA from dung samples (Ramón-Laca et al. 2015; Chiou and Bergey 2018; Bourgeois et al. 2019). In the first instance DNA quality is affected by the fact that samples cannot always be collected immediately after defecation. Prior to collection samples could have been exposed to environmental factors such as the sun and moisture (Fernando et al. 2003; Kovach et al. 2003; Piggott 2004; Hájková et al. 2006; Jeffery et al. 2007; Wultsch et al. 2015). These factors lead to degradation or complete removal of the intestinal epithelial cells on the outer surface of the pellets. Under some environmental conditions, dung is very quickly exposed to saprophytes or coprophagous species. In addition to these environmental factors, PCR inhibitors originating from for example the plant material in the diet of herbivores, can cause problems including PCR amplification failure and allelic drop out or false alleles leading to genotyping errors (Fernando et al. 2003; Lampa et al. 2007; Chiou and Bergey 2018; Ramón-Laca et al. 2018; Bourgeois et al. 2019).

In an attempt to correct for artefacts that may result from lower quality DNA, many dung DNA based studies have advocated for repeated extractions and repeated genotyping of samples, but this is time consuming and adds to the costs of analysis (Goossens et al. 1996; Fernando et al. 2003). Another obstacle that can be encountered during gNIS is resampling. This occurs when multiple samples are inadvertently collected from the same individual. Hence, it is advised to use

highly variable markers such as microsatellites, to help identify different individuals (Bhagavatula and Singh 2006; Ramón-Laca et al. 2018). However, resampling can become valuable when the number of individuals and movement patterns in a species are being investigated.

Many studies have contributed to overcoming the challenges of gNIS. For example, DNA extraction protocols have been altered (Ramón-Laca et al. 2015; Chiou and Bergey 2017; Costa et al. 2017; Chiou and Bergey 2018) so that DNA yield could be improved. Oher studies have changed their sampling method, making use of swabs, in order to only collect the critical section of the sample thereby reducing exposure to or co-extraction of PCR inhibitors (Ramón-Laca et al. 2015; Bourgeois et al. 2019). Furthermore studies have also tried to reduce the cost associated with these analyses (Fernando et al. 2003; Chiou and Bergey 2017; Chiou and Bergey 2018; Ferreira et al. 2018; Bourgeois et al. 2019). However, very few studies have exclusively focussed on highly degraded and exposed samples.

In this chapter I aimed to develop guidelines for optimization of DNA extraction and PCR amplification specifically for highly degraded, small dung samples. The main goal of this study was to show what changes could be made at different stages of handling such degraded material. I focused on extractions, mitochondrial DNA cytochrome *b* amplification, molecular sexing, microsatellite locus selection, amplification and fragment analysis. This chapter also gives some insight into how 16 microsatellite markers were selected for subsequent investigation of the genetic diversity of suni populations (Chapter 3).

Materials and methods

Sample collection

Dung sampling, representative of the distribution of suni in north-eastern South Africa, took place in four different nature reserves: Tembe Elephant Park (TEM), Tshanini Nature Reserve (TSH), Ndumo Game Reserve (NDU) and Phinda Private Game Reserve (PHI), all of which are in the northern part of KwaZulu-Natal (KZN). Additional samples were collected in Mozambique (MOZ) (tissue samples) and Mnemba Island (MNE) (dung samples). The different sampling localities are shown in Fig. 2.1 and sample information given in Table 2.1 and Table S2.1. For the first two nature reserves mentioned earlier transects were not used during sample collection. Instead previously identified localities in TEM and TSH, from pilot studies done in 2008 and 2013, were used as sampling starting points. From those points two to five researchers spread out and collected any fresh material that was available and recorded the location. This method was chosen to ensure that the whole area was covered. In some rare instances, suni tracks were used to locate suni dung. Due to rain during the sampling period, the sand was wet, enabling the field sampling team to identify and follow tracks in a much more effortless manner. By doing this the freshest samples possible were collected.

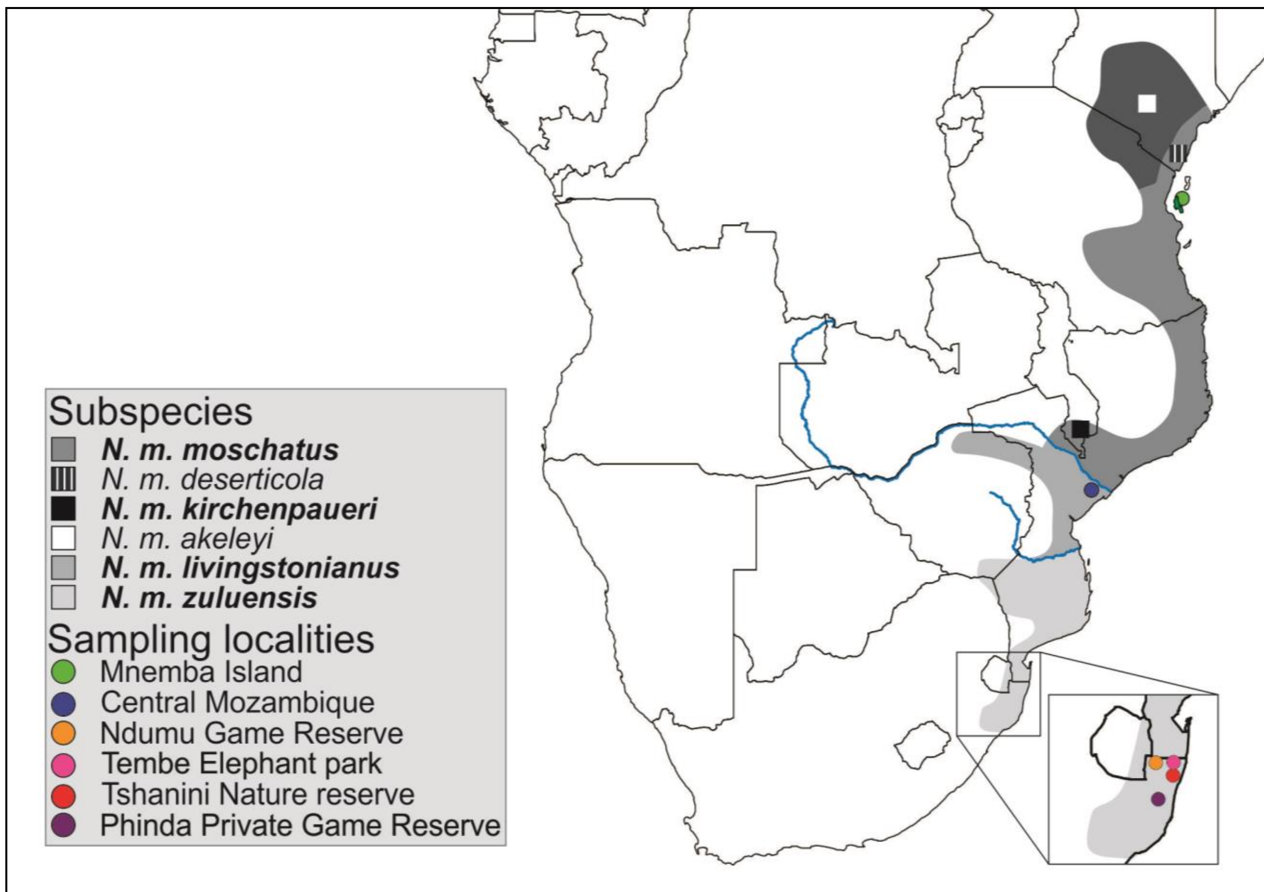


Fig. 2.1 Distribution map of suni subspecies and sampling localities. The subspecies shown in bold are the four currently recognised subspecies (Ansell 1971; Skinner and Chimimba 2005; Groves and Grubb 2011; Frost and Carnaby 2015).

Table 2.1 Tissue and dung samples collected from different locations used during extraction protocol testing

Location	# Samples	Sample type
MOZ	19	Tissue
TEM	18	Dung
NDU	20	Dung
PHI	24	Dung
TSH	19	Dung
MNE	14	Dung
Total	**114**	

Unfortunately for NDU no GPS coordinates were available to use as starting locations. Instead the game rangers of NDU indicated where suni were last observed and these observations were used as starting points. Sample collection in the field was conducted in the exact same manner as discussed earlier. Once enough samples were collected at a site the next sampling site was identified/chosen by moving 200-400 m further down the road searching for samples on both sides. This method was conducted over 5.53 km.

The collection at PHI and MNE on the other hand was done in a different manner than above; instead three different sampling sites of approximately 600 m x 600 m were sampled. At each site, the area was divided into 16 grids of 150 m x 150 m, of which each grid was sampled for 45 min, searching for the freshest samples available. In summary, each grid was sampled every 24 hours, resulting in each grid being sampled at least three to four times.

During the collection of dung samples, pellets were chosen according to freshness, as well as size. The appearance of suni dung is very similar to other small antelopes' dung e.g. red duiker (*Cephalophus natalensis*), in which case they only differ in pellet size. This can be observed in Fig. 2.2. Pictures, as well as size, freshness and locality of each sample were recorded. Throughout sampling sterilized tweezers were used to collect the suni dung pellets (five to ten) and placed into either empty collection tubes or tubes containing 70% EtOH. All samples were kept in an icebox until it could be stored in a freezer.

Fig. 2.2 Comparison of different dung samples, a) suni dung pellets and b) red duiker dung pellets

Tissue samples from MOZ in the form of ear clippings were provided by individual trophy hunters (qualified to shoot suni under legal permit conditions at concessions). Samples were collected by making use of sterilised blades and tweezers in between each tissue sample collected and placed into 5 ml tubes containing coarse salt and stored at room temperature. Sample collection was done with ethical clearance given by the Ethics Committee of the University of Pretoria (EC040-16).

From all of the above it could be noted that there were multiple samples collected in the different areas mentioned previously. In this study where different extraction protocols were tested I made use of 95 dung samples and 19 tissue samples. The different extraction protocols used are discussed in more detail below.

Sample preparation

Each type of sample required different extraction protocols to be performed. For dung samples two different extraction protocols were compared in order to identify which kit provided the best quality DNA and at what conditions each kit performed optimally. Once the best kit was identified, that kits' protocol was optimised so that the specific type of samples collected during this study could be extracted with the aim of obtaining the highest possible quality and quantity DNA. This is important

to note since the majority of samples collected in this study were not the freshest, which could as a result cause difficulty in downstream steps such as PCR. Hence, it was very important to work carefully with each specific sample.

Preparation of samples for extraction varied since the samples were collected and stored in different ways (as explained above). Samples stored in 70% EtOH were placed in a petri dish covered by a new piece of paper towel. This step was included to enable the EtOH to evaporate and the dung pellets to dry. Samples not collected in EtOH did not undergo this step. In general, two pellets were used for each extraction. In some instances where the pellets were a lot bigger, only one pellet was used. Epithelial cells were then recovered from the pellets by scraping the outside layer and weighing the scrapings (approximately 60-70 mg) before placing them in an Eppendorf tube.

All extractions were performed in a sterilized environment with no other lab procedures taking place at the bench other than extractions. Between extractions the bench as well as the gloves were cleaned with 70% EtOH. All extractions were performed making use of tweezers and scalpel blades that were sterilized with 96% EtOH and flamed between samples. A new separate petri dish was used for each new extraction to ensure that no cross contamination of samples could take place. All these steps were included to lower the risk of contamination.

DNA extraction from dung samples

The first kit tested was the QIAamp DNA mini stool kit (Qiagen). This is one of the most utilised kits for dung extractions, producing credible results. These results include high quality and quantity DNA yield, but the use of this kit can be time consuming and expensive. The second kit tested was the NucleoSpin DNA Stool kit (Macherey-Nagel). This kit is still very new on the market and is also designed with specialised protocols particularly focussing on work done on different types of dung samples e.g. origin (carnivore or herbivore) and type (dry or medium water content). The next section focuses on how each kit was tested and what alterations had to be made to each kit for optimal results.

QIAamp DNA mini stool kit: Samples was prepared as above, and DNA was extracted from the pellets following the manufacturer's instructions with only one modification. An additional overnight incubation step was added after the pellet scrapings were placed into a 2 ml Eppendorf tube and 1.6 ml of stool lysis buffer (Buffer ASL) was added. These samples were then vortexed continuously for 1 min until the whole sample was thoroughly homogenized and placed in the incubator overnight at 65°C. Subsequent steps indicated problems regarding the DNA concentration (low) and changes were made. More specifically in cases where extractions were already performed indicating very low concentrations, decisions were made to rather spin the samples down and evaporate them at the same time using the Savant Svc-100h SpeedVac Concentrator (Gemini BV). This was done to increase their final concentration and finally eluting the pelleted DNA in 50 µl of water (Sabax). Later samples were only eluted in 60 µl of elution buffer (EU buffer) instead of 200 µl.

Even though the above alterations made a difference during extractions, problems were still occurring. Consequently, other alterations as suggested by Costa et al. (2017) were implemented. These included the addition of Proteinase K, as well as additional wash steps during the extraction procedure. These steps were included to reduce the amount of inhibitors present in the extracts and to increase the amount of DNA recovered. The final protocol followed with the QIAamp DNA mini stool kit including all the steps and alterations that were made is illustrated in Fig. 2.3.

Once the first round of extractions was completed, samples were analysed on a 1% agarose gel followed by quantification of the nucleic acid concentration and purity ratios using the NanoDrop 1000 (Thermo Fisher Scientific) (Desjardins and Conklin 2010). Nanodrop readings are not very accurate with degraded samples thus, it was decided to amplify the cytochrome *b* (cyt *b*) gene to test whether DNA could be successfully amplified. The PCR consisted of a 25 µl reaction volume, containing the following: 1X PCR Buffer (Separation Scientific), 2.5 mM $MgCl_2$ (Separation Scientific), dNTP mixture (200 µM of each) (Promega), 0.1 µg/µl BSA (Roche diagnostics), 1 U of Super-Therm *Taq* DNA Polymerase (Separation Scientific), 10 pM of each suni cyt *b* F107 and suni cyt *b* R573 primer, 3 µl of template DNA and Sabax® water (Adcock Ingram) to the final volume of 25 µl. These PCR reactions were then carried out in the ABI 2720 Thermo Cycler with the following conditions: 94°C for 5 min, 38 cycles of 94°C for 30 sec, 53°C for 30 sec, 72°C for 30 sec, followed by a final extension of 72°C for 7 min and a holding step at 4°C. Amplified PCR products were then analysed on a 1% agarose electrophoresis gel using GelRed (Biotium, Anatech Instuments) staining along with a 100 bp DNA ladder (The Scientific Group).

Nucleospin® DNA stool kit: DNA was extracted from the dung pellets according to the manufacturers' instructions including some minor alterations. For this study in particular I made use of the protocol which was specialised for samples that are very hard and dry. For each extraction the pellets' outer scrapings were placed into a NucleoSpin® bead tube and 500 µl of lysis buffer (Buffer ST1) and 500 µl of distilled water (Sabax®) were added. These tubes were then shaken horizontally for 2-3 seconds to ensure that the sample and lysis buffer was completely mixed. Samples were then placed into an incubator at 70°C for 5 min followed by homogenisation of the samples. This was done by agitating the samples in the vortex adapter (Multi-tube holder of Scientific Industries) for 10 min at room temperature. From this step on the remainder of the manufacturer's instructions were followed. In the final step 75 µl of elution buffer (SE buffer) was added to the column and incubated at room temperature for 30 minutes before being centrifuged. The volume of elution buffer was later reduced to 40 µl to increase the concentration of the DNA obtained. Once the extractions were completed, samples were analysed and checked with agarose gels, Nanodrop and PCRs similar to that of the previous kit. In addition, each sample was also quantified using the Qubit ® 2.0 kit (Invitrogen) following the manufacturers' recommendation.

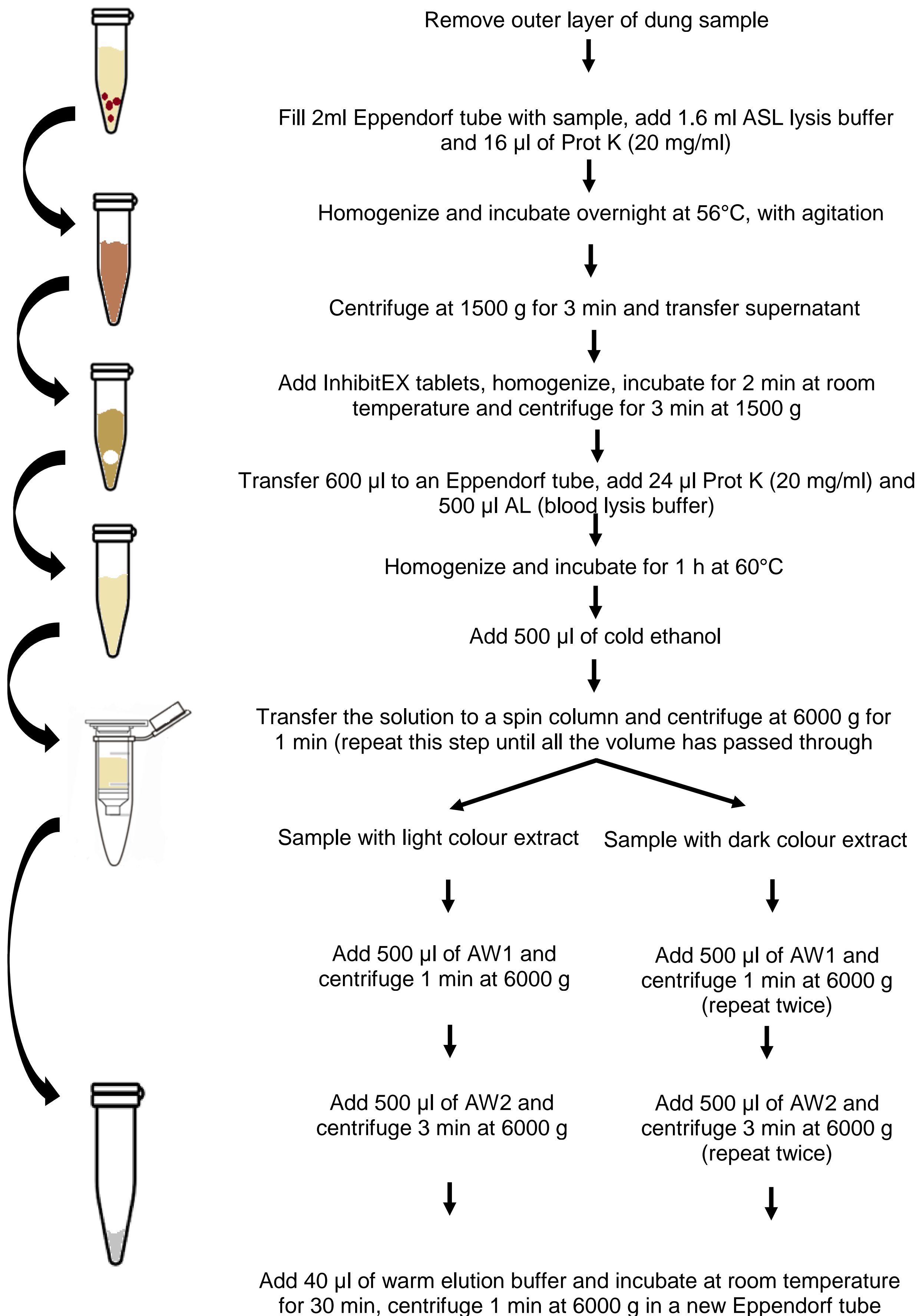


Fig. 2.3 Extraction protocol using the QIAamp DNA mini stool kit from Qiagen. Adapted from Costa et al. (2017)

In summary, this study was mainly focussed on degraded samples, therefore only working with the two above mentioned extraction kits. Even though both kits yielded DNA of good quality and quantity, it was decided to continue with the Nucleospin DNA stool kit since it was the least altered, fastest and most cost-effective kit to use.

DNA extraction for tissue samples

For tissue samples DNA was extracted from 25 mg of tissue using the Animal Tissues Spin-Column Protocol (Qiagen) as recommended by the manufacturer. DNA was eluted in 200 µl of elution buffer (Buffer AE) and quantified using the Qubit ® 2.0 kit following the manufacturer's recommendation.

Species identification using mtDNA

To confirm the species of the sample (tissue and dung), a small section of mtDNA (476 bp of cytochrome *b*), was amplified and sequenced using the primers indicated below (Table 2.2). These primers were designed and tested during a pilot study (Reitmann 2008) which showed successful identification of suni and red duiker. PCR amplifications were carried out in 25 µl reactions containing 1X PCR Buffer, 2.5 mM $MgCl_2$, dNTP mixture (200 µM of each), 0.1 µg/µl BSA, 1 U of Super-Therm *Taq* DNA Polymerase, 10 pM each of forward and reverse cyt *b* primer (Table 2.2), 5 µl of template DNA and Sabax® water added to the final volume. PCR reactions were carried out in the ABI 2720 Thermo Cycler with a denaturing step at 94°C for 5 min. Samples were then subjected to 38 cycles of 94°C for 30 sec, 53°C for 30 sec and 72°C for 30 sec. It was then followed by a final extension step of 72°C for 7 min and a holding step at 4°C. PCR products were analysed on a 1% agarose electrophoresis gel using GelRed staining along with a 100 bp DNA ladder.

Table 2.2 Primers used for mtDNA cytochrome *b* gene amplification of *Nesotragus moschatus*

Primer name	Primer sequence (5'-3')
Suni cyt *b* F107	TCCTAGGCATCTGCCTAATCC
Suni cyt *b* R573	TGCTGCGATGATAAATGG

PCR reactions with optimized products at 467 bp were then precipitated by adding 90 µl 96% EtOH, 10 µl distilled water and 0.05 M NaAc to the PCR product. It was then followed by a centrifugation step (20 000 g for 20 min) and removal of supernatant. The remaining pellet was then washed with 90 µl 70% EtOH and centrifuged for 10 min at 20 000 g. The supernatant was then removed, the pellet left to dry at room temperature and eluted in 20 µl distilled water once it was dry. Precipitated samples were visualised on a 1% agarose gel. Diluted cycle sequencing reactions (0.5X) were set up with the BigDye™ Terminator v.1.1 and v.3.1 5X Sequencing Buffer (Thermo Fisher Scientific) following manufacturer's instructions. Cycle sequencing reactions of 10 µl were set up using 3 µl of precipitated PCR product. Cycle sequencing reactions were carried out on the ABI 2770 Thermo Cycler following conditions stated in the users guidelines. Products were then run on the ABI 3500xl automated sequencer (Applied Biosystems). Both forward and reverse cyt *b*

sequences of individual samples were aligned in CLC Main Workbench v.8.1 (Qiagen Bioinformatics) and compared against the available data on NCBI using BLAST (Altschul et al. 1990) to confirm the likely species identity of the sample.

Microsatellite marker selection

No microsatellite primers specific for suni is available. For this reason, I decided that it would be the best alternative to make use of markers that were previously tested on suni material during a pilot study in 2008. This pilot study in particular selected markers that amplified consistently and were found to be polymorphic in other bovids such as caprine, gazelle, impala and ovine. I then tested 16 of these previously tested markers (including one sex marker) on both dung and blood samples of which only 11 amplified (including one sex marker).

PCR amplification reactions (final volume of 5 µl) consisted of 0.6 X multiplex mix (Qiagen Multiplex PCR Master Mix), 1 pM of each forward and reverse primer and 2.5 µl of template DNA. Using the ABI 2720 Thermo Cycler the PCR reactions were carried out with the following conditions: 95°C for 15 min, 40 cycles of 94°C for 1 min, 60°C for 1 minute and 30 sec followed by a final holding step at 4°C. PCR products were then examined on a 1% agarose gel as before and further analysed for amplification success by Genescan analysis. Making use of the ABI 3500xl automated sequencer GENESCAN software all PCR amplicons were separated by size with the GENESCAN LIZ 500 (ABI™/ Life Technologies) marker. Subsequently only five (four microsatellites and one sex marker) of the eleven markers tested were selected for further analysis according to criteria such as amplification success and presence of visible peaks after fragment separation.

In addition to this other microsatellite markers available at the Veterinary Genetics Laboratory (VGL), at the University of Pretoria, were also tested on five freshly extracted dung extracts and one blood sample. These primers that were designed for different bovids including cattle, impala, wildebeest, sable antelope, goat and sheep were previously tested in other species (eland, gazelle, gemsbok roan, buffalo), and were found to amplify consistently with different polymorphism levels (Miller et al. 2016). A total of 43 microsatellite markers (including one sex marker) that were polymorphic in other bovids were tested in the suni at the VGL. In order to obtain the best possible markers for this study three separate rounds of tests and evaluation were found necessary. In the first round of testing markers were amplified in multiplex following the exact same PCR set up as in Miller et al. (2016). PCR reactions were carried out in the ABI 2720 Thermo Cycler with the following conditions: 95°C for 15 min, 35 cycles of 95°C for 45 sec, 60°C for 45 sec, 72°C for 1 min, followed by 72°C for 60 min and a final holding step at 4°C. These products were then checked and analysed as above. After the first round of testing I reduced the 43 markers to 19 (only microsatellites) according to amplification success and peak height. All the marker's peaks had to be higher than 100 fluorescent units in order to be eligible for further testing.

These 19 markers were then amplified for a second time in which the PCR amplifications were carried out in 10 µl reactions consisting of 0.6 X multiplex mix, 2 µl of template DNA, Sabax® water

to the final volume and forward and reverse primer (final concentrations in Table S2.2). PCR reactions were carried out with the following conditions: 95°C for 15 min, 35 cycles of 95°C for 45 sec, 60°C for 45 sec, 72°C for 1 min and 30 sec, followed by 72°C for 60 min and a final holding step at 4°C. Successful amplicons were then analysed as above and only markers that presented signs of polymorphism (two or more alleles) were kept for further analysis. This reduced the 19 markers to 14. A third and final round of testing was performed as in round two but on eight additional dung samples. As a result, only nine of the 14 microsatellite markers were selected according to the polymorphism criteria.

Lastly three other microsatellite markers were added to the list resulting in a total of 16 dinucleotide microsatellites (four from the pilot study in 2008, nine from VGL and three from a second pilot study in 2013) and one sexing marker for further analysis. These three markers were previously investigated during a pilot study that took place in our laboratory in 2013. These markers selected showed a lot of potential since they amplified consistently and were highly polymorphic. The final list of microsatellite markers can be observed in Table 2.3. Once the list of markers was finalised, multiplex manager v.1.2 (Holleley and Geerts 2009) was used to create two groups of multiplexes. The first panel consisted of eight markers and the second panel of nine (Fig. 2.4). Take note that the fluorescence labels of each marker in Table 2.3 were the final colours used when ordered, not the labels shown in Fig. 2.4. To observe all the markers tested and used in this study please refer to Table S2.3.

Microsatellite- and sex marker optimization

The labelled microsatellite markers as well as the sex marker were tested separately on tissue extractions to ensure optimal amplification. This was done by setting up PCRs (final volume 10µl) consisting of 1X PCR Buffer, 1.5 mM $MgCl_2$, dNTP mixture (200 µM of each), 0.2 µg/µl BSA, 2.5 units of Super-Therm *Taq* DNA Polymerase, 10 pM of each forward and reverse primer, 2 µl of template DNA and Sabax® water added to the final volume. The selected amplification conditions consisted of 95°C for 7 min, 35 cycles of 95°C for 30 sec, locus-specific annealing temperature for 30 sec, 72°C for 30 min, followed by 72°C for 7 min and a final holding step at 4°C. Annealing temperatures recommended by the labelled primers' manufactures (ABI™/ Life Technologies) were used and therefore each marker had a different annealing temperature, indicated in Table S2.3.

From the above PCR protocol, not all of the markers amplified initially. To improve amplification success of these markers, gradient PCRs were set up for each in which case the initially tested annealing temperatures above were both decreased and increased by two degrees. Once this alteration was made some markers still did not amplify optimally. In order to improve specificity and the strength of amplification, temperatures were then increased further until the optimal temperature was identified.

Table 2.3 Final set of dinucleotide microsatellite markers and one sexing marker (Amelogenin) used in this study

Marker	Forward Primer 5'-3'	Reverse Primer 5'-3'	Reference	Fragment size range	Label	Panel
CT07	CTACCTGGGAAACCCATAT	GTGTCTTTGTTGGTTTCTGCCATACAA	(Røed et al. 2011)	110-125	FAM	2c
INRA006	AGGAATATCTGTATCAACCTCAGTC	CTGAGCTGGGGTGGGAGCTATAAATA	(Vaiman et al. 1992)	100-130	VIC	2c
SRCRSP24	AGCAAGAAGTGTCCACTGACAG	TCTAGGTCCATCTGTGTTATTGC	(Yeb et al. 1997)	140-160	VIC	1b
ETH10	GTTCAGGACTGGCCCTGCTAACA	CCTCCAGCCCACTTTCTCTTCTC	(Toldo et al. 1993)	220-230	PET	2c
BM2113	GCTGCCTTCTACCAAATACCC	CTTCCTGAGAGAAGCAACACC	(Bishop et al. 1994)	140-156	NED	2c
TGLA53	GCTTTCAGAAATAGTTTGCATTCA	ATCTTCACATGATATTACAGCAGA	(Barendse et al. 1994)	156-170	VIC	2c
SPS115	AAAGTGACACAACAGCTTCTCCAG	GTGTCTTAACGAGTGTCCTAGTTTGGCTGTG	(Mommens et al. 1998)	260-280	FAM	2c
I206	ATTAGGAAAAGCAATGTGAATGG	GTGTCTTCACTCCTGTATTCTGCCTGG	(Huebinger et al. 2006b)	150-165	PET	2c
SRCRSP8	TGCGGTCTGGTTCTGATTTCAC	CCTGCATGAGAAAGTCGATGCTTAG	(Bhebhe et al. 1994)	230-240	VIC	1b
OARCP26	GGCCTAACAGAATTCAGATGATGTTGC	GTCACCATACTGACGGCTGGTTCC	(Ede et al. 1995)	140-150	NED	1a
ILST87	AGCAGACATGATGACTCAGC	CTGCCTCTTTTCTTGAGAGC	(Kemp et al. 1995)	130-140	PET	1a
SPS113	CCTCCACACAGGCTTCTCTGACTT	CCTAACTTGCTTGAGTTATTGCCC	(Roeder et al. 2001)	143-155	FAM	1a
TGLA122	CCCTCCTCCAGGTAAATCAGC	AATCACATGGCAAATAAGTACATAC	(Georges and Massey 1992)	119-155	FAM	2d
INRA63	ATTTGCACAAGCTAAATCTAA CC	AAACCACAGAAATGCTTGGAAG	(Vaiman et al. 1994b)	180-210	NED	1b
BM757	TGGAAACAATGTAAACCTGGG	TTGAGCCACCAAGGAACC	(Bishop et al. 1994)	180-190	PET	1a
F10	TGTCCAGCAGCTTCACCATTACGCC	GCCTTCATGATCTTGTCGATCCACT	(Dietz et al. 1992)	215-245	VIC	2d
AMELOGENIN	CAGCCAAACCTCCCTCTGC	CCCGCTTGGCTTGTCTGTTGC	(Weikard et al. 2006)	215; 260	FAM	1a

The number (1 and 2) refers to the two different panels designed. The letters a, b, c and d indicate the further division of the panels when it was necessary.

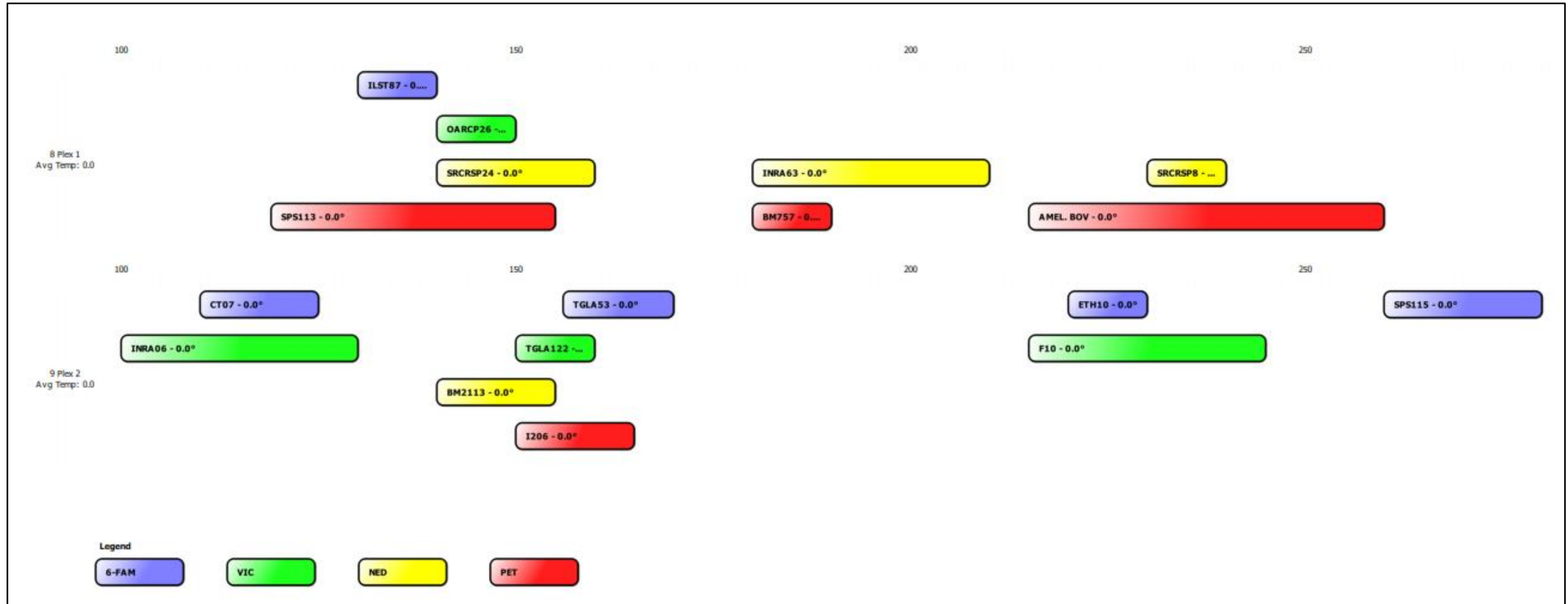


Fig. 2.4 Designed multiplex panel for 16 dinucleotide microsatellite markers and one sexing marker (Amelogenin) using Multiplex Manager v.1.2

The initially planned panels above (Fig. 2.4) were both separated into two separate panels (Table 2.3) thus, resulting into four different panels, each panel with their own separate annealing temperature. PCRs were then set up as stated above with the exception that 10 pM of each forward and reverse primer present in each respective panel was added. Amplification conditions were the same as above except for the annealing temperatures which were different for each panel: panel 1a 60°C, panel 1b 54°C, panel 2c 52°C and panel 2d 60°C.

Once PCR amplification was completed, plates for Genescan analysis was set up according to manufacturer's instructions using the GENECAN LIZ 500 standard and changing the amount of PCR product added. This was performed in two different ways. The first setup only consisted of PCR products produced in single PCR reactions. Each panel's markers (PCR products) were pooled together. In each event only 1 µl of each markers' PCR products were added to the plate. For the second setup, which only consisted of PCR products produced by multiplexes, panel 1a and 1b were pooled together and the same for panel 2c and 2d. This was done by only adding 1 µl of each panels' PCR products e.g. 1 µl of panel 1a and 1 µl of panel 1b being added to the plate. Each sample was then analysed by Genescan analysis making use of the ABI 3500xl automated sequencer and GENESCAN software as above.

From the results obtained it was noted that the first setup, which consisted of the markers amplified in single PCR reactions pooled together, performed a lot better than the second setup based on the number of markers that amplified successfully (multiplexes pooled together). This led to the decision to continue with the single PCR setups instead of multiplexes. From this step forward, single PCR reactions were set up but resulted in inadequate amplification and therefore had to be optimized further. This was done by using multiplex mix instead of the above-mentioned polymerase with the same amplification conditions as earlier. These PCRs consisted of 0.6 X multiplex mix, 2 µl of template DNA, Sabax® water to the final volume and 10 pM of each forward and reverse primer.

Coupled with this, further optimization on dung samples was also necessary. It initially started with the exact same steps as was optimized for tissue samples. In the case where certain markers did not amplify as was necessary, changes were made such as temperature increase and decrease, as well as replacing the polymerase with multiplex mix were performed. This was a long tedious process and since the markers' PCR setup and conditions differed a lot, a decision was made to test the effect of Hot Start *Taq* (Invitrogen™ Platinum™ II *Taq* Hot-Start DNA polymerase, Thermo Fisher Scientific).

In short, PCR reactions were set up using the Invitrogen™ Platinum™ II *Taq* Hot-Start DNA polymerase protocol as stated in the manual. The PCR conditions consisted of 94°C for 2 min, 35 cycles of 94°C for 15 sec, locus-specific annealing temperature for 15 sec, followed by 72°C for 30 min and a final holding step at 4°C. This was performed in the SimpliAmp™ Thermal Cycler (ABI Thermo Fisher Scientific). For the markers which did not amplify immediately with this protocol, PCR reactions were performed again but now using the suggested annealing temperature in the protocol. Again, if amplification still did not occur, a gradient PCR was set up between 50°C and 62°C at 2°C

intervals. This was done so that the optimal binding temperatures could be identified. From this step onwards, microsatellite markers with the same binding temperatures, as well as multiplex panel were amplified together. Hence, reverting back to the multiplex panels only for the markers with the same binding temperatures.

Results and discussion

Sample quality

In the past three decades gNIS has developed substantially, creating opportunities to do research on wildlife not previously possible (Kovach et al. 2003; Chiou and Bergey 2018; Ramón-Laca et al. 2018). Currently gNIS is being implemented on various animals, big and small (Ferreira et al. 2018), overcoming hurdles such as invasive sampling and the cost connected to it (Ferreira et al. 2018). Although it has been a developing field, problems are still experienced such as low quality DNA yield (Flagstad et al. 1999; Rivière-Dobigny et al. 2009; Chiou and Bergey 2018; Bourgeois et al. 2019) and PCR amplification (Fernando et al. 2003; Ramón-Laca et al. 2018), all due to the samples type (Hájková et al. 2006), age (Kovach et al. 2003; Piggott 2004), collection climate (Piggott 2004; Hájková et al. 2006), season (Kovach et al. 2003; Piggott 2004), species (Kovach et al. 2003) and their diet (Murphy et al. 2003). These studies came to the conclusion that the freshest samples should be used in studies (Kovach et al. 2003; Piggott 2004; Lampa et al. 2007; Rivière-Dobigny et al. 2009; Reddy et al. 2012). Due to this evaluation, for this study only the freshest samples were collected daily for PHI and MNE, however, this was not always possible for the other reserves. Thus, the same problems as was stated in the above-mentioned articles were experienced.

Moreover, on agarose gels differences in quality and concentration could be observed in the DNA extracted from the different reserves (Fig. 2.5). Yet, the nanodrop readings were very low and similar for all making it uninformative. Furthermore, TEM samples (all old) presented clear solid bands with very little smears present whereas PHI and MNE (all fresh) presented with very faint bands and large smears. Cyt *b* PCR reactions (Fig. 2.6) were also set up for all the extracts shown in Fig. 2.5. Samples from PHI and TEM both yielded the same difficulty and success during amplification. Thus, despite the lower quality DNA extracts from PHI, samples were still good for amplification. Genotyping results obtained from eight microsatellite markers tested on suni dung (Table 2.4) also indicated similar results as another article, showing that sampling methods influence genotyping success (Piggott 2004). Markers were considered successful, once genotypes could be identified. A clear difference can be observed in the genotyping success rate once the type of samples (tissue or dung) are considered at the ≥75% success rate for the markers that amplified. Additionally, I also observed an increase in genotyping success from old (35%) to fresh (62%) samples, similar to what another article (Piggott 2004) has found.

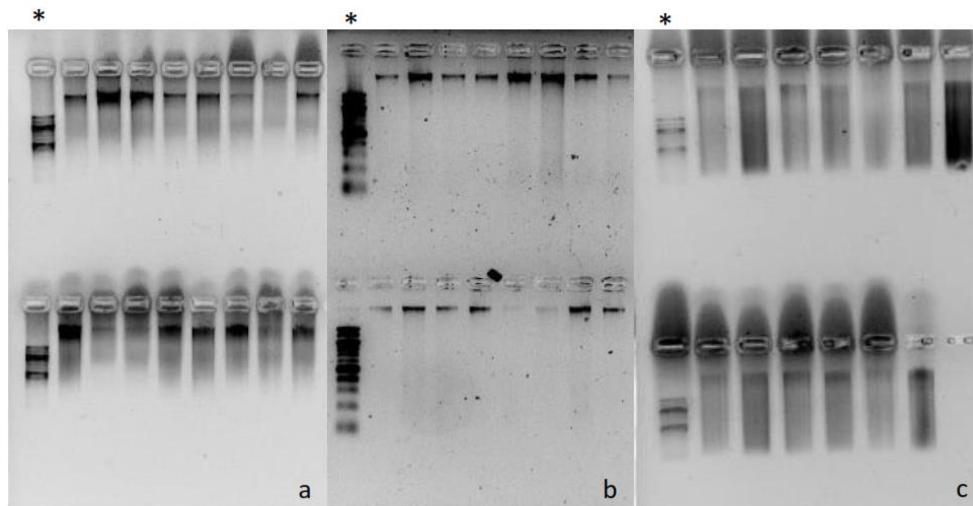


Fig. 2.5 Agarose gels representing DNA extractions of samples from three different nature reserves: a) PHI b) TEM and c) MNE Island using the Nucleospin® DNA stool kit (Macherey-Nagel). *- lanes with 100 bp DNA ladder for both the top and bottom sections of each separate gel

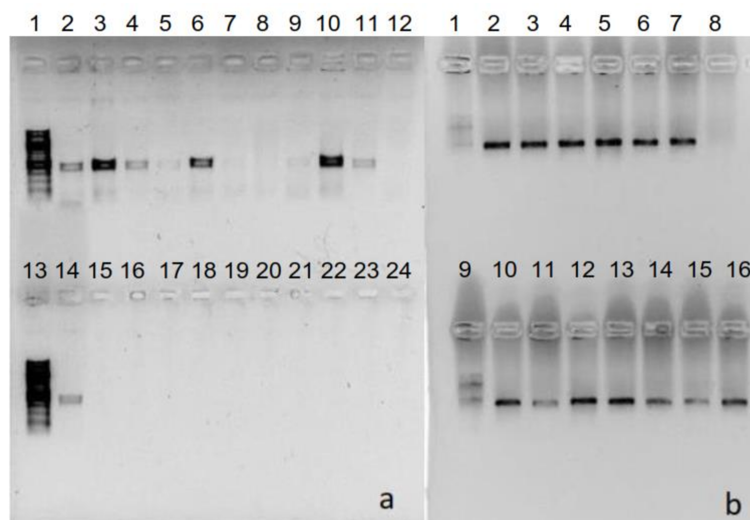


Fig. 2.6 Agarose gels representing amplicons of the cyt *b* gene from three different nature reserves: TEM (lanes a2-7), PHI (lanes a8-10 and a14) and MNE (lanes b2-6 and b10-16). 100 bp DNA ladder = lanes a1, a13, b1 and b9. Positive control = lanes a11 and b7. Negative controls = lanes a12 and b8.

Table 2.4 Comparison of genotyping success over different nature reserves, sample types and freshness using eight microsatellite loci.

Nature reserve	Sample type (state)	*n*	Percentage of individuals of which:	
			≥50% markers amplified	≥75% markers amplified
MOZ	Tissue (Fresh)	19	100	100
MNE	Dung (Fresh)	12	100	58.33
PHI	Dung (Fresh)	12	100	66.67
TEM	Dung (Old)	14	92.86	42.86
NDU	Dung (Old)	11	90.91	27.27

n- number of individuals

This study has shown that sample collection does play an important role as was expected (Kovach et al. 2003; Piggott 2004; Hájková et al. 2006; Lampa et al. 2007). Samples that were collected in a fresh state definitely had a higher chance of successful DNA extraction and subsequent amplification. Even though the PCR amplification of the cyt *b* gene did not clearly support this, overall it could be concluded that the fresher the sample is at collection, the better the chances are for the sample to amplify with ease and to obtain more accurate results. Older samples might require more repeats yet, accurate and complete results can still be generated.

DNA extractions

Comparison between kits

From multiple articles it has been pointed out that degraded dung samples can yield lower quality and quantity DNA during extraction (Kovach et al. 2003; Hájková et al. 2006; Wultsch et al. 2015). Due to this, many different kits have been developed specifically for the use of dung samples. Additionally articles have also compared some of these available kits (Bhagavatula and Singh 2006; Lampa et al. 2007; Puechmaille et al. 2007) however, none used the Nucleospin® DNA stool kit from Macherey-Nagel with the focus on degraded samples. Coupled with this, many studies have been using a variety of dung extraction kits yet, all were utilised on some of the freshest samples (Van Vliet et al. 2008; Bowkett et al. 2009; Reddy et al. 2012; Ramón-Laca et al. 2014; Ramón-Laca et al. 2015; Ramón-Laca et al. 2018; Proença-Ferreira et al. 2019), with limited cases including highly degraded samples (Lampa et al. 2007; Bourgeois et al. 2019). The Qiagen kit is one of the most utilised kits to date (Kovach et al. 2003; Wehausen et al. 2004; Costa et al. 2017; Chiou and Bergey 2018; Bourgeois et al. 2019). Additionally, the results obtained in a pilot study previously done in our laboratory (Reitmann 2008) also supported this choice with an additional incubation step, similar to what was proposed by Costa et al. (2017), who improved the protocol further. For these reasons this study compared the Nucleospin® DNA stool kit (Macherey-Nagel) and the QIAamp DNA mini stool kit (Qiagen) with the use of highly degraded dung samples.

At first the DNA extracts (Fig. S2.1a) produced from the Qiagen kit indicated a clear band of high molecular weight however, the PCR bands were very faint (Fig. S2.1b). These were improved after concentrating the samples (Fig. S2.1c) yet, some samples still didn’t amplify, even when BSA was added. Nevertheless, with the use of additional wash steps (Costa et al. 2017) the quality of DNA extracted improved, as was expected. The Macherey-Nagel kit on the other hand only needed one alteration (reduction in elution buffer used) to obtain good quality DNA. In fact, this alteration was only necessary in a few instances. In Fig. 2.7 the comparison between the Qiagen kit (including incubation and wash steps) and the Macherey-Nagel kit (reduced elution buffer) can be observed. Here both kits performed well however, the low-quality samples could explain why the Qiagen kit didn’t perform as in other studies (Costa et al. 2017; Bourgeois et al. 2019). This could be said since the Qiagen kit has mainly been used on the freshest samples (Ramón-Laca et al. 2015; Ramón-Laca et al. 2018; Bourgeois et al. 2019) and has been designed for the general use of extracting

from dung samples. This is in contrast with the Macherey-Nagel kit that has specifically designed protocols for different types of samples (old versus fresh). This may explain why the Macherey-Nagel kit required fewer alterations to generate good quality DNA instantaneously with highly degraded samples.

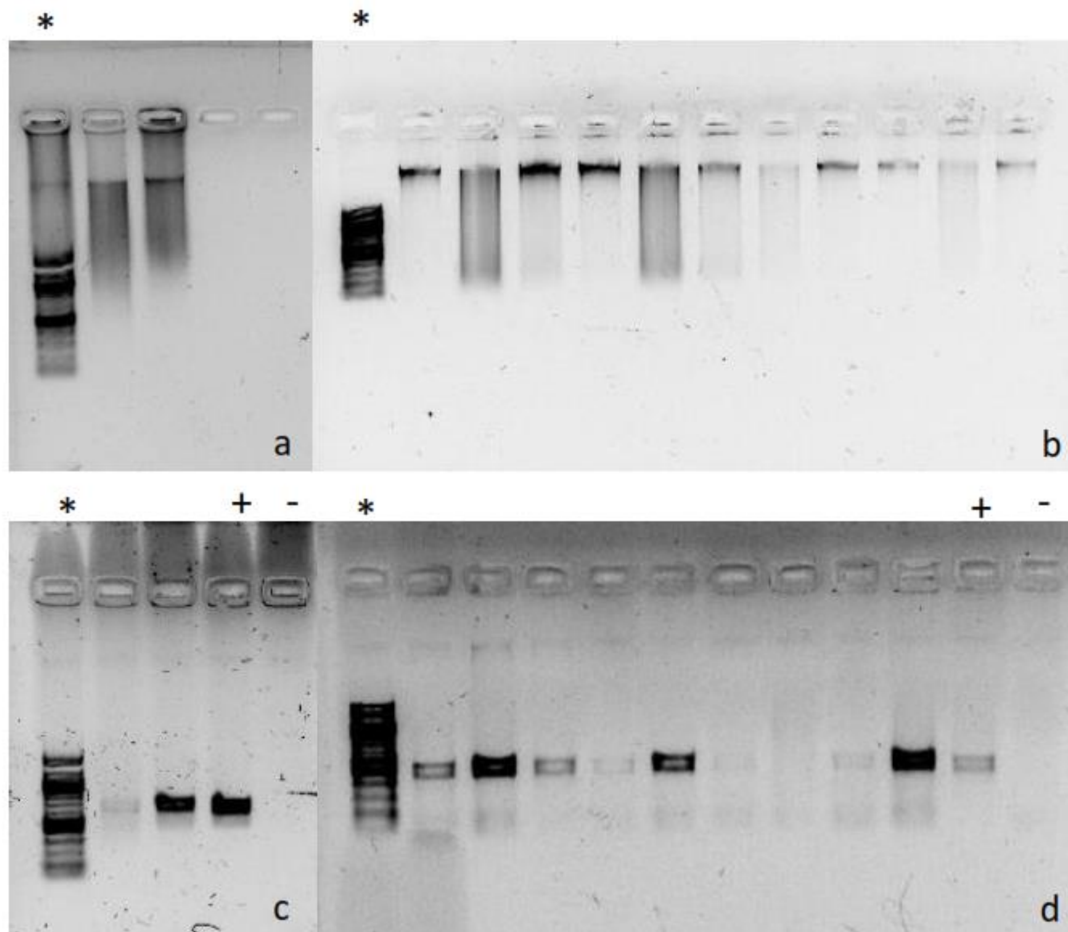


Fig. 2.7 Agarose gel images of DNA extractions and cyt *b* PCR amplifications illustrating a comparison between the NucleoSpin DNA Stool kit (Macherey-Nagel) and the QIAamp DNA mini stool kit (Qiagen). a) DNA extracts from the Qiagen kit, b) DNA extracts from Macherey-Nagel kit, c) PCR amplicons from DNA extracted with Qiagen kit, d) PCR amplicons from DNA extracted with Macherey-Nagel kit. *- lanes with 100 bp DNA ladder, positive (+) and negative (-) control.

DNA quantification

Even though the Macherey-Nagel extraction kit performed well, some samples still did not amplify the first time during PCR amplification. In order to resolve this samples were both quantified with a Nanodrop and Qubit. Initially DNA extracts looked very promising (Fig. 2.7a, b) yet the Nanodrop readings (Table S2.4) were not as expected (Desjardins and Conklin 2010). Instead DNA concentrations were low and ranged between 29-33 ng/µl (Table S2.2). Coupled with this the 260/280 ratios were adequate, but the 260/230 ratios were a lot lower than expected (Desjardins and Conklin 2010). These ratios could indicate that other organic contaminants might be present (Desjardins and Conklin 2010). This would be expected given the diet of species of interest (Lawson 1989), the starting material and how these two factors can affect PCR amplification (Fernando et al. 2003; Murphy et al. 2003).

Due to the inconsistent PCR amplifications and Nanodrop values not being informative enough, the relationship between Nanodrop and Qubit values, respectively, with the PCR strength of the same samples were investigated (Table 2.5). Here a positive correlation between the Nanodrop and Qubit values with PCR strength was expected however, no apparent relationship was observed (Fig. 2.8). As a result, DNA extractions were immediately followed by PCR amplification instead of quantifying the extracts. For this reason, DNA concentrations could not be standardised specifically for the use in PCR reactions. Hence, PCR reactions were set up with a pre-established volume of DNA instead of using DNA concentration, which generally generated bands at the expected sizes in this study.

Table 2.5 Comparison of Nanodrop values, Qubit values and PCR amplification strength

Sample ID	DNA concentration (ng/µl)		PCR strength
	Nanodrop value	Qubit values	
210	8	4.32	3
216	12.9	1.82	1
217	9.5	4.76	2
219	19.3	6.46	2
221	13.8	4.46	2
225	12	3.88	3

1-Average (very faint PCR bands), 2-high (bright PCR bands) and 3-very high (very bright PCR bands)

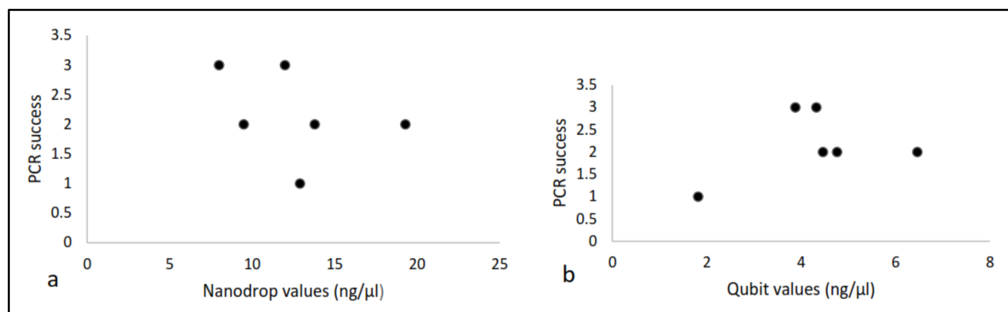


Fig. 2.8 Scatter plots of PCR success and DNA concentration determined by a) Nanodrop and b) Qubit.

Species identification

Since this study made use of dung samples, it was important to determine whether the samples collected were of the correct species. A few previous articles did suggest samples to be genetically tested since there is an increased chance of collecting other species' dung (Van Vliet et al. 2008; Ramón-Laca et al. 2014; Spitzer et al. 2019). Moreover, a mitochondrial gene, cytochrome *b* was suggested for the purpose of species identification (Reitmann 2008). No PCR optimization was necessary however, some difficulty was experienced in cases where the samples were older and more degraded, which was expected (Kovach et al. 2003). In the event where samples yielded adequate quality and quantity DNA, amplification occurred successfully producing an amplicon of 467 bp similar to what was found in another study performed on suni (Reitmann 2008) (Fig.2.7c, d).

During the confirmation/sequencing and BLAST step the same observations were made as in other studies focussed on ungulates (Van Vliet et al. 2008; Spitzer et al. 2019). Not only the species of interest (suni) dung was collected, red duiker was also collected due to the similar appearances of the dung in the field. Of all the samples sequenced I was able to assign 81.74% of the samples to a species (suni or red duiker) of which 91.49% was assigned to suni and only 8.51% to red duiker (Table S2.1). The remaining 18.26% could not be assigned to a species due to low quality sequences. Inclusion of the cyt *b* PCR and species identification step did not only identify samples which were adequate for genotyping, but also identified samples which were of the correct species.

Marker selection

For this study microsatellite marker selection and optimization was very important since these markers were intended for use in a population study on suni antelope (see Chapter 3). Only a few microsatellite markers have been tested in suni making use of dung samples (Reitmann 2008). Hence, for this study only the most variable markers (Kolodziej et al. 2012), as advised, had to be identified and optimised (Ogliari et al. 2000). Optimization is not only done to obtain accurate data for this species, but also performed to reduce the amount of time and cost necessary especially since dung samples can cause amplification failure and genotyping error (Chiou and Bergey 2018; Ramón-Laca et al. 2018; Bourgeois et al. 2019). Furthermore heterologous markers could also increase the chances of non-amplification (Barbara et al. 2007; Narina et al. 2011).

The markers tested on tissue samples all amplified very well during the microsatellite marker identification and selection phases. However, some markers did amplify better than others as can be observed in Fig. 2.9. This could be expected due to the cross-species amplification of microsatellites (Barbara et al. 2007; Narina et al. 2011). Even though the intensity of the bands were different for each marker this did not appear to negatively impact genotyping. The electropherogram of marker BM757, with the lightest bands compared to the others, clearly showed that good quality scoring of fragments could still be performed even if the bands were faint (Fig. 2.9a, b). Additionally, dung samples were also included and did not always amplify, allowing identification of only the markers which had a higher chance of amplification. Even though these markers were identified with dung samples during the test phases, it is important to note that this phase was performed with the freshest dung samples, as well as other solutions and polymerases which were not used later in the study. Different polymerases (Wolffs et al. 2004) and fresh dung samples (Kovach et al. 2003; Piggott 2004) can have a big influence on the amplification success.

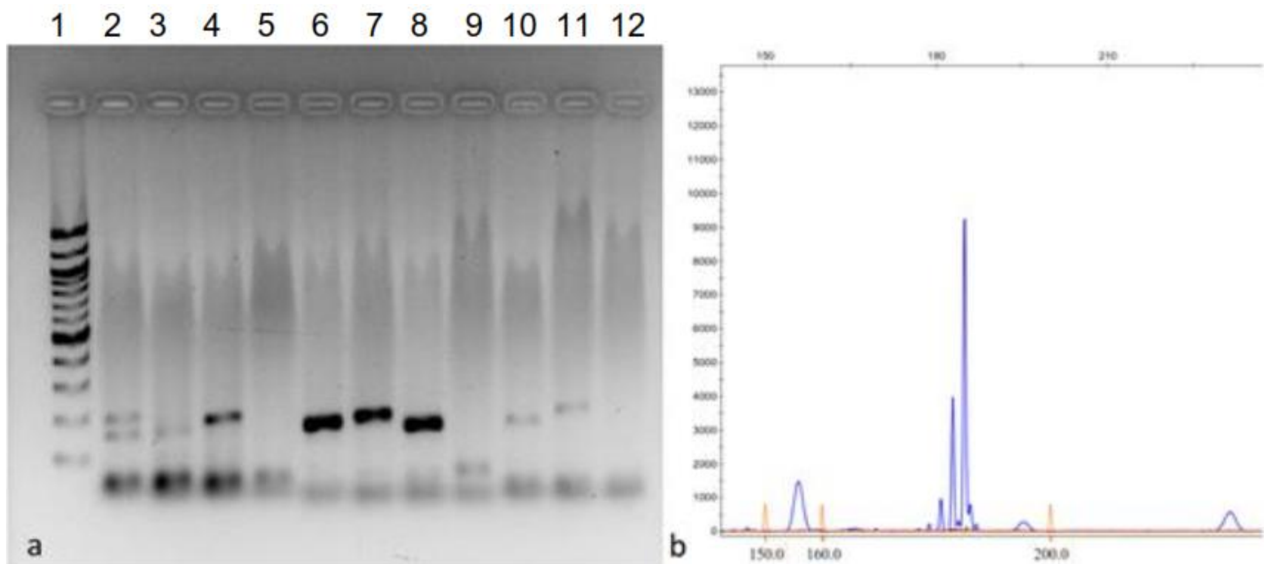


Fig. 2.9 a) Agarose gel comparing the PCR amplification of three microsatellite markers performed on tissue samples. Lane 1: 100 bp DNA ladder, lane 2 and 3: BM415 samples, lane 4: BM415 positive control (blood sample), lane 5: BM415 negative control, lane 6 and 7: SRCRSP24 samples, lane 8: SRCRSP24 positive control, lane 9: SRCRSP24 negative control, lane 10: BM757 sample, lane 11: BM757 positive control and lane 12: BM757 negative control. b) an electropherogram of the sample in lane 11.

Marker optimisation

Optimisation of amplification and genotyping of tissue samples did not take much time. These markers were also amplified in multiplex but did not perform as well as pooled single reactions. Instead markers done in single reactions increased 21-36% of data that was generated compared to multiplexes (Table 2.6), giving a clear indication that further PCR amplifications had to be performed in single reactions especially in dung samples which had a higher chance of non-amplification and genotyping errors (Chiou and Bergey 2018; Ramón-Laca et al. 2018; Bourgeois et al. 2019). The sex marker (Amelogenin) on the other hand amplified the most successfully and consistently out of all markers tested, even when the type of *Taq* was changed.

Table 2.6 Comparison of amplification success in tissue samples for markers amplified in single reactions as well as multiplex reactions. A total of 14 microsatellite markers were evaluated.

Sample	# Markers amplified and scoreable	Percentage of markers amplified and scoreable (%)
Tissue Multiplex	6	43
Tissue Single reactions	11	79
Positive control Multiplex	7	50
Positive control single reactions	10	71

With regards to types of samples used, a clear difference was observed concerning the amplification success as found in other studies (Fernando et al. 2003; Bhagavatula and Singh 2006). All markers amplified more effectively and at a higher intensity for all tissue samples in comparison to dung samples. One such example can be observed for the microsatellite marker BM2113

(Fig. 2.10). All the tissue samples in this figure, as well as the positive control (blood sample) amplified successfully for this marker, with only two dung samples that amplified. Since the exact same PCR conditions were applied for all of these samples, these results confirmed that sample type influence amplification success (Fernando et al. 2003; Bhagavatula and Singh 2006).

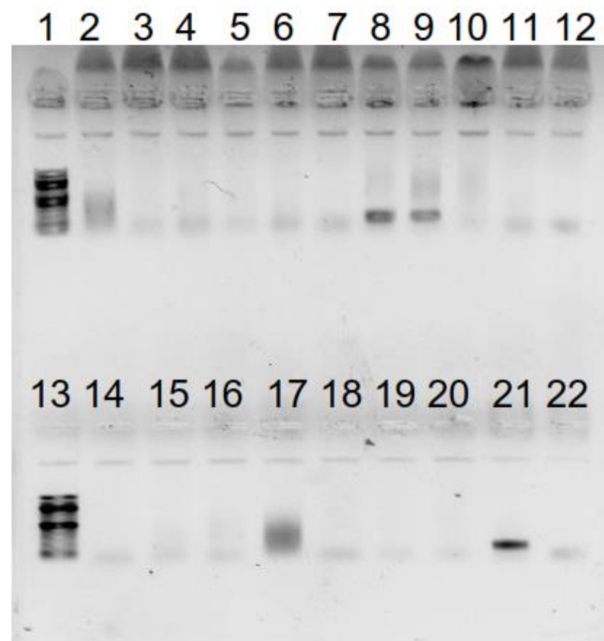


Fig. 2.10 Comparison of PCR amplification of the microsatellite marker, BM2113, in tissue and dung samples. Lanes 1 and 13: 100 bp DNA ladder, lanes 2-7, 10-12 and 14-20: dung samples, lanes 8-9: tissue samples, lane 21: positive control and lane 22: negative control.

In addition to this the results generated here also indicated that the quality of genotyping results can also be affected by sample type (Fernando et al. 2003; Bhagavatula and Singh 2006) (Fig. 2.11). Moreover, in the case of dung samples, identifying the peak shapes of the microsatellite markers were difficult due to stutter, as well as non-specific amplification peaks, not to mention the very low RFUs observed. In contrast for tissue samples each marker's shape could be easily identified with acceptable RFUs. In Fig. 2.11 a difference can be observed in the peak shape as well as the RFU height. The tissue samples had peaks above the size standard whereas the dung samples' peaks were below the size standard, making these difficult to score.

With further investigation, the genotyping results (low RFUs, inconsistent peak shapes and non-specific amplification) indicated that the annealing temperatures optimised in tissue samples could not be optimally used in dung samples. By doing a gradient PCR with Super-Therm *Taq* the ideal annealing temperatures were identified for dung samples (Fig. 2.12). Yet once it was applied on more samples, amplification failure took place even when multiplex mix was incorporated using the newly found annealing temperatures. The incorporation of Platinum Hot Start *Taq* on the other hand significantly increased amplification success especially once it was tested on multiple individuals for all markers. Not only was I able to show that it improved the PCR success in dung samples, I was also able to amplify two markers in the same reaction. However, multiplexing with Platinum Hot Start *Taq* was only considered once markers were optimised at the same annealing temperatures.

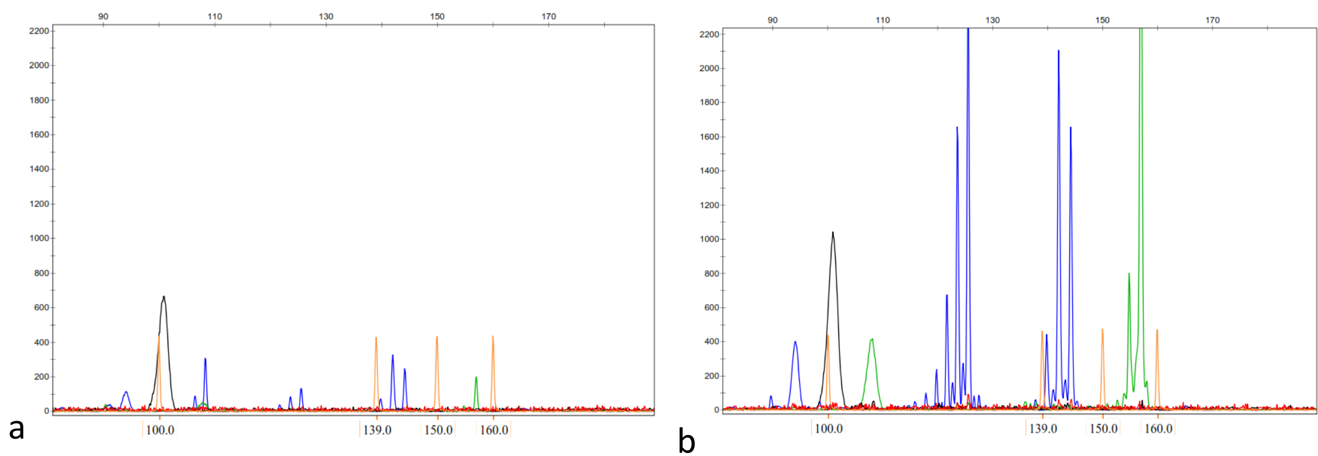


Fig. 2.11 Electropherograms of the marker INRA6 (two blue peaks between 120-150 bp) in a) a dung and b) blood sample

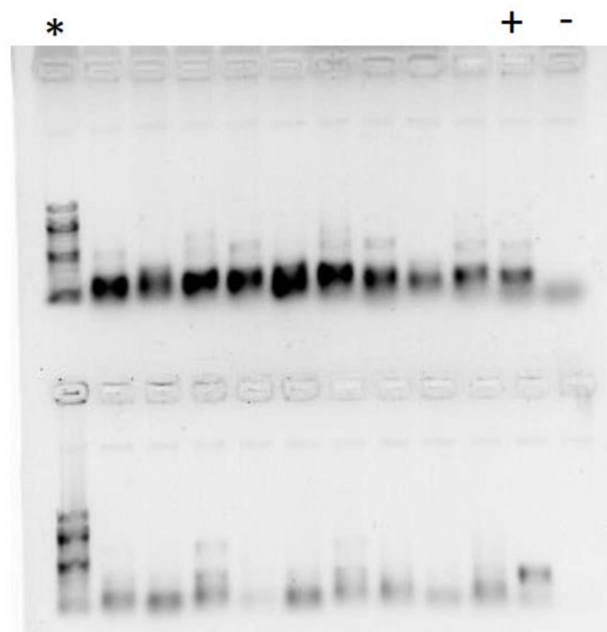


Fig. 2.12 Gradient PCR in dung samples for marker OARCP26 with normal *Taq*. *- lanes with 100 bp DNA ladder, positive (+) and negative (-) control.

With the addition of BSA in PCR reactions performed on dung samples, the contrary to what another study found was observed (Reitmann 2008). BSA unfortunately did not make any difference to these markers' amplification success. The polymerase was also changed (Super-Them *Taq* to multiplex mix), in which case the amplification success of some samples did increase yet, this was only observed for a few markers (Fig. 2.13). This can be observed in Fig. 2.13 which clearly shows that multiplex mix can improve PCR amplification (INRA6) however, it can also have no affect (I206).

As a result of the optimization steps in the present study, it can be suggested that the following steps should be followed in order to obtain consistent results across markers when amplifying DNA from dung samples. First, only freshly extracted dung should be used, this reduces the chances of non-amplification and genotyping errors (Kovach et al. 2003; Piggott 2004; Bourgeois et al. 2019). Secondly, use samples which vary in age (old and fresh) during the gradient PCR to identify the optimal annealing temperatures. As mentioned earlier using tissue samples can only be misleading.

However, tissue samples can be considered as a positive control but should not be used as an indication as to which annealing temperatures are best in dung samples; only dung samples' amplification results should be considered. Once the optimal temperature for each marker has been identified in dung, markers which had the same annealing temperatures, as well as multiplex panels can be amplified together. These markers can subsequently be pooled together with all of the other single amplified markers for fragment analysis.

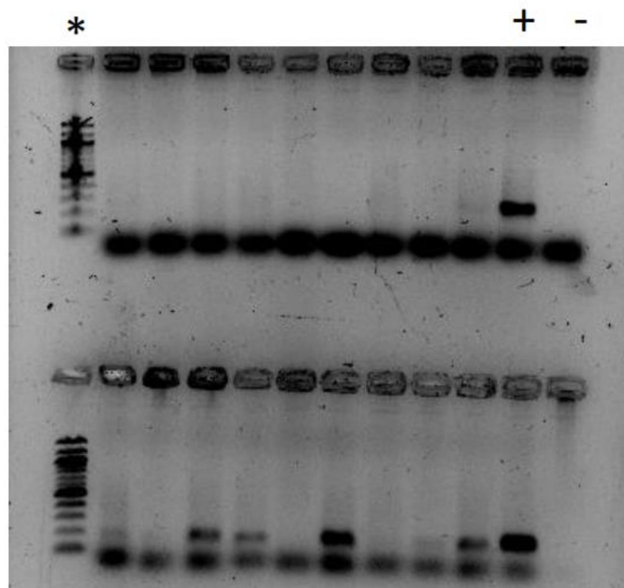


Fig. 2.13 PCR amplification for markers I206 (top) and INRA6 (bottom) with the use of multiplex mix in dung samples. * - 100 bp DNA ladder, positive (+) and negative (-) control.

Conclusion

In this study I have shown and discussed how difficult working with dung samples can be, especially if they are not the freshest. Additionally, this study has also shown the important role sampling can play in downstream steps e.g. PCR, emphasizing the importance of using only the freshest samples. Two extraction kits were compared, and both produced credible results, yet the NucleoSpin DNA Stool kit was considered to be more optimal for suni dung DNA. I was able to optimize amplification of microsatellite markers for dung and tissue samples. These microsatellite markers were subsequently further tested in suni and only six were successfully optimised (discussed in Chapter 3). The observations in this study, specifically for highly degraded dung samples, can be valuable for similar studies in other taxa.

Supplementary material

Table S2.1 Tissue and dung samples collected from different locations used for extraction protocol testing. Unknown = quality of sequence was too low to identify its species

No.	Individual	Origin	Species	Sample type	Sample collection date
1	4	MOZ	Suni	Tissue	2017/11/20
2	6	MOZ	Suni	Tissue	2017/11/20
3	11	MOZ	Suni	Tissue	2017/11/20
4	12	MOZ	Suni	Tissue	2017/11/20
5	13	MOZ	Suni	Tissue	2017/11/20
6	14	MOZ	Suni	Tissue	2017/11/20
7	21	MOZ	Suni	Tissue	2017/11/20
8	22	MOZ	Suni	Tissue	2017/11/20
9	24	MOZ	Suni	Tissue	2017/11/20
10	27	MOZ	Suni	Tissue	2017/11/20
11	32	MOZ	Suni	Tissue	2017/11/20
12	33	MOZ	Suni	Tissue	2017/11/20
13	34	MOZ	Suni	Tissue	2017/11/20
14	35	MOZ	Suni	Tissue	2017/11/20
15	36	MOZ	Suni	Tissue	2017/11/20
16	42	MOZ	Suni	Tissue	2017/11/20
17	43	MOZ	Suni	Tissue	2017/11/20
18	44	MOZ	Suni	Tissue	2017/11/20
19	46	MOZ	Suni	Tissue	2017/11/20
20	101	TSH	Unknown	Dung	2016/11/13
21	103	TSH	Suni	Dung	2016/11/13
22	105	TSH	Unknown	Dung	2016/11/13
23	108	TSH	Red duiker	Dung	2016/11/13
24	109	TSH	Unknown	Dung	2016/11/13
25	111	TSH	Unknown	Dung	2016/11/13
26	112	TSH	Unknown	Dung	2016/11/13
27	115	TSH	Unknown	Dung	2016/11/13
78	116	TSH	Unknown	Dung	2016/11/13
29	118	TSH	Suni	Dung	2016/11/13
30	119	TEM	Suni	Dung	2016/11/14
31	128	TEM	Suni	Dung	2016/11/14
32	129	TEM	Suni	Dung	2016/11/14
33	134	TEM	Suni	Dung	2016/11/14
34	136	TEM	Suni	Dung	2016/11/14

35	137	TEM	Suni	Dung	2016/11/14
36	144	TEM	Suni	Dung	2016/11/14
37	145	TEM	Suni	Dung	2016/11/14
38	147	TEM	Red duiker	Dung	2016/11/14
39	149	TEM	Suni	Dung	2016/11/14
40	153	TSH	Suni	Dung	2016/11/15
41	155	TSH	Suni	Dung	2016/11/15
42	156	TSH	Unknown	Dung	2016/11/15
43	157	TSH	Suni	Dung	2016/11/15
44	158	TEM	Suni	Dung	2016/11/16
45	159	TEM	Suni	Dung	2016/11/16
46	160	TEM	Suni	Dung	2016/11/16
47	162	TEM	Suni	Dung	2016/11/16
48	164	NDU	Suni	Dung	2016/11/17
49	165	NDU	Unknown	Dung	2016/11/17
50	166	TSH	Red duiker	Dung	2016/11/15
51	167	TSH	Unknown	Dung	2016/11/15
52	169	TSH	Unknown	Dung	2016/11/15
53	170	TSH	Suni	Dung	2016/11/15
54	172	TSH	Unknown	Dung	2016/11/15
55	174	TEM	Unknown	Dung	2016/11/16
56	176	TEM	Suni	Dung	2016/11/16
57	177	TEM	Unknown	Dung	2016/11/16
58	178	TEM	Suni	Dung	2016/11/16
59	180	NDU	Suni	Dung	2016/11/17
60	181	NDU	Suni	Dung	2016/11/17
61	183	NDU	Suni	Dung	2016/11/17
62	185	NDU	Suni	Dung	2016/11/17
63	186	NDU	Unknown	Dung	2016/11/17
64	187	NDU	Suni	Dung	2016/11/17
65	188	NDU	Unknown	Dung	2016/11/17
66	191	NDU	Red duiker	Dung	2016/11/17
67	194	NDU	Suni	Dung	2016/11/17
68	196	NDU	Red duiker	Dung	2016/11/17
69	197	NDU	Suni	Dung	2016/11/17
70	199	NDU	Suni	Dung	2016/11/17
71	201	NDU	Suni	Dung	2016/11/17
72	203	NDU	Unknown	Dung	2016/11/17
73	205	NDU	Suni	Dung	2016/11/17

74	207	NDU	Red duiker	Dung	2016/11/17
75	208	NDU	Suni	Dung	2016/11/17
76	209	NDU	Unknown	Dung	2016/11/17
77	210	PHI	Suni	Dung	2016/09/03
78	211M	MNE	Suni	Dung	2016/11/02
79	211	PHI	Unknown	Dung	2016/09/03
80	216	PHI	Suni	Dung	2016/09/03
81	217	PHI	Suni	Dung	2016/09/03
82	218M	MNE	Suni	Dung	2016/11/02
83	218	PHI	Unknown	Dung	2016/09/03
84	219	PHI	Suni	Dung	2016/09/03
85	221	PHI	Suni	Dung	2016/09/03
86	225	PHI	Suni	Dung	2016/09/03
87	337	MNE	Suni	Dung	2016/11/02
88	338	MNE	Suni	Dung	2016/11/02
89	339	MNE	Suni	Dung	2016/11/02
90	340	MNE	Suni	Dung	2016/11/02
91	341	MNE	Suni	Dung	2016/11/02
92	342	MNE	Suni	Dung	2016/11/02
93	343	MNE	Suni	Dung	2016/11/03
94	344	MNE	Suni	Dung	2016/11/03
95	345	MNE	Suni	Dung	2016/11/03
96	346	MNE	Suni	Dung	2016/11/03
97	347	MNE	Suni	Dung	2016/11/04
98	348	MNE	Suni	Dung	2016/11/04
99	611	PHI	Red duiker	Dung	2017/08/22
100	614	PHI	Red duiker	Dung	2017/08/23
101	615	PHI	Suni	Dung	2017/08/24
102	616	PHI	Suni	Dung	2017/08/28
103	617	PHI	Suni	Dung	2017/08/28
104	619	PHI	Unknown	Dung	2017/08/29
105	621	PHI	Suni	Dung	2017/09/04
106	622	PHI	Suni	Dung	2017/09/05
107	633	PHI	Suni	Dung	2017/09/24
108	636	PHI	Suni	Dung	2017/09/25
109	639	PHI	Suni	Dung	2017/09/27
110	640	PHI	Suni	Dung	2017/09/27
111	648	PHI	Suni	Dung	2017/10/02
112	663	PHI	Suni	Dung	2017/10/05

113	671	PHI	Suni	Dung	2017/10/06
114	674	PHI	Suni	Dung	2017/10/06

Table S2.2 Selected primers from VGL with their final primer concentrations

Marker	Final concentration (uM)
BM2113	0.08
BM4028	0.06
BM719	0.06
CSSM19	0.06
CT03	0.05
CT07	0.1
CT13	0.1
DIK20	0.12
D5S2	0.12
ETH10	0.1
I206	0.2
I9	0.1
ILST028	0.05
ILST26	0.08
INRA23	0.08
INRA6	0.08
SPS115	0.1
TGLA126	0.1
TGLA263	0.1

.

Table S2.3 Microsatellite loci tested in suni. Ta = Annealing temperatures for only the final 16 microsatellites and sex marker chosen

Marker	Forward Primer 5'-3'	Reverse Primer 5'-3'	Reference	Ta (°C)
32HDZ688	CCGAGGAGGAGAAAAAGGTC	TGTTGTGTAATCATCAGTCCCC	(Huebinger et al. 2006b)	
32HDZ707	CATTCCCTGGCCACTGTC	CAAGGGGATAGTGATGGAAAG	(Huebinger et al. 2006b)	
33HDZ496	GTTTTTCCAGATGGTATTTTCCTC	GTATTCGGCTGAAGGGACC	(Huebinger et al. 2006a)	
33HDZ749	GTGTGCCATGCTTCTCTATCTG	GTGTGTGAGACGAGGTGCAAG	(Huebinger et al. 2006a)	
AMELOGENIN	CAGCCAAACCTCCCTCTGC	CCCGCTTGGCTTGTCTGTTGC	(Weikard et al. 2006)	60
BM1818	AGCTGGGAATATAACCAAAGG	GTGTCTTGTGTCTTAGTGCTTTCAAGGTCCATGC	(Bishop et al. 1994)	
BM1824	GAGCAAGGTGTTTTTCCAATC	GTGTCTTCATTCTCCAACTGCTTCCTTG	(Bishop et al. 1994)	
BM2113	GCTGCCTTCTACCAAATACCC	CTTCCTGAGAGAAGCAACACC	(Bishop et al. 1994)	60
BM3205	TCTTGCTTCCTTCCAAATCTC	GTGTCTTTGCCCTTATTTTAACAGTCTGC	(Bishop et al. 1994; Van Hooft et al. 1999)	
BM3517	GTGTGTTGGCATCTGGACTG	GTGTCTTTGTCAAATTCTATGCAGGATGG	(Bishop et al. 1994; Van Hooft et al. 1999)	
BM4028	ACGGAAGCAGCATCTCTTAC	GTGTCTTATGGAAACATGGTCTCCTGC	(Bishop et al. 1994; Van Hooft et al. 1999)	
BM415	GCT ACA GCC CTT CTG GTT TG	GAG CTA ATC ACC AAC AGC AAG	(Bishop et al. 1994)	
BMS4008	CGGCCCTAAGTGATATGTTG	GAAGAGTGTGAGGGAAAGACTG	(Kappes et al. 1997)	
BM719	TTCTGCAAATGGGCTAGAGG	GTGTCTTCACACCCTAGTTTGTAAGCAGC	(Bishop et al. 1994; Van Hooft et al. 1999)	
BMB757	TGGAAACAATGTAAACCTGGG	TTGAGCCACCAAGGAACC	(Bishop et al. 1994)	60

BOLA DRBP1	ATGGTGCAGCAGCAAGGTGAGCA	GGGACTCAGTCTCTCTATCTCTTTG	(Creighton et al. 1992)	
BRY	GGATCCGAGACACAGAACAGGCTGC	TTGATCAAGCTAATCCATCCATCCTAT	(Schwerin et al. 1992)	
CSRM60	AAGATGTGATCCAAGAGAGAGGCA	GTGTCTTAGGACCAGATCGTGAAAGGCATAG	(Barendse et al. 1994)	
CSSM19	TTGTCAGCAACTTCTTGTATCTTT	GTGTCTTTGTTTTAAGCCACCCAATTATTTG	(Moore et al. 1994)	
CSSM66	AATTTAATGCACTGAGGAGCTTGG	GTGTCTTACACAAATCCTTTCTGCCAGCTGA	(Barendse et al. 1994)	
CT03	CCATTCTAATGGGACCCTTG	GTGTCTTAGCACCCAACGAAACTAACC	(Røed et al. 2011)	
CT07	CTACCTGGGAAACCCATAT	GTGTCTTTGTTGGTTTCTGCCATACAA	(Røed et al. 2011)	52
CT08	CAGGGTGAAGGAAGTCAGAAA	GTGTCTTGGGTCACCATTACCACCTTTT	(Røed et al. 2011)	
CT10	CCTCTTTGCCCTGTTCTTCA	GTGTCTTGGTATTGGTGACCACCTGCT	(Røed et al. 2011)	
CT12	GATCAATGCTTGCAAATCCA	GTGTCTTGACAGGCAAACAGTGATTGAAA	(Røed et al. 2011)	
CT13	TTTACCGTCTGAGCCACACA	GTGTCTTCCAGAGCAGAATTTTGAGCA	(Røed et al. 2011)	
DIK20	AACCAGTAATCGTGAGAGGA	GTGTCTTAAGAAAGTCCCTACCATGAG	(Inoue et al. 1995)	
D5S2	TACTCGTAGGGCAGGCTGCCTG	GTGTCTTGAGACCTCAGGGTTGGTGATCAG	(Toldo et al. 1993)	
ETH03	GAACCTGCCTCTCCTGCATTGG	GTGTCTTACTCTGCCTGTGGCCAAGTAGG	(Toldo et al. 1993)	
ETH10	GTTCAGGACTGGCCCTGCTAACA	CCTCCAGCCCACTTTCTCTTCTC	(Toldo et al. 1993)	60
ETH225	GATCACCTTGCCACTATTTCCT	ACATGACAGCCAGCTGCTACT	(Steffen et al. 1993)	
F10	TGTCCAGCAGCTTCACCATTACGCC	GCCTTCATGATCTTGTCGATCCACT	(Dietz et al. 1992)	60
I122	CAGCGTACCCTTCAGAAGCA	GTGTCTTGACCCAATGACCCAAAACCAT	(Huebinger et al. 2006b)	
I14	CCTGCAGTGATTTTCTCAAGGC	GTGTCTTAAGGGAAAGGCTGATGCTAC	(Huebinger et al. 2006b)	
I206	ATTAGGAAAAGCAATGTGAATGG	GTGTCTTCACTCCTGTATTCTGCCTGG	(Huebinger et al. 2006b)	55

I264	GAGACCCCAAAGCCCACATT	GTGTCTTGTTGCCATGATGTTTTGCTTTG	(Huebinger et al. 2006b)	
I701	ACTCTTCTGGCTCAAATAATACTGC	GTGTCTTGCGGGCAAGTTCTTTACCAC	(Huebinger et al. 2006b)	
I707	ATTCCCTGGCCACTGTCTTG	GTGTCTTAAGGGGATAGTGATGGAAAGCC	(Huebinger et al. 2006b)	
I9	AAGGTGGGATGTTAGGCAGC	GTGTCTTCTCCTCTCCCTCCCTCTTCC	(Huebinger et al. 2006b)	
ILST028	GCCTAACTAAGTTATTGAGATGACACA	GTGTCTTAAAATTAGTTCATACCCTTGCATGT	(Kemp et al. 1995)	
ILST26	CTGAATTGGCTCCAAAGGCC	GTGTCTTAAACAGAAGTCCAGGGCTGC	(Kemp et al. 1995)	
ILST6	TGTCTGTATTTCTGCTGTGG	ACACGGAAGCGATCTAAACG	(Brezinsky et al. 1993)	
ILST87	AGCAGACATGATGACTCAGC	CTGCCTCTTTTCTTGAGAGC	(Kemp et al. 1995)	57
INRA23	GAGTAGAGCTACAAGATAAACTTC	GTGTCTTTAACTACAGGGTGTTAGATGAACTC	(Vaiman et al. 1994a)	
INRA5	TTCAGGCATACCCTACACCACATG	GTGTCTTAAATATTAGCCAACTGAAAACTGGG	(Vaiman et al. 1992)	
INRA6	AGGAATATCTGTATCAACCTCAGTC	CTGAGCTGGGGTGGGAGCTATAAATA	(Vaiman et al. 1992)	60
INRA63	ATTTGCACAAGCTAAATCTAA CC	AAACCACAGAAATGCTTGGAAG	(Vaiman et al. 1994b)	59
OARCP26	GGCCTAACAGAATTCAGATGATGTTGC	GTCACCATACTGACGGCTGGTTCC	(Ede et al. 1995)	56
RM067	TGAGTAATGCAATAGATACAGTAT	GCTTTGGCCATATGAAGAGCTTT	(Kossarek et al. 1993)	
SPS113	CCTCCACACAGGCTTCTCTGACTT	CCTAACTTGCTTGAGTTATTGCCC	(Røed et al. 2011)	58
SPS115	AAAGTGACACAACAGCTTCTCCAG	GTGTCTTAACGAGTGTCCTAGTTTGGCTGTG	(Mommens et al. 1998)	54
SRCRSP24	AGCAAGAAGTGTCCACTGACAG	TCTAGGTCCATCTGTGTTATTGC	(Yeb et al. 1997)	60
SRCRSP7	TCTCAGCACCTTAATTGCTCT	GGTCAACACTCCAATGGTGAG	(Bhebhe et al. 1994)	
SRCRSP8	TGCGGTCTGGTTCTGATTTCAC	CCTGCATGAGAAAGTCGATGCTTAG	(Bhebhe et al. 1994)	56

SRCRSP9	AGAGGATCTGGAAATGGAATC	GCACTCTTTTCAGCCCTAATG	(Engel et al. 1996)	
TGLA057	GCTTTTTAATCCTCAGCTTGCTG	GTGTCTTGCTTCCAAAACTTTACAATATGTAT	(Barendse et al. 1994; Van Hooft et al. 2000)	
TGLA122	CCCTCCTCCAGGTAAATCAGC	AATCACATGGCAAATAAGTACATAC	(Georges and Massey 1992)	59
TGLA126	CTAATTTAGAATGAGAGAGGCTTCT	TTGGTCTCTATTCTCTGAATATTCC	(Bishop et al. 1994)	
TGLA159	GCATCCAGGGAACAAATTACAAAC	GTGTCTTTTTATTTCGAATCTCTTGAGTACAG	(Barendse et al. 1994; Van Hooft et al. 2000)	
TGLA263	CAAGTGCTGGATACTATCTGAGCA	GTGTCTTTTAAAGCATCCTCACCTATATATGC	(Mommens et al. 1998; Van Hooft et al. 2000)	
TGLA53	GCTTTCAGAAATAGTTTGCATTCA	ATCTTCACATGATATTACAGCAGA	(Barendse et al. 1994)	54
TGLA73	GAGAATCACCTAGAGAGGCA	CTTTCTCTTTAAATTCTATATGGT	(Georges and Massey 1992)	

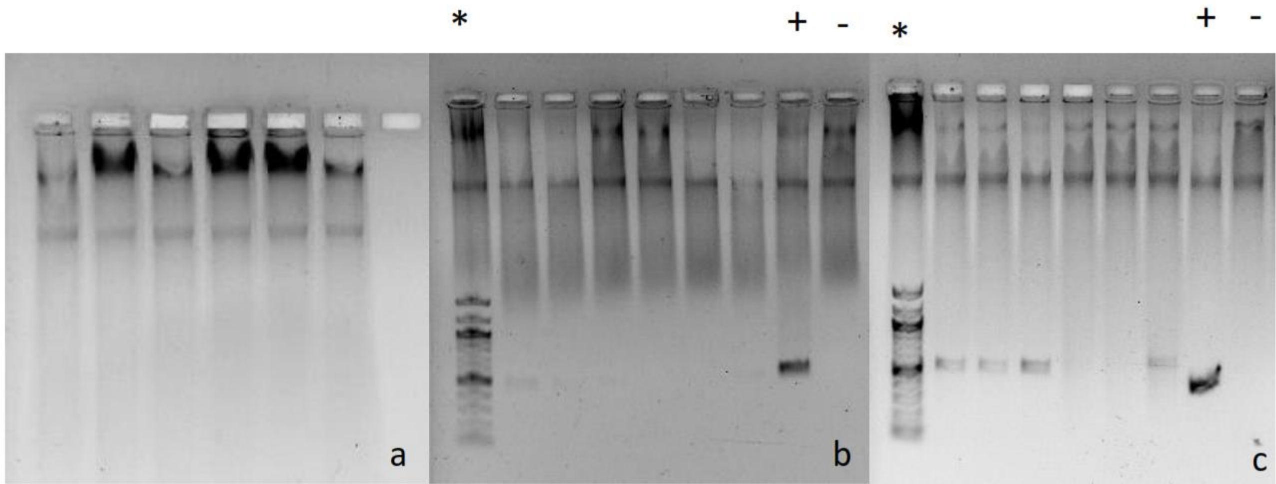


Fig. S2.1 Agarose gels of a) DNA extracts, b) cyt *b* amplifications before and c) cyt *b* amplifications after increasing the DNA concentration with the SpeedVac concentrator. All extractions were performed with the Qiagen extraction kit *- lanes with 100 bp ladder, positive (+) and negative (-) control.

Table S2.4 Nanodrop readings of DNA extracted with the two extraction kits (NucleoSpin DNA Stool kit and the QIAamp DNA mini stool kit)

Sample ID	DNA concentration (ng/µl)	260/280	260/230
118	30.67	2.07	0.31
129	29.88	2	0.29
153	29.13	2.09	0.29
164	30.51	2.05	0.31
173	30.25	1.9	0.32
199	32.86	2.04	0.35

Chapter 3 Conservation genetics of the suni antelope *Nesotragus moschatus* von Dueben, 1846

Abstract

Very little information is currently available for the suni (*Nesotragus moschatus*) especially concerning the valid subspecies taxonomy as well as how suni populations are connected. In this study I made use of genetic non-invasive sampling methods along with conservation genetics approaches to study a few of the suni subspecies. With the use of 67 dung and 19 tissue samples, collected over eastern and southern Africa, I performed the first population genetics study on the suni making use of five microsatellite markers amplified in 64 individuals. In addition, phylogenetic relationships were also assessed with the use of the mitochondrial gene, cytochrome *b*, in 70 individuals (dung and tissue samples), consisting of 63 short and 7 long fragments, along with published sequences. Two Evolutionarily Significant Units were identified in eastern and southern Africa, as well as two possible Management Units in southern Africa. In South Africa alone, limited geographic structuring was identified, indicating that gene flow is likely still occurring in the highly fragmented sand forests of Maputaland. The discoveries made in this study provide valuable insights that can be considered in future conservation management plans designed for the suni as well its habitat.

Introduction

Since the 1970's conservation genetics as a discipline developed along with the growing interest in conservation biology (Frankham 2010a, b). The main objective of this field of study was and remains to comprehend and determine the processes causing species extinction (Hedrick and Hurt 2012). Conservation genetics contributes information on evolutionary and ecological scales and hence, to the development of effective management plans (Hedrick and Hurt 2012). However, it was not until the 1900s that the Convention on Biological Diversity and the IUCN (CBD, 1992) formally recognized the role and importance of biological diversity within conservation. Biodiversity overall is an invaluable source of information and by conserving it, it can provide a society with a wide variety of goods and services (Council 1999). Biodiversity should be considered at three different levels: genes, species and ecosystems (McNeely et al. 1990).

Moreover, conservation genetics focuses on three major concepts: The first is to ensure high levels of genetic diversity along with low levels of inbreeding (Woodruff 2001; Frankham et al. 2002); secondly, to investigate any uncertainties or discrepancies present in the taxonomy and in such a manner identify appropriate conservation management units (Frankham et al. 2002; Frankham 2010a); thirdly, utilising molecular techniques and approaches to understand species biology (Frankham et al. 2002; Frankham 2010a). Thus, current species and population status are determined, which in the end can enable conservation managers to design better management strategies (Frankham 2010a, b). In the event where no research has been performed on population status and genetic variation of a species, limited information about the species' viability and adaptive capabilities would be available, making it difficult to create effective management plans. One such species is the suni *Nesotragus moschatus* von Dueben, 1846.

Nesotragus moschatus forms part of the Bovidae family, which includes the buffalo and antelope species (Fernández and Vrba 2005). The subfamily Antilopinae previously included the tribes Antilopini and Neotragini (Gatesy et al. 1997; Matthee and Robinson 1999), although recent classifications treat the Antilopinae and Neotraginae as separate subfamilies (IUCN 2016). The relationships among members of the Neotragini (small antelopes of the genera *Ourebia*, *Oreotragus*, *Madoqua*, *Raphicerus*, *Neotragus* and *Dorcatragus*) are not very well resolved however, a recent study has shown that *N. batesi* and *N. moschatus* formed a monophyletic group not closely related to *N. pygmaeus*. Thus, the authors recommended resurrecting the genus *Nesotragus*, first described by von Dueben in 1846, for *N. moschatus* and *N. batesi* (Bärmann and Schikora 2014). This recommendation has since been adopted by the IUCN (2016) and is the treatment followed in the present study.

The suni is a very shy and secretive animal and can be found in the eastern parts of Africa, from Kenya south to KwaZulu-Natal in South Africa (Skinner and Chimimba 2005; Frost and Carnaby 2015) (see Fig. 1.7). A population of suni can also be found on Mnemba Island ever since three suni were first introduced in 1991 (Fiske 2011). Since then the suni populations have been growing in such a way that four translocations was deemed necessary to avoid the risk of overgrazing (Fiske 2011). This species is usually associated with enclosed canopy woodlands and dense sand forest patches (Skinner and Chimimba 2005). However suni have been found mostly restricted to the sand forests in South Africa and Mozambique (Prins et al. 2006; Belton et al. 2008; Ramesh et al. 2016). Hence, enabling me to use this relationship between suni and sand forests to estimate the time when suni entered South Africa. To summarise no information is currently available as to when the first suni arrived in South Africa yet, with the use of the sand forests age (Matthews 2005) it can be hypothesized that the maximum time the suni could have inhabited South Africa is between 3 million to 125 000 BP.

Coming back to its habitat, sand forests are highly fragmented and could also cause the suni populations to become more isolated (Jewitt et al. 2015; Hunnicutt et al. 2016). Due to its elusive nature it is very rare to observe this animal in the field and thus the estimated population numbers, stated to be 365 000 individuals, likely represents an overestimation (IUCN 2016). In addition to this, the conservation status for this species on the Global Red List (2016) and the Regional Red List (2016) has been declared as of Least Concern and Endangered, respectively (Hunnicutt et al. 2016). The regional (assessed for mammals of South Africa, Lesotho and Swaziland) status was recently changed from Least Concern to Endangered due to the drastic decline in population numbers during a study performed on dung abundance counts in both Tembe Elephant Park and Mkhuze Game Reserve (Hunnicutt et al. 2016).

Several subspecies have been described (Ansell 1971; Wilson and Reeder 2005; Groves and Grubb 2011), but the distinction between these still remains questionable. Currently only four of the subspecies, *N. m. kirchenpaueri, N. m. moschatus, N. m. livingstonianus and N. m zuluensis* are recognized (Skinner and Chimimba 2005; Groves and Grubb 2011; Frost and Carnaby 2015;

Hunnicutt et al. 2016) (see Fig. 1.7). Furthermore, all of the above mentioned subspecies were classified according to morphological differences such as size and hair patterns, with no genetic background taken into consideration (Groves and Grubb 2011). Nonetheless, one study did utilise genetic information from captive suni individuals. It was discovered that some subspecies differed in diploid chromosome numbers (2n = 52-56) of which 2n = 52 and 2n = 56 belonged to *N. m. akeleyi* and *N. m zuluensis* respectively (Kingswood et al. 1998). The study also concluded that the cytotype variations 2n = 53, 54 and 55 observed had to be due to hybridization between the two subspecies in captivity (Kingswood et al. 1998). These hybridization events were suggested to result in an increased perinatal mortality as well as a decrease in fitness of those individuals (Kingswood et al. 1998). This may reflect the development of reproductive barriers between some of the subspecies and emphasizes the importance of studying the phylogenetic relationships between suni subspecies, as well as their genetic diversity and population structure.

Due to all of the above, this study was conducted on six different populations covering both the eastern and southern parts of Africa (Mnemba Island (MNE), central Mozambique (MOZ), Tembe Elephant Park (TEM), Tshanini Nature Reserve (TSH), Ndumo Game Reserve (NDU) and Phinda Private Game Reserve (PHI)). Samples (dung and tissue) were collected from each site of which the majority consisted of dung samples. As described in Chapter 2 good quality DNA could be extracted from dung samples and these extracts could be used for phylogenetic and population analyses.

The main aims of the present chapter were (1) to determine the phylogenetic relationships between suni subspecies (thought to represent *N. m. moschatus, N. m. livingstonianus* and *N. m. zuluensis*) and (2) to identify the degree of connectivity between the different suni populations, especially across the naturally fragmented range of the species in southern Africa. These analyses enabled me to identify conservation units within suni, as well as to assess the diversity and distribution of one of the suni subspecies.

Materials and methods

Sample collection and selection

For this study non-invasive sampling (of dung) took place at various sites across the species' distribution (Fig. 3.1). Localities in South Africa included the following nature reserves: TEM, TSH, NDU and PHI (Fig. 3.1b). Since suni have such a shy and secretive nature, obtaining fresh dung only occurred on rare occasions. In addition to this dung samples were also collected on MNE, Tanzania. During sampling, different approaches were used at each site, detailed in the methodology section of Chapter 2. Coupled with the dung samples, fresh material (tissue) was also collected from a hunting concession in MOZ (Fig. 2.1 and 3.1a). A total of 67 dung- and 19 tissue samples were used during this study (Table S3.1). Sample coordinates for each reserve was plotted to select only a few samples (approximately 20) per reserve to cover the entire sampling area in each reserve. A minimum distance of 80 m between samples was chosen to minimise resampling of the same

individuals. Information regarding the samples used per locality can be observed in Table 3.1 and Table S3.1.

Fig. 3.1 Google earth images showing the fragmented nature of the forests between the different sampling localities (red pins) in a) southern Africa and b) South Africa (KZN)

Table 3.1 Tissue and dung samples used for sequencing (mtDNA cytochrome *b*) and microsatellite genotyping over six different localities in eastern and southern Africa

Location	# Samples	Sample type
Mozambique	19	Tissue
Tembe	16	Dung
Ndumo	13	Dung
Phinda	19	Dung
Tshanini	7	Dung
Mnemba	12	Dung
Total	**86**	

DNA extractions and species identification

DNA extractions for all dung samples selected were performed by using the Nucleospin® DNA stool kit (Macherey-Nagel) following the manufacturer's instructions with a few alterations (see Chapter 2). Moreover, I made use of the protocol specifically designed for samples that are very hard and dry. For each extraction the pellets' outer scrapings were placed into a NucleoSpin® bead tube and 500 µl of lysis buffer (Buffer ST1) and 500 µl of distilled water (Sabax®, Adcock Ingram) was added, respectively. Tubes were then shaken horizontally for 2-3 seconds to ensure that the sample and lysis buffer was completely mixed. Samples were then placed into an incubator at 70°C for 5 min

followed by a homogenisation step. This was achieved by making use of a vortex adapter (Multi-tube holder of Scientific industries) and agitating the samples for 10 min at room temperature. From this step onwards, the remainder of the manufacturer's instructions were followed with only the DNA elution step changed. In this step each DNA extraction was eluted in 40 µl of elution buffer (buffer SE) by adding it to the column and incubating it at room temperature for 30 min followed by centrifugation at 13000 g for 1 min. DNA extractions on tissue samples (25 mg) were performed using the Animal Tissues Spin-Column Protocol (Qiagen) as recommended by the manufacturer.

For species identification, a part of the mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) gene (467 base pairs (bp)) was amplified with the primers shown in Table 2.2. Amplification and sequencing were performed as described in Chapter 2. Individual's cyt *b* sequences were then aligned in CLC Main Workbench v.8.1 and compared against the data available on NCBI, making use of the BLAST search (Altschul et al. 1990) to confirm the species of origin. Only the samples confirmed to be of suni origin were retained in the study (other samples were identified as red duiker and stored for future analyses).

Phylogenetic relationships and haplotype network

In order to assess the phylogenetic relationships longer fragments of the cyt *b* gene was deemed necessary. Two separate fragments (Fig. 3.2) of the cyt *b* gene were amplified and sequenced using the primers indicated below (Table 3.2). Amplification of one large fragment is very difficult to obtain when working with degraded dung samples, therefore two smaller fragments were amplified. These were performed on seven chosen samples (2 PHI, 2 MOZ, 1 TEM, 1 NDU, 1 MNE) which were previously identified as suni and amplified very well during the species identification step. PCR reactions were set up using the Invitrogen™ Platinum™ II Taq Hot-Start DNA polymerase (Thermo Fisher Scientific) protocol as stated in the manual. For each PCR reaction 2 µl of template DNA were used since DNA quantification was uninformative (see Chapter 2). PCR reactions were carried out in the ABI 2720 Thermo Cycler (Thermo Fisher Scientific) with a denaturing step at 94°C for 5 min. Samples were then subjected to 38 cycles of 94°C for 30 sec, 50°C for 30 sec and 72°C for 30 sec. It was then followed by a final extension step of 72°C for 7 min and a final holding step at 4°C. PCR products were electrophoretically analysed on 1% agarose gels using GelRed (Biotium, Anatech Instuments) staining; a 100 bp DNA ladder (The Scientific Group) was included for confirming the size of the amplicons.

PCR reactions with optimized bands were then precipitated by adding 90 µl 96% EtOH, 10 µl distilled water and 0.05 M NaAc to the PCR product. It was followed by a centrifugation step (20 000 *g* for 20 min) and removal of the supernatant. The remaining pellet was then washed with 90 µl 70% EtOH and centrifuged for 10 min at 20 000 *g*. The supernatant was removed, and the pellet dried at room temperature and eluted in distilled water. Precipitated samples were visualised on a 1% agarose gel. Diluted cycle sequencing reactions (0.5X) were set up with the BigDye™ Terminator v.1.1 and v.3.1 5X Sequencing Buffer (Thermo Fisher Scientific) following the

manufacturer's instructions. Cycle sequencing reactions of 10 µl were set up using 3 µl of precipitated PCR product. Cycle sequencing reactions were carried out on the ABI 2770 Thermo Cycler following conditions stated in the users guidelines. Products were separated on an ABI 3500xl automated sequencer (Applied Biosystems). Both the forward and reverse sequences, of each pair of primers, of individual samples were aligned in CLC Main Workbench v.8.1 (Qiagen).

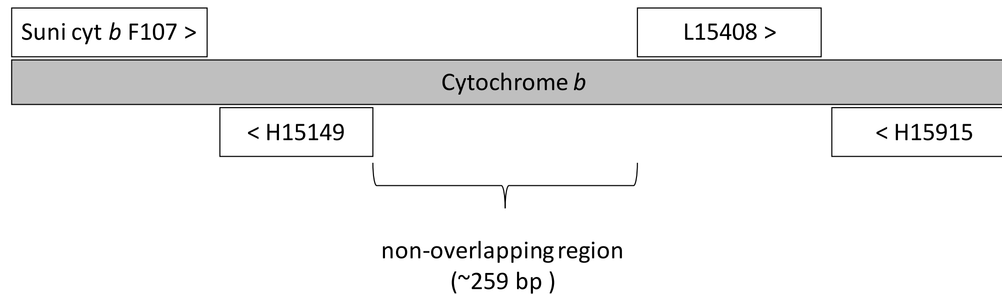


Fig. 3.2 Illustration of the two primer pairs used to amplify two separate cyt *b* gene fragments, separated by a non-overlapping region of approximately 259 bp. Primer positions are labelled according to the human mtDNA sequences (Kocher et al. 1989).

Table 3.2 Two primer pairs used for the amplification of separate cyt *b* fragments in suni

Primer name	Primer pair	Primer sequence (5'-3')
Suni cyt *b* F107	1	TCCTAGGCATCTGCCTAATCC
H15149	1	GCCCCTCAGAATGATATTTGTCCTCA
L15408	2	ATAGACAAA.ATCCCATTCCA
H15915	2	TCTCCATTTCTGGTTTACAAGAC

The seven longer sequences along with NCBI reference sequences (Table S3.2) for *N. moschatus, N. batesi* and *Neotragus pygmaeus* were all aligned in MEGA v.7 (Tamura et al. 2011) using ClustalW. Sequences were then truncated to a fragment of 1034 bp. With the use of MEGA the best model for these sequences was determined; the best fit model based on the AIC and likelihood ratio test was the Hasegawa-Kshino-Yano (HKY) model with Gamma distribution (G). Evolutionary relationships were determined by creating a Maximum Likelihood tree (HKY + G) in MEGA using all sites. In conjunction a pairwise distance matrix using the Maximum Composite Likelihood model with Gamma distribution was also constructed with all 12 sequences to determine the percentage sequence divergence. In addition to this a haplotype network was also drawn with TCS v.1.21 (Clement et al. 2000). This was done with the use of 70 cyt *b* gene sequences (produced during species identification), covering six locations, along with NCBI reference sequences of suni (Table S3.2). These sequences were aligned and truncated (254 bp) in MEGA in order to create a homogeneous dataset which could be used for drawing a haplotype network in TCS.

PCR optimization and amplification of microsatellites and one sex marker

No microsatellite markers specifically designed for suni was available for this study. Instead I made use of microsatellite markers developed for or used in other species such as goat, sheep, gazelle, impala, wildebeest, sable antelope, roan, eland, cattle, gemsbok and buffalo. These markers were all tested and selected in a specific manner, discussed in Chapter 2, so that only the most variable and easily amplifiable markers were chosen. As a result, a total of 16 dinucleotide microsatellites markers were selected for further analyses (Table 2.3) and with the use of Multiplex manager v.1.2 (Holleley and Geerts 2009) two separate multiplexes were designed. One panel contained eight markers and a second panel contained nine (including the sex marker) (Fig. 2.4).

The 16 microsatellite markers along with one sex marker were each tested separately on tissue and dung samples (see Chapter 2). The markers that did not amplify successfully or consistently were removed from further testing. The remaining markers were further tested in multiplex as in Fig. 2.4 but regrettably did not work as expected. Instead markers were amplified in single reactions followed by co-loading of the PCR products during plate preparations for Genescan. In short, PCR reactions were set up using the Invitrogen™ Platinum™ II *Taq* Hot-Start DNA polymerase protocol as stated in the manual. The PCR conditions consisted of 94°C for 2 min, 35 cycles of 94°C for 15 sec, the optimized annealing temperature for 15 sec, followed by 72°C for 30 min and a final holding step at 4 °C. This was performed in the SimpliAmp™ Thermal Cycler (ABI Thermo Fisher Scientific). The annealing temperature differed for each marker and can be observed in Table 3.3. Markers which had the same annealing temperature and occurred together in the original multiplex panels were amplified together. In addition to this it is also important to note that for the majority of samples 2-3 PCR repetitions were performed, however the seven repetitions suggested by (Taberlet et al. 1996) were not reached since consistent genotypes were obtained.

Once PCR amplification was completed, plates for Genescan analysis was set up with LIZ 500 size standard (ABI™/ Life Technologies) and Hi-Di™ Formamide (ABI™/ Life Technologies) according to manufacturer's instructions with the amount of PCR product added being altered. Alternatively, each panels' markers (PCR products) were pooled together as follows: 1 µl of each marker's PCR products were added directly to the Genescan plate. The PCR products were then separated with the ABI 3500xl Genetic Analyser along with the GeneScan™ LIZ 500 size standard. The results generated were analysed using GeneMapper software v.4.1 (Rinehart 2004). Markers that were not polymorphic or did not amplify consistently were eliminated, ending with six microsatellite markers that were successfully optimised for this study (Table 3.3).

Scoring, binning and comparing microsatellites

From the data generated, standard bins and reading rules were developed for each individual marker. In order to ensure accurate scoring, all the data obtained was manually checked after GeneMapper automatically scored each individuals' genotypes. With the use of peak heights (requirement of at least 100 RFUs), stutter peaks and peak shapes, each marker was consistently

Table 3.3 Final set of 16 dinucleotide microsatellite markers and one sexing marker used in this study. Grey = Optimised and remaining markers

Marker	Panel	Forward Primer 5'-3'	Reverse Primer 5'-3'	Reference	Label	T_A
CT07	2	CTACCTGGGAAACCCATAT	GTGTCTTTGTTGGTTTCTGCCATACAA	(Røed et al. 2011)	FAM	56
INRA006	2	AGGAATATCTGTATCAACCTCAGTC	CTGAGCTGGGGTGGGAGCTATAAATA	(Vaiman et al. 1992)	VIC	57
SRCRSP24	1	AGCAAGAAGTGTCCACTGACAG	TCTAGGTCCATCTGTGTTATTGC	(Yeb et al. 1997)	VIC	58
ETH10	2	GTTCAGGACTGGCCCTGCTAACA	CCTCCAGCCCACTTTCTCTTCTC	(Toldo et al. 1993)	PET	53
BM2113	2	GCTGCCTTCTACCAAATACCC	CTTCCTGAGAGAAGCAACACC	(Bishop et al. 1994)	NED	56
TGLA53	2	GCTTTCAGAAATAGTTTGCATTCA	ATCTTCACATGATATTACAGCAGA	(Barendse et al. 1994)	VIC	50
SPS115	2	AAAGTGACACAACAGCTTCTCCAG	GTGTCTTAACGAGTGTCCTAGTTTGGC TGTG	(Mommens et al. 1998)	FAM	52
I206	2	ATTAGGAAAAGCAATGTGAATGG	GTGTCTTCACTCCTGTATTCTGCCTGG	(Huebinger et al. 2006b)	PET	52
SRCRSP8	1	TGCGGTCTGGTTCTGATTTCAC	CCTGCATGAGAAAGTCGATGCTTAG	(Bhebhe et al. 1994)	VIC	50
OARCP26	1	GGCCTAACAGAATTCAGATGATGTTGC	GTCACCATACTGACGGCTGGTTCC	(Ede et al. 1995)	NED	56
ILST87	1	AGCAGACATGATGACTCAGC	CTGCCTCTTTTCTTGAGAGC	(Kemp et al. 1995)	PET	53
SPS113	1	CCTCCACACAGGCTTCTCTGACTT	CCTAACTTGCTTGAGTTATTGCCC	(Roeder et al. 2001)	FAM	58
TGLA122	2	CCCTCCTCCAGGTAAATCAGC	AATCACATGGCAAATAAGTACATAC	(Georges and Massey 1992)	FAM	58
INRA63	1	ATTTGCACAAGCTAAATCTAA CC	AAACCACAGAAATGCTTGGAAG	(Vaiman et al. 1994b)	NED	58
BM757	1	TGGAAACAATGTAAACCTGGG	TTGAGCCACCAAGGAACC	(Bishop et al. 1994)	PET	54
F10	2	TGTCCAGCAGCTTCACCATTACGCC	GCCTTCATGATCTTGTCGATCCACT	(Dietz et al. 1992)	VIC	60
AMELOGENIN	1	CAGCCAAACCTCCCTCTGC	CCCGCTTGGCTTGTCTGTTGC	(Weikard et al. 2006)	FAM	56

scored to ensure the best possible accuracy. In order to reduce genotyping error, each individual was genotyped at least three times for each marker. Once all were scored and binned each locus was evaluated separately according to the number of alleles present over all suni individuals tested and compared to what was recently found in other ungulate species (Miller et al. 2016; Van Wyk et al. 2018). In addition to this, each sample's genotype was compared to the rest in order to identify the different individuals within each nature reserve (i.e. to ensure that, based on the set of markers, there were no identical genotypes), reducing the risk of working on the same sample.

Marker selection and power

Before any further downstream analyses were performed, markers were checked to ensure that they were adequate for this study. This included loci being tested for linkage disequilibrium (LD) with the use of Genepop v.4.7 (Raymond 1995). Thereafter markers were tested for genotyping errors such as null alleles, allelic dropout and false alleles with the use of Micro-checker v.2.2.3 (Van Oosterhout et al. 2004). In the case where null alleles were detected the same data set was assessed by FreeNA (Chapuis and Estoup 2006) to determine whether these markers' null allele frequencies could be corrected, if not the markers were removed.

The final set of markers were then tested for statistical power with the use of Powsim v.4.1 (Ryman and Palm 2006). For this analysis, only the data of individuals with complete genotypes were used. This was performed to assess the power these loci possessed to detect a specific level of genetic differentiation (F_{ST}) between populations. By making use of both the Fisher's exact test and chi-square test (1000 burn-ins, 100 batches and 1000 iterations) a wide variety of F_{ST} values were evaluated. This was done using the same effective population (N_e = 1000) and changing the number of generations of drift (t = 20, 45, 50, 55, 60, 100), as well as the sample sizes to either the same for all (n = 9) or different (n = 11, 19, 12, 9 and 11). The percentage of significant outcomes (α = 0.05, 1000 runs) for the selected range of F_{ST}s' generated by Powsim was used to determine the power of the loci to detect a specific level of genetic differentiation. Lastly the power of the loci to identify individuals, as well as siblings were tested by Cervus v.3.0 (Kalinowski et al. 2007) using only the data for individuals with complete genotypes.

Genetic diversity and statistics

Summary statistics for each locality was determined using Genepop. The diversity indices for each locus were done using Cervus and HP-RARE (Kalinowski 2005) to calculate allelic richness within each locality. To assess the levels of gene flow present, Nm values were deduced from the F_{ST} according to the Wright's island model of population structuring (Wright 1931) using the expression $F_{ST} \approx 1/(1+4Nm)$

Genetic structure and connectivity

In order to assess the genetic structure over all five different localities, the generated dataset over all five selected loci had to be reduced first. This was done by removing individuals who had more

than 60% of their data missing. Overall this dataset finally consisted of 64 individuals (11 MNE, 19 MOZ, 12 TEM, 10 NDU and 12 PHI) for which only 15% of the data was missing. With the use of STRUCTURE v.2.3.4 (Pritchard et al. 2000) all runs performed for the five mentioned localities, were assessed using an admixture model with correlated allele frequencies with parameters set to a burn in period of 100 000 and 500 000 sampled MCMC steps, with *K* set from 1 to 6 possible clusters, at 20 iterations each. In order to determine the most probable number of clusters present, the results obtained in STRUCTURE were further analysed with Structure Harvester (Earl 2012) available online. In conjunction with this, GenALex v.6.5 (Peakall and Smouse 2006; Smouse and Peakall 2012) was used to construct a Principal Coordinates Analysis (PCoA) plot to assess the genetic similarity amongst all of the individuals using the pairwise genetic distance (Codom-genotypic).

In the case of assessing the genetic structure of only the South African (SA) localities (TEM, NDU and PHI) the same parameters were used as above for STRUCTURE with $K = 1$ to $K = 4$ possible clusters. With the use of Structure Harvester, the results obtained in STRUCTURE were used to evaluate the most probable number of clusters present. In addition to that, Geneland v.4.08 (Guillot et al. 2005), a package run in *R*, was also used for a spatially explicit analysis. It was used to assess the genotype data for $K = 1$ to $K = 4$ possible populations. With the use of the correlated allele frequency model, 10 independent runs were performed with a 100 000 MCMC steps, a thinning value of 100 and a burn in period of 200. In addition to this GenALex was used to perform both PCoA and spatial autocorrelation analysis. The spatial autocorrelation analysis was conducted to determine the maximum distance in which genetic similarities between different pairs of suni individuals can be found. The spatial autocorrelation coefficient (*r*) was assessed at each distance class (10, 20, 30, 40, 50, 60, 70, 80, 90, 100 and 110 km) against the null hypothesis (no correlation present between geographic and genetic distance) with the parameters set at 9999 permutations and 9999 bootstraps.

Identifying ESUs and MUs

For this study I decided to combine two previously described definitions (Moritz 1994b; Fraser and Bernatchez 2001) created for conservation and use these to identify potential ESUs and MUs. Moreover I tried to recognise conservation units (ESUs and MUs) by identifying reciprocal monophyletic groups present in the mtDNA of suni (Moritz 1994b). In addition conservation units were also identified by searching for lineages which demonstrated highly restricted gene flow from other lineages using both nuclear and mitochondrial DNA (Fraser and Bernatchez 2001).

Results

DNA extractions and species identification

In the event where red duiker samples were identified, these samples were removed from further analyses. In this study a total of 86 suni DNA samples (19 tissue and 67 dung) were extracted successfully. The results obtained for some of these samples can be observed in Chapter 2. In addition to this all of these samples were successfully amplified and sequenced for the cyt *b* gene

(467 bp). Overall, not all extractions and PCR amplifications worked instantaneously and had to be repeated (see Chapter 2).

Cytochrome *b* analysis

Evolutionary relationships between all selected samples were investigated by generating a phylogenetic tree (Fig. 3.3) including NCBI reference sequences (Table S3.2), as well as a pairwise distance matrix (Table S3.3). From the results obtained a clear separation was observed between the southern (SA and Mozambique) and eastern African (MNE) samples with bootstrap values of 57% and a 5-6% sequence divergence (Table S3.3). These results indicate two ESUs. In addition to this, the MOZ samples (thought to represent *N. m. livingstonianus*) unexpectedly clustered together with the South African samples (*N. m. zuluensis*) (bootstrap value of 70%). Due to the lack of statistical support for reciprocal monophyly, these regions may represent separate MUs but not ESUs. All South African localities clustered together with high bootstrap support (95%).

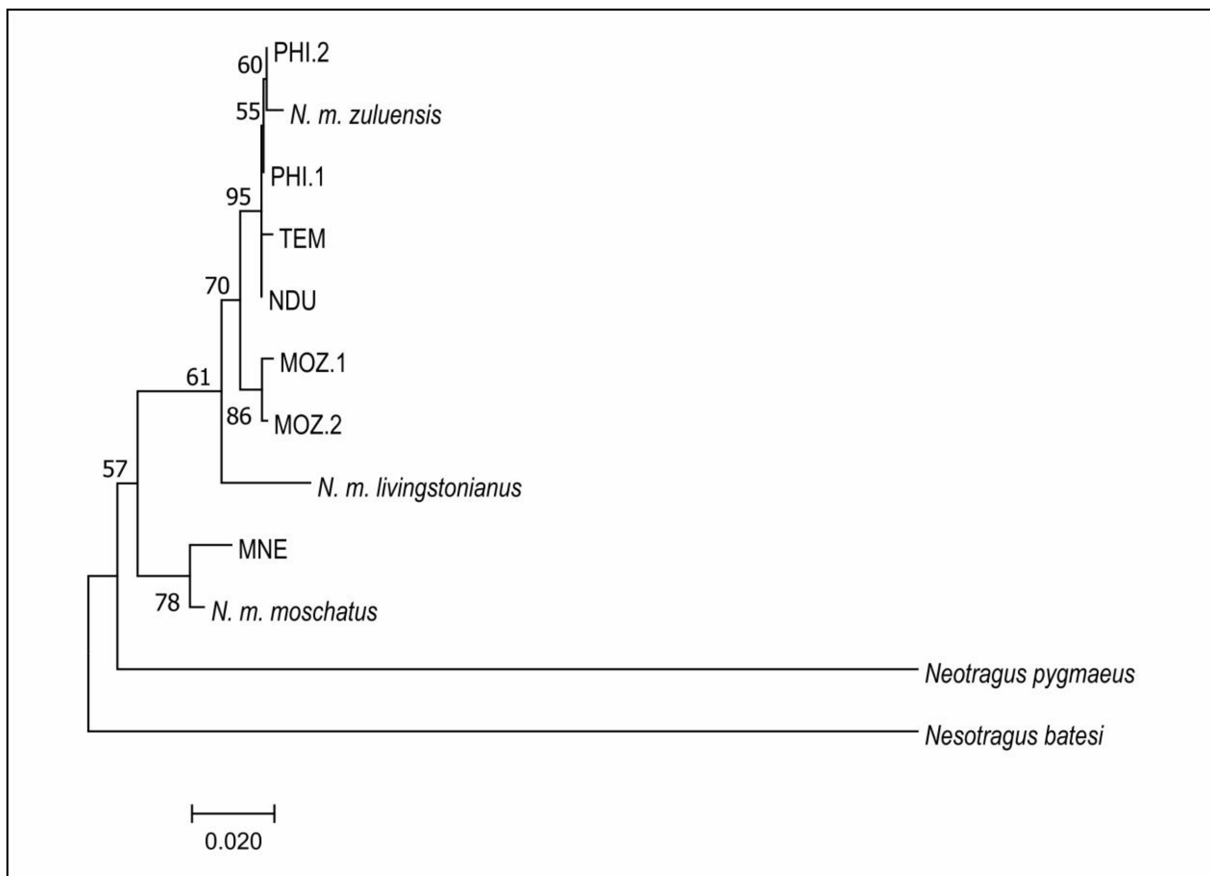


Fig. 3.3 Phylogenetic tree based on part of the mtDNA cyt *b* gene (1034 bp) in suni individuals from five locations (including NCBI sequences).

A haplotype network was drawn so that different haplotypes as well as how each of them were connected could be identified. From the samples analysed, 12 different haplotypes were identified from the short cyt *b* sequences (Table 3.4). Once sequences from NCBI were added to the list the number of haplotypes increased to 16 (Fig. S3.1 and Table S3.4). Moreover, one common ancestral

haplotype was identified for three of the SA nature reserves (TEM, NDU and TSH) (Fig. 3.4). In addition, locality-specific haplotypes, connected to the previously mentioned haplotype, were identified for MOZ and PHI (Fig. 3.4). MNE also had a locality-specific haplotype but could not be connected at the 95% cut-off level (Fig. 3.4).

Table 3.4 Different cyt *b* haplotypes present in six localities in southern and eastern Africa. *n* = Number of individuals with a specific haplotype.

Haplotype	*n*	Location
1	21	NDU, TEM, TSH
2	5	MOZ
3	4	MOZ
4	8	MOZ
5	1	MOZ
6	1	TEM
7	1	TEM
8	1	TSH
9	4	NDU
10	11	PHI
11	1	PHI
12	12	MNE

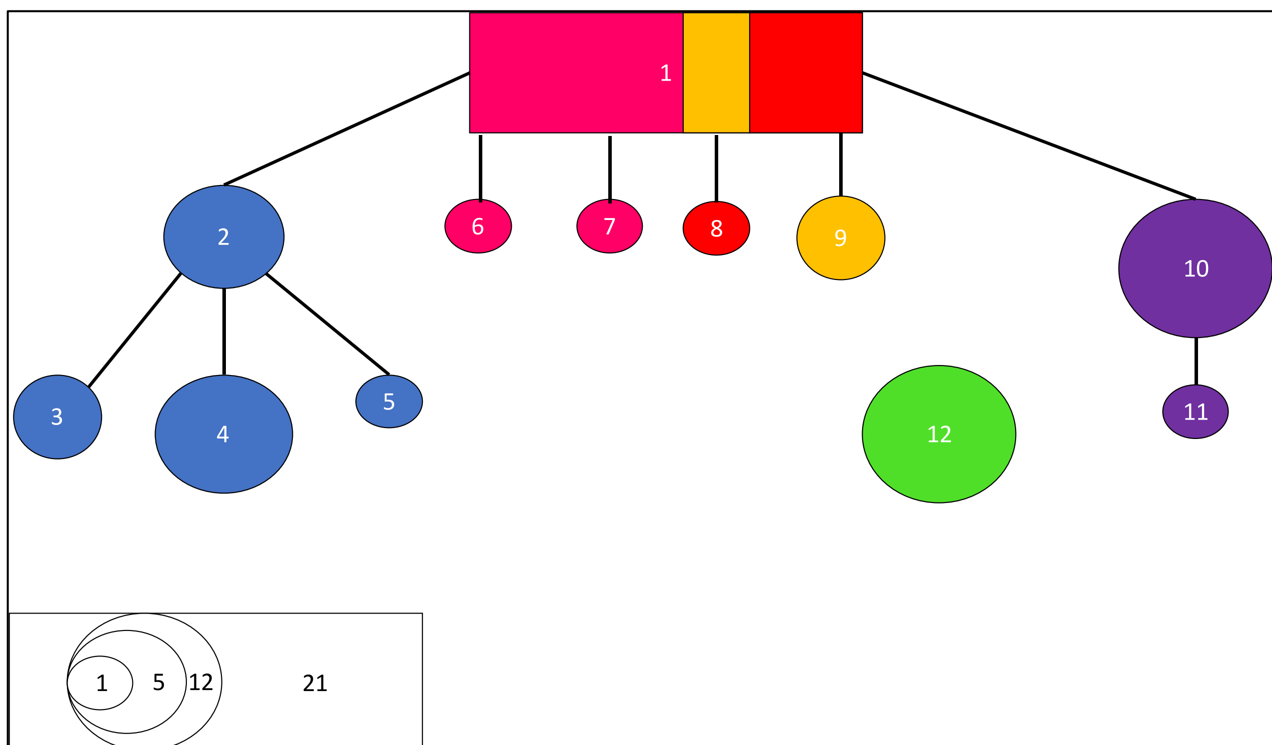


Fig. 3.4 Haplotype network of cyt *b* sequences in six localities in eastern and southern Africa. The white numbers inside each circle/ block corresponds to the haplotype number shown in Table 3.4, with the size of each circle proportional to the number of individuals displaying that specific haplotype. White block/circles = size scale with black numbers indicating number of individuals. Rectangle = haplotype with the highest frequency in southern Africa. Blue = MOZ, Pink = TEM, Red = TSH, Yellow = NDU, Purple = PHI and Green = MNE.

Microsatellite marker optimization and amplification

Microsatellite markers had to be optimised to ensure that only the best performing, and polymorphic markers (more than 2 alleles) were used in the population analyses section of this study. In the end only six of the initial 16 microsatellite markers remained. Markers F10 and ETH10 were removed from the panel due to false alleles that lead to initial classification as polymorphic markers. In the case for markers BM757, INRA63 and TGLA53, these markers were removed due to inconsistent amplification. Markers CT7, INRA6, I206, SRCRSP8 and ILST87 on the other hand either did not amplify in highly degraded samples or amplified non-specifically.

In conjunction with the selected markers, a sex marker (Amelogenin) was also added to this study. By doing this, each samples' sex could be determined, which in turn adds more value to population analyses, as well as contribute to the current knowledge on suni social structure and behaviour. During the amplification process, this marker amplified 95% of the time but, all samples analysed presented two peaks at the expected fragment sizes. This led to the inference that every sample analysed was a male individual, which is highly unlikely. Molecular sexing in suni thus needs further optimization that could not be done within the scope of the present study. For more information regarding selection and optimization of all markers please refer to Chapter 2.

Microsatellite validation, genotyping and comparison

All microsatellite markers used in this study were validated and scored consistently throughout to reduce genotyping errors. Four (SPS113, BM2113, OARCP26 and TGLA122) of the six microsatellite markers used, presented prominent stutter peak patterns making scoring very easy. Although the majority of the markers had these prominent stutter peaks this was not the case for marker SRCRSP24 who had inconsistent allele peak shapes throughout this study (Fig. S3.2). To reduce genotyping errors from occurring for this marker, alleles were scored with reading rules designed here and applied consistently. Marker SPS115 (dinucleotide repeat) on the other hand scored very easily in the beginning due to its shape, but caution had to be taken since some alleles appeared to differ by one base pair instead of the expected two base pair difference for a dinucleotide repeat. For this study a total of 67 samples were successfully genotyped at six microsatellites (Table S3.5). Due to missing data only 64 and 62 of these samples were used respectively in the analyses where five or six loci were considered (Table S3.5).

From the data available only 16 individuals' (complete genotypes) data were used during the assessment of diversity indices in Cervus (Table 3.5). Both the results shown in Table 3.5 as well as in Fig. S3.3 clearly show that all microsatellite markers used in this study were polymorphic. From the results generated by Cervus it can be observed that allele numbers per locus ranged from 3-9 (mean = 5.83) across all six loci. For all of the markers it was observed that heterozygosity was lower than what was expected and could be due to null alleles being present. The null allele results generated by Cervus clearly showed that null alleles should be expected in three (OARCP26, SRCRSP24 and SPS113) of the six markers. Marker SRCRSP24 presented a very high null allele

frequency (Fnull) of 0.7938. In addition to this only two of the six markers were in HWE and one test could not be performed since there were too few individuals present for the program to proceed for that locus.

Table 3.5 Diversity indices generated by Cervus for 16 selected individuals from all different localities

Locus	N_a	*n*	H_O	H_E	HWE	F(Null)
OARCP26	7	16	0.375	0.778	*	0.3443
SRCRSP24	3	16	0.063	0.567	**	0.7938
SPS113	6	16	0.375	0.774	*	0.3451
SPS115	6	16	0.563	0.823	NS	0.1657
BM2113	9	16	0.625	0.865	Nd	0.1385
TGLA122	4	16	0.438	0.712	NS	0.2197

n = Sample size, N_a = Number of alleles, H_O = Observed Heterozygosity, H_E = Expected Heterozygosity, F(null) = Null Allele Frequency and HWE = Hardy-Weinberg Equilibrium (NS = Not Significant, Nd = Not Determined, * P<0.05, ** P<0.01)

Lastly genetic diversity in the suni antelope was also investigated by comparing the number of alleles found here in suni for each locus to those previously found in other ungulates (Miller et al. 2016; Van Wyk et al. 2018) (Table 3.6). From all the loci analysed in suni only two, OARCP26 and BM2113, had more than the average number of alleles for that specific locus found across seven different ungulates (Table 3.6). The other markers generally showed similar results found in one or more of the other ungulates.

Table 3.6 Comparison of number of alleles found in suni (this study) and combined data of other previously studied ungulates (Miller et al. 2016; Van Wyk et al. 2018). * = not available, Imp = Impala, Roa = Roan, Buf = Buffalo, Sab = Sable, Gnu = Wildebeest, Ela = Eland and Gem = Gemsbok.

Marker	**# of alleles present**							**Mean # alleles per locus**	**# of alleles**
	Imp	Roa	Buf	Sab	Gnu	Ela	Gem		Suni
SPS113	6	6	*	12	*	10	8	**6.8**	6
TGLA122	13	*	*	*	7	*	*	**10**	4
OARCP26	*	*	*	*	3	*	*	**3**	7
SRCRSP24	*	6	*	5	11	*	*	**7.3**	3
BM2113	*	8	*	8	6	*	14	**5.5**	9
SPS115	*	*	15	*	*	*	3	**7.5**	6

Marker selection and power analysis

Before any population analyses could be performed, all the markers had to be checked to ensure that the power for downstream inferences was suitable for this study. From Table 3.5 it was noted that null allele frequencies were high for some of the markers and was again confirmed when all data were analysed with Micro-checker (Table 3.7). Markers OARCP26 and SRCRSP24 both identified in Cervus for null alleles, again presented Fnull > 0.3. Interestingly all the samples collected from TEM and NDU also showed high Fnull over five loci. The markers were also tested for linkage

disequilibrium and no significant results were observed. Based on null allele frequency, only five of the six markers were deemed suitable, with marker SRCRSP24 being removed. These five markers were then lastly tested for statistical power, resulting in a 95% probability of these markers being able to detect genetic structure for any $F_{ST} \geq 0.0223$ (Fig. 3.5). Since all pairwise F_{ST} values generated by Genepop (Table 3.8) were above 0.0223 this gave a clear indication that these five loci had the statistical power to detect levels of differentiation within the larger dataset. Furthermore, *Nm* values were calculated and ranged between 0.26 and 1.87 with some location pairs exhibiting *Nm* values of higher than 1 (Table 3.8), suggesting that gene flow are occurring under this model: TEM vs. NDU and PHI and TEM + NDU vs. PHI. In addition to this, Cervus also showed that these five loci had a mean polymorphic information content (PIC) value of 0.7745, a probability to identify individuals at $9{,}2 \times 10^{-7}$, as well as a probability to identify siblings at $6{,}9 \times 10^{-3}$.

Table 3.7 Null allele frequencies generated for all genotyped individuals with the Oosterhout test in Micro-checker. (NS = Not Significant, ND = Not Determined, * P<0.01, ** P<0.001)

Marker	MNE	MOZ	TEM + NDU	PHI
OARCP26	0.3086 ND	0.027 ND	0.2445**	0.092 NS
SRCRSP24	-0.4226 ND	-0.0318 ND	0.3055 ND	0 ND
SPS113	0.0253 ND	0.0213 NS	0.2444**	0.1184 ND
SPS115	0.1441 ND	-0.2302 ND	0.2989**	-0.2929 ND
BM2113	0 ND	-0.0006 NS	0.2353*	0.0527 ND
TGLA122	0.1959 NS	-0.0609 ND	0.0722 NS	-0.142 ND

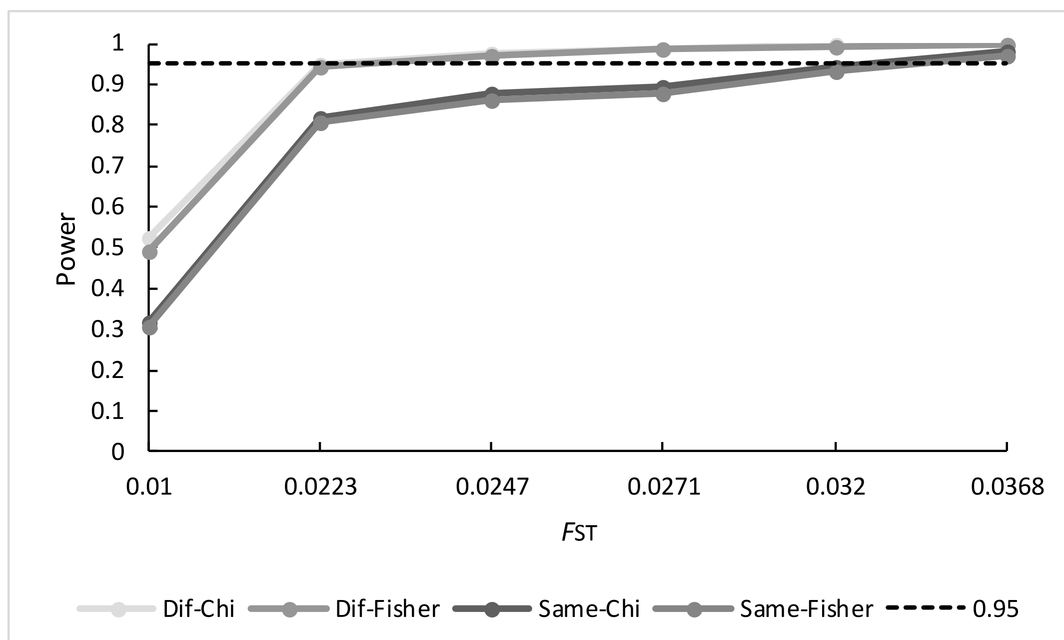


Fig. 3.5 Statistical power comparison of five microsatellite loci to detect a predefined level of genetic differentiation (F_{ST}). Dashed line indicates a 95% probability of detecting genetic structure with approximate F_{ST} values of 0.02254 and 0.0334 respectively when using the same or different sample sizes. Dif = Different sample sizes, Same = Same sample sizes, Chi =Chi-square approach and Fisher = Fisher's exact test

Table 3.8 Pairwise F_{ST} and corresponding *Nm* values between different suni localities based on five microsatellite loci. F_{ST} and *Nm* values are shown above and below the diagonal, respectively. *Nm* values above 1 are shown in bold.

	MNE	MOZ	TEM	NDU	TEM + NDU	PHI
MNE		0.3738	0.3515	0.3876	0.3163	0.4858
MOZ	0.42		0.2509	0.3241	0.2451	0.4017
TEM	0.46	0.75		0.1212		0.1306
NDU	0.39	0.52	**1.81**			0.2104
TEM + NDU	0.54	0.77				0.1178
PHI	0.26	0.37	**1.66**	0.94	**1.87**	

Genetic diversity and structuring of suni samples over all populations

From the data generated in Genepop the observed (H_O) and expected heterozygosity (H_E) over all five localities respectively ranged from 0.2449-0.6506 and 0.4620-0.7533 (Table 3.9). All localities' inbreeding coefficients (F_{IS}), except for MOZ and PHI, were higher than 0.25, with MNE having the highest value of 0.47 (Table 3.9). In addition to this, private alleles were relatively low for all localities with PHI having no private alleles present (Table 3.9). MOZ on the other hand had seven private alleles which was relatively higher than the rest of the localities number of private alleles (Table 3.9). Allelic richness was also very low over all localities ranging from 2.37-3.71 (Table 3.9). These results overall indicate low levels of genetic diversity and possible inbreeding occurring within each location.

Table 3.9 Summary statistics of five localities based on five microsatellite loci generated in Genepop, GenALEx and HP-RARE

Population	n	P_A	A_R	H_O	H_E	F_{IS}
MNE	12	4	2.48	0.2449	0.4620	0.4700
MOZ	19	7	3.24	0.6506	0.6380	-0.0197
TEM	13	4	3.71	0.4107	0.7533	0.4548
NDU	9	4	3.25	0.4000	0.6359	0.3710
PHI	12	0	2.37	0.4375	0.5083	0.1394

n = Sample size, P_A = Private alleles, A_R = Allelic richness, H_O = Observed heterozygosity, H_E = Expected heterozygosity and F_{IS} = Inbreeding coefficient

The PCoA plot generated by GenALEx for 64 suni samples (five different localities) is depicted in Fig. 3.6. Three major clusters can be observed, with MNE forming the first cluster, MOZ the second and TEM, NDU and PHI together forming the third cluster (Fig. 3.6). All three clusters were well defined with the second cluster consisting of all MOZ samples, as well as one TEM sample (Fig. 3.6). The two axes shown in the PCoA plot cumulatively explain 31.95% of the overall variance found in the genotypic data. In addition to this all genotypic data (64 individuals) were analysed in STRUCTURE to see if the same results as in Fig. 3.6 was obtained. The bar plots of each individual's assignment probability (q) for $K = 2$, $K = 3$ and $K = 4$ can be observed in Fig. 3.7. From these results the most probable number of clusters present is $K = 3$ (Fig. 3.7), with the first cluster consisting of

all MNE samples, the second cluster formed by all the MOZ samples and the third cluster made up of all South African samples (TEM, PHI AND NDU), although some samples from TEM and NDU assigned to more than one cluster ($q < 0.8$). The Delta K (Evanno et al. 2005) showed the largest value at $K = 2$ whereas the log probability of the data indicated $K = 4$ as the most likely number of clusters (Fig. S3.4a, b). Once MNE was removed delta K showed the largest value at $K = 2$ whereas the log probability of the data indicated $K = 3$ as the most likely number of clusters (Fig. S3.4c, d).

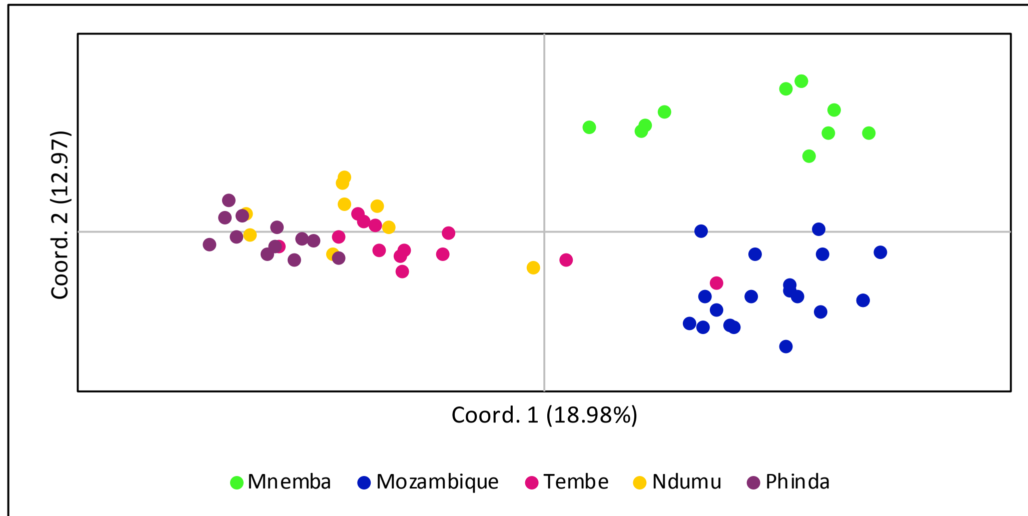


Fig. 3.6 Principal Coordinates Analysis (PCoA) plot clustering 64 suni individuals from five different localities based on genetic distances between genotypes at five microsatellite loci

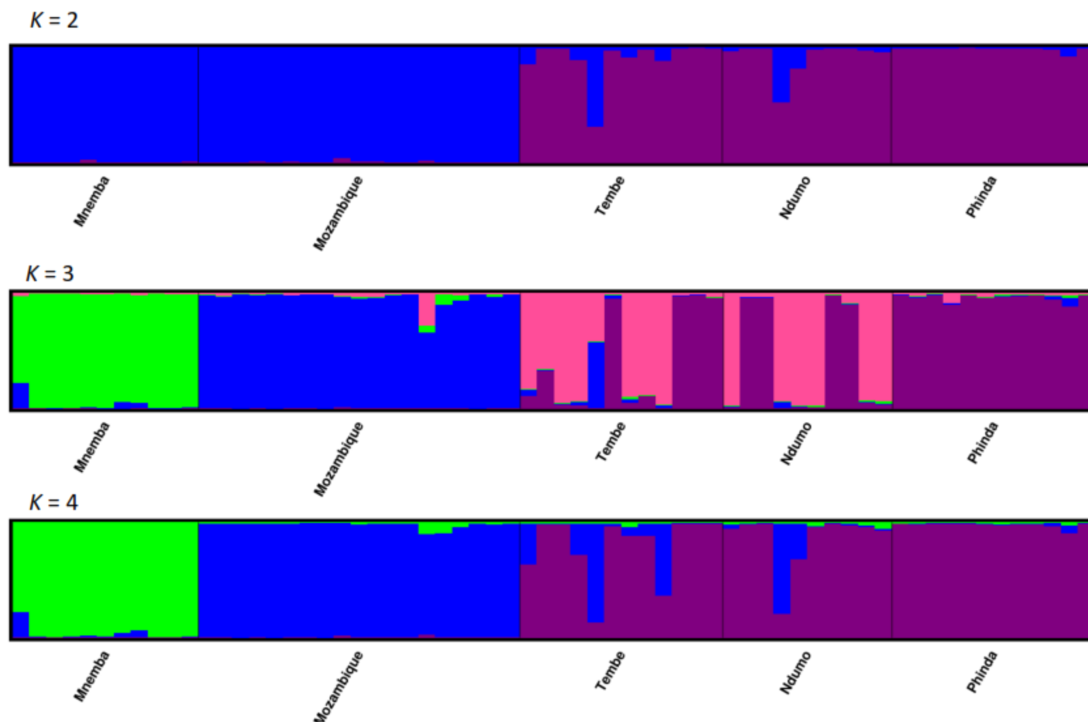


Fig. 3.7 STRUCTURE results depicting individual-based clustering of 64 suni individuals from five localities based on genotypes at five microsatellite loci for $K = 2$, $K = 3$ and $K = 4$ clusters. Each vertical bar along the X-axis represents an individual and the Y-axis is the probability of assignment (q) to K clusters. Location was not specified as prior in the analysis. Localities are arranged from north to south.

Spatial analysis and genetic structure of the South African localities

In order to determine the degree of connectivity between SA localities as well as the structure present, a few different analyses had to be performed. The first of which was a PCoA plot generated in GenALex. The graph obtained does not show clear separation of individual genotypes from the three localities (Fig. 3.8). The axes from the PCoA plot explain approximately 34.3% of the variance observed in the genotypic data. In the interest of the previous results obtained the same data were evaluated in STRUCTURE, which identified $K = 2$ as the most possible number of clusters present (Fig. 3.9). The first cluster consisted of only PHI samples and the second consisted of both TEM and NDU samples, of which about a third of TEM and NDU samples assigned to the first cluster (PHI) or indicated assignment to both clusters. Coupled with this, Geneland analysis was also performed, indicating $K = 3$ as the most probable number of populations across all SA samples with a 56% confidence over all ten runs (Fig. S3.5). In addition to this the posterior probabilities graphs in Fig. 3.10 supports this by identifying three clusters. The first cluster consisted of TEM and PHI, the second cluster of the remaining TEM samples and the third cluster consisting of NDU. Lastly, a spatial autocorrelation analysis was performed to determine the maximum distance in which related individuals could be found. From Fig. 3.11 it could be seen that a positive correlation was identified between genetic and geographic distances for less than 18 km.

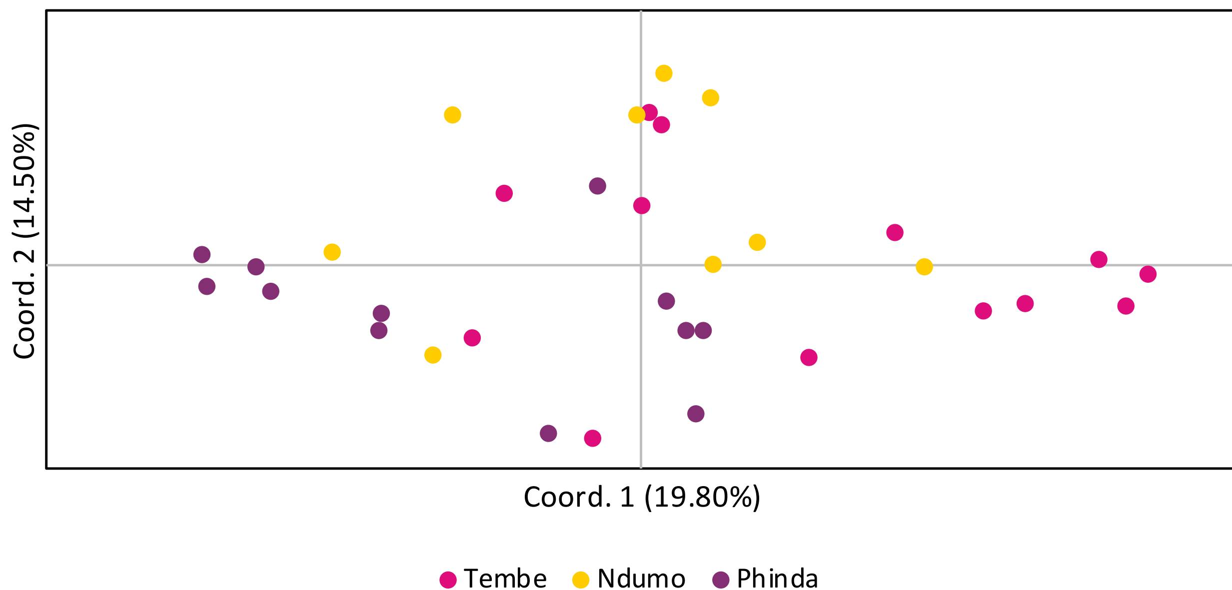


Fig. 3.8 Principal Coordinates Analysis (PCoA) plot clustering 34 suni individuals from three different localities in South Africa based on genetic distances between genotypes at five microsatellite loci.

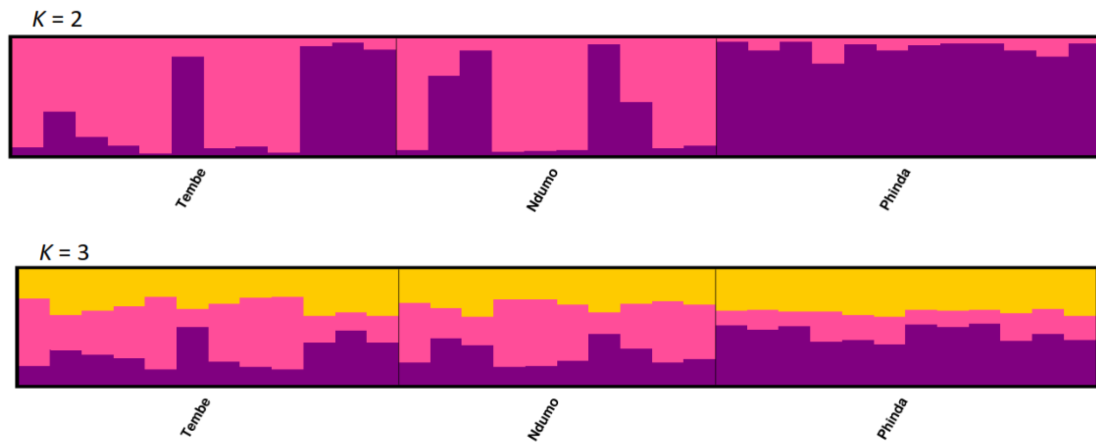


Fig. 3.9 STRUCTURE results depicting individual-based clustering of 34 suni individuals from three localities in South Africa based on genotypes at five microsatellite loci for $K = 2$ and $K = 3$ clusters. Each vertical bar along the X-axis represents an individual and the Y-axis is the probability of assignment (q) to K clusters. Location was not specified as prior in the analysis. Localities are arranged from north to south

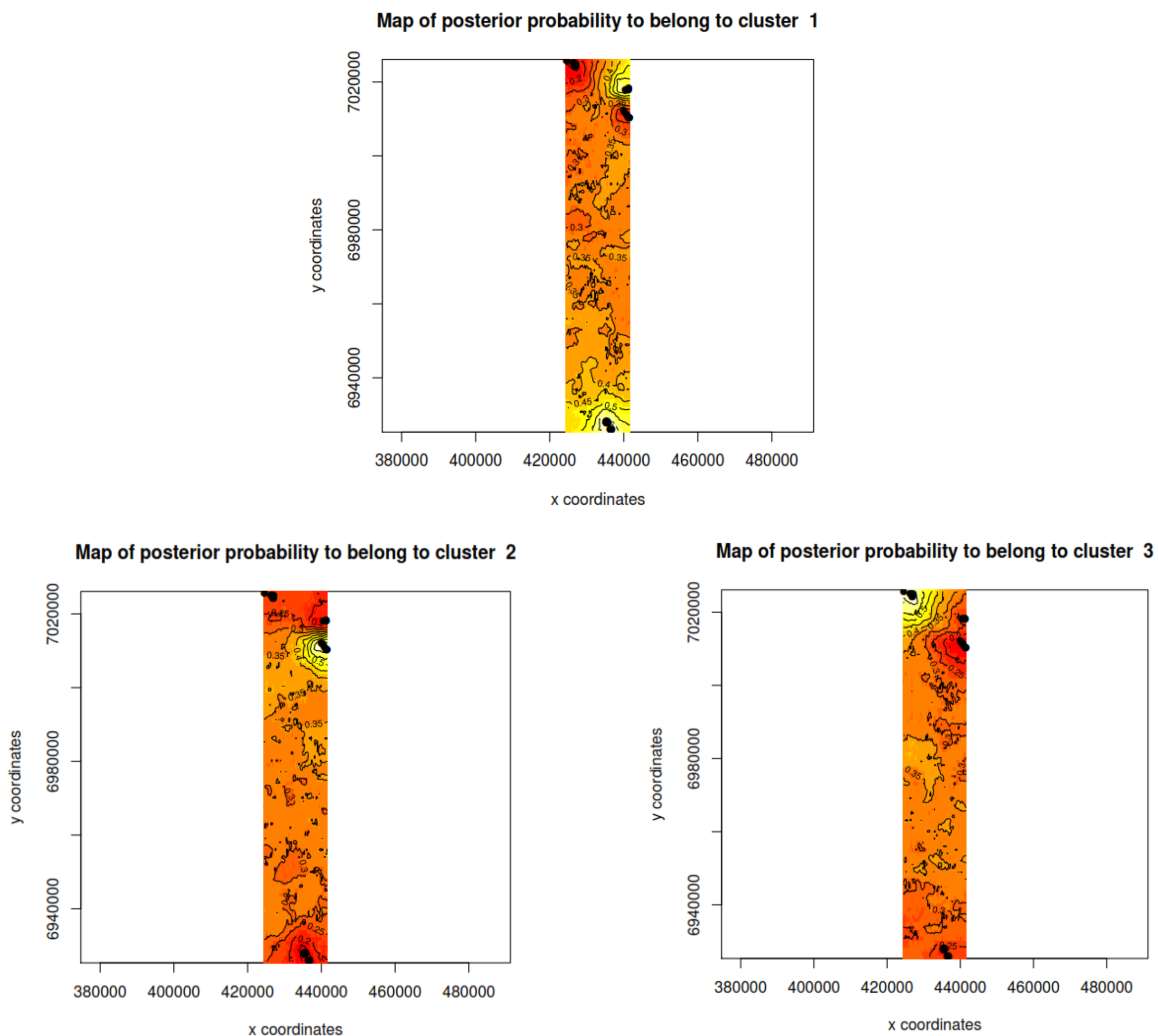


Fig. 3.10 Maps of posterior probabilities of population membership at $K = 3$, based on 34 sunis' microsatellite genotypes at five loci from three South African localities. The black dots represent the different sampling localities while the white-yellow colour indicates a high probability of population membership and the orange-red colour indicates a low probability of population membership. Contour lines indicate posterior probability.

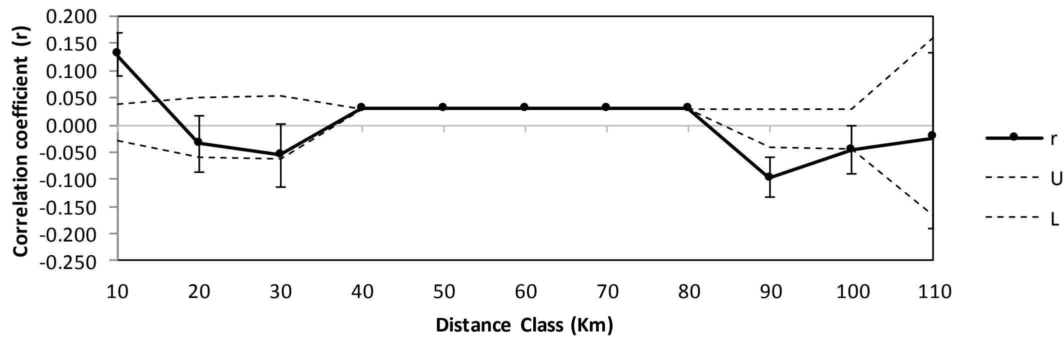


Fig. 3.11 Spatial autocorrelation analysis in suni populations from South Africa. The correlogram indicates the correlation coefficient (*r*) at various distance classes (km). Dashed lines U and L indicate the upper and lower confidence limits.

Discussion

In this study the main aims were to assess the phylogenetic relationships between three suni subspecies and, to investigate the connectivity between separated forest patches in northern KwaZulu-Natal as well as Mozambique. With the use of 64 (dung and tissue) samples collected from MNE (Tanzania), central MOZ and northern KwaZulu-Natal I was able to identify two Evolutionarily Significant Units (ESU), as well as two possible Management Units (MU) in the suni. In addition to this I also studied the genetic diversity within the five localities and investigated the spatial structuring between the southern Africa locations. The results obtained in this study have been informative and will add value to the conservation management of the species. Moreover, no other genetic study has been performed on the suni focussing on the subspecies using wild populations, or to perform population analyses.

Implications of using dung DNA analysis

Since there is so little known about the suni, its subspecies and population connectivity (broad and fine scale), the species was an important focal species for conservation genetics analyses. One of the reasons why there is so little known about this animal is because of its very shy and secretive nature, thus making it very difficult to find and study (Frost and Carnaby 2015). Nevertheless, recent studies have shown that the usual material (blood and tissue) used for genetic studies does not have to be a limitation for wildlife studies anymore (Chiou and Bergey 2017; Chiou and Bergey 2018). Animals can now be studied in the field without observing or hurting them by just collecting their dung samples (Bourgeois et al. 2019). Although this might sound very promising and straightforward, most studies that utilised this sampling method were on large animals (Ferreira et al. 2018; Bourgeois et al. 2019).

While it might seem strange for this method being utilised mostly in big animals this could be due to the ease experienced in finding large animals and following them to collect fresh dung samples. Whereas with small mammals, this is not the case. In the present study I experienced exactly this because as mentioned earlier the species is very shy and if samples were only collected when the animals were observed I would have very limited data. Thus, I had to search all possible patches

intensively to find some samples, and consequently the material was not always the freshest. Various environmental factors such as UV rays and rain could also have affected the quality of these samples (Kovach et al. 2003; Piggott 2004; Wultsch et al. 2015). With this in mind, as well as the fact that the TEM and NDU samples were not always the freshest samples collected, the high null allele frequencies observed (for TEM and NDU) could potentially be explained by allelic drop-out since the environmental factors affected DNA quality and downstream analyses (Kovach et al. 2003; Jeffery et al. 2007; Wultsch et al. 2015). Yet, other factors such a sample size (Pruett and Winker 2008) could also explain the high null allele frequencies observed. However, this will be discussed later.

Microsatellite and sexing markers

The markers used in this study were all meticulously selected and tested in Chapter two. Within this chapter, 16 microsatellite markers were tested further and additional marker filtering was necessary to prevent any biases in the analyses (such as large amounts of missing data). Due to this decision only six microsatellite markers remained along with a sexing marker. In the results it was observed that even though the sex marker did amplify at the expected sized alleles, all individuals evaluated were inferred to be males. This seems unlikely and hence, further optimization (using known sex individuals from captive populations or other observational data) is needed. In addition, the locus should be cloned and sequenced in some of the known-sex individuals, to design a more optimal primer set for application to the wild suni population.

The six microsatellite markers on the other hand could be reliably amplified and genotyped but created some concern once their null allele frequencies (Fnull) were estimated. Three of the six markers presented with relatively high Fnull (>0.34) and was therefore further investigated. The marker SRCRSP24 which had the highest value in Cervus (0.7938) and second highest in Micro-checker (0.3055), provided justification to remove this marker from further population analyses. Redesign of the primers may improve resolution at this locus. Marker OARCP26 also showed some concern in the MNE population (Fnull = 0.3086) but, since there were no missing data and because this population is on a separate island, it would be expected that the genetic diversity to be lower. It was decided to keep OARCP26 in the analyses since such high values were not observed in other localities. Values of $0.2 < \text{Fnull} < 0.3$ were observed in the TEM and NDU populations for five of the six markers. Although these results could be due to lower DNA quality, as seen in Chapter 2, each sample was genotyped at least three times to reduce allelic drop out and false alleles from occurring.

With the above considered only five of the six microsatellite markers were retained for further analyses. These markers were then tested for their statistical power and were considered appropriate for broad and finer scale genetic analyses. Many studies performed in wildlife forensics make use of six to ten microsatellite markers (Kolodziej et al. 2012) and the aim of my study was to optimize at least 10 loci. Nevertheless a study done on dung samples in boars have shown that a minimum of four microsatellites can be used to perform individual identification, especially in closely

related individuals (Kolodziej et al. 2012). Additionally it has also been shown that potential genotyping error can increase with an increase of microsatellite markers (Taberlet and Luikart 1999) hence, minimizing microsatellite markers can be one way of reducing genotyping errors (Kolodziej et al. 2012). Going forward it is recommended that the number of variable microsatellite markers are increased instead of increasing the number of samples (Landguth et al. 2012), especially when landscape genetic inferences are being made.

Genetic diversity of the suni

With the results obtained for standard summary statistics (including H_O, H_E and F_{IS}), I observed that the genetic diversity present in suni is very low. Since the scoring was done consistently the results obtained is reliable hence, sample type and environmental factors can be excluded as possible reasons for the low levels of diversity. Moreover these low levels of genetic diversity found in MNE, TEM, NDU and PHI, located at the edges of this species' distribution, does not come as a surprise especially since such populations do show lower levels of genetic diversity compared to more centrally located (within species distribution) populations (Hardie and Hutchings 2010). A second explanation could be the sample sizes having an effect (Pruett and Winker 2008). MOZ had 6-10 more samples than the other locations allowing it to have a possible effect on genetic diversity levels. In addition to this MNE is an isolated population on Mnemba Island and SA at the edge of the suni distribution hence, the low variability observed could be due to small Ne and the founder effect (Sonsthagen et al. 2017). However, the chances for this being true for SA might be slim since it is hypothesized that suni only moved into SA 3 million to 125 000 BP. The last possible reason could be due to cross-species amplification which could have played a significant role. This can be said since this study made use of markers not designed for suni, increasing the chances of non-amplification therefore resulting in null alleles and low genetic diversity observed. The suni is not phylogenetically closely related to any of the species whose genomes have been mined for microsatellite loci (Matthee and Robinson 1999; Matthee and Davis 2001). Future research should thus aim to mine the suni genome for polymorphic markers.

ESUs and MUs in the suni antelope

To date, very few studies have been performed on the suni, except for phenotypes, habitat, distributions, food resources and behaviour (Ansell 1971). Suni subspecies descriptions have been based on phenotype and distribution only (Ansell 1971). In previous studies six different suni subspecies have been described, of which only four are currently recognised (Ansell 1971; Wilson and Reeder 2005; Groves and Grubb 2011). One genetic study performed on captive individuals, identified a variety of cytotypes and came to the conclusion that mating between different suni subspecies could be the reason for the increased perinatal mortality, as well as increased infertility they observed (Kingswood et al. 1998). This raises questions about the distinctiveness of the subspecies.

Since subspecies delineation is challenging, units such as Evolutionarily Significant Units (ESUs) and Management Units (MUs) were created to provide an objective method which can help prioritise conservation below the taxonomic level of species (Ryder 1986). Results based on mtDNA sequences and microsatellite markers unfortunately weren't consistent. According to the mtDNA sequences a genetic distinction between the East African and southern African suni were observed, however this split only had a bootstrap value of 57%. Whereas with microsatellite data (STRUCTUTRE) $K = 2$ clustered MOZ and MNE together and SA on its own. Nevertheless, when $K = 3$ was performed MOZ, SA and MNE clustered on their own. Hence due to the latter, the reciprocal monophyly in mtDNA, haplotype network and the significant differences found in nuclear allele frequencies it could be said that Mnemba Island should be considered a separate ESU from the southern African lineage.

With regards to MUs I reanalysed the results within each of the ESUs identified, which in this case was only for the ESU identified in southern Africa, where multiple populations had been sampled. The results obtained in the PCoA plot along with STRUCTURE results for the five populations indicated a clear separation between MOZ and the SA populations. This split was also supported by the phylogenetic tree which had a bootstrap value of 70% for MOZ and South Africa as well as a divergence of 0.8-1.2%. Hence, indicating that MOZ could be a distinct MU.

Some interesting observations were also made once the STRUCTURE, PCoA plot, Geneland and phylogenetic analyses results were looked at for the SA localities only. Overall it can be deduced that PHI, TEM and NDU should be seen as one MU. Although some analyses clustered them all together (PCoA and STRUCTURE), other analyses (STRUCTURE and Geneland) distinguished them as separate clusters. However, the ways in which TEM, NDU and PHI were separated did not show consistency between the different analyses. I therefore suggest that these three groups (PHI, NDU and TEM) should currently be treated as one MU, although PHI is geographically separated from the other two reserves and with an increase in markers and samples, may represent a unique MU.

Even though identifying ESUs and MUs was an important part of this study, any information gathered regarding the subspecies could also add value. According to the results obtained from the haplotype network and phylogenetic tree the overall comparison between the two were very similar except for one difference, the position of the NCBI sequence for *N. m. livingstonianus* was not consistent overall. This sequence could be found in the haplotype network grouped together with the MOZ samples whereas in the phylogenetic tree this sequence could be found separate from MOZ. This could be explained by the fragment lengths which were used to do these analyses (254 bp for the haplotype network and 1034 bp for the phylogenetic tree) and from this as well as the bootstrap value (61%) supporting this separation, it could be said that the results of the phylogenetic tree for now should be considered as final in this study.

Due to this decision the following observation should also be taken into account. According to the described distributions of the different suni subspecies (Ansell 1971; Skinner and Chimimba

2005; Groves and Grubb 2011; Frost and Carnaby 2015) it would be expected that samples taken immediately south of the Zambezi river would represent *N. m. livingstonianus*. However, the MOZ samples which were taken south of the Zambezi River did not cluster along with the NCBI sequence of *N. m. livingstonianus*, instead it clustered along with *N. m. zuluensis*. This may indicate that the available information on the distribution of the subspecies is incorrect or that they represent synonyms. Unfortunately, no information about the NCBI sequence's (*N. m. livingstonianus*) sampling locality is available. Additional sampling from the type localities of the subspecies and especially sampling from parts of northern Mozambique, as well as increasing fragment lengths would resolve the current uncertainties discussed above.

Degree of connectivity across South African suni localities

Studying both genetic connectivity as well as structure can provide information about the processes occurring in different populations. Moreover, reliable information regarding the distribution and connectivity of populations have become valuable since it could affect species persistence, especially for those found in fragmented landscapes (Cushman et al. 2013; Wegmann et al. 2014). Due to the above-mentioned facts it was important to study the genetic connectivity as well as structure present in the suni populations since their habitat in South Africa consists mostly of sand forest (Matthews et al. 2001; Gaugris et al. 2004; Skinner and Chimimba 2005; Frost and Carnaby 2015), which is highly fragmented (Eeley et al. 1999; Jewitt et al. 2015; Hunnicutt et al. 2016), placing the emphasis on maximised connectivity and maintenance between these fragmented landscapes. In addition to this, the status of the suni has also been declared regionally as endangered according to the Red List of Mammals of South Africa, Lesotho and Swaziland (Hunnicutt et al. 2016), emphasising the importance of studying the SA populations.

For all the SA locations it could be said there are moderate levels of connectivity present since there are very few genotypes that overlap in the PCoA plot generated for the SA locations only. Low numbers of private alleles (especially in PHI) and the low confidence level in identifying three clusters in Geneland, appear to be indicative of ongoing gene flow between the different reserves. In addition, the results obtained in STRUCTURE (SA only), as well as the haplotype network indicated some separation. The results of both analyses grouped NDU and TEM together with PHI clustering on its own. The grouping of NDU and TEM can be expected since the minimum distance needed for correlation between genetic and geographic distance is approximately 18 km and the maximum distance between the furthest samples in TEM and NDU was 22 km apart. Hence, supporting the grouping observed. In addition to this, the F_{ST} values were lower compared to the broader scale comparisons, ranging between 0.1212-0.2104, indicating gene flow between all populations; the highest levels of gene flow were inferred between TEM vs. NDU and NDU + TEM vs. PHI.

Some degree of structure can thus be observed in the SA populations but that they should rather be taken as one MU present based on the current analyses. It should be advised that generating more complete genotypes, using more loci and decreasing the Fnull would better inform future

conservation decisions for the SA reserves. Even so, these results improved the understanding of the structure present in the suni populations of South Africa and what factors need to be taken into consideration in future decision making.

Conclusion

In conclusion, this was the first population genetics study performed on suni, in which I was able to identify two ESUs present in southern and eastern Africa, as well as two possible MUs. In addition, the study revealed some structuring between localities in SA, with finer scale resolution needed by including larger samples sizes and additional markers. The discoveries made in this study are important and can be considered in conservation management plans designed in future. Future studies should also increase the number of localities, not studied in this chapter, and include investigation of the karyotypes present in wild individuals for each subspecies previously described. By doing this, potential reproductive barriers could be identified and the conservation efforts for the species improved.

Supplementary data

Table S3.1 Suni samples used for sequencing (cytochrome *b*) and genotyping over six different localities in eastern and southern Africa

No.	Individual	Origin	Species	Sample type	Genotyped	Collection date
1	4	MOZ	Suni	Tissue	Yes	2017/11/20
2	6	MOZ	Suni	Tissue	Yes	2017/11/20
3	11	MOZ	Suni	Tissue	Yes	2017/11/20
4	12	MOZ	Suni	Tissue	Yes	2017/11/20
5	13	MOZ	Suni	Tissue	Yes	2017/11/20
6	14	MOZ	Suni	Tissue	Yes	2017/11/20
7	21	MOZ	Suni	Tissue	Yes	2017/11/20
8	22	MOZ	Suni	Tissue	Yes	2017/11/20
9	24	MOZ	Suni	Tissue	Yes	2017/11/20
10	27	MOZ	Suni	Tissue	Yes	2017/11/20
11	32	MOZ	Suni	Tissue	Yes	2017/11/20
12	33	MOZ	Suni	Tissue	Yes	2017/11/20
13	34	MOZ	Suni	Tissue	Yes	2017/11/20
14	35	MOZ	Suni	Tissue	Yes	2017/11/20
15	36	MOZ	Suni	Tissue	Yes	2017/11/20
16	42	MOZ	Suni	Tissue	Yes	2017/11/20
17	43	MOZ	Suni	Tissue	Yes	2017/11/20
18	44	MOZ	Suni	Tissue	Yes	2017/11/20
19	46	MOZ	Suni	Tissue	Yes	2017/11/20
20	103	TSH	Suni	Dung	No	2016/11/13
21	118	TSH	Suni	Dung	No	2016/11/13
22	119	TEM	Suni	Dung	No	2016/11/14
23	128	TEM	Suni	Dung	Yes	2016/11/14
24	129	TEM	Suni	Dung	Yes	2016/11/14
25	134	TEM	Suni	Dung	Yes	2016/11/14
26	136	TEM	Suni	Dung	Yes	2016/11/14
27	137	TEM	Suni	Dung	Yes	2016/11/14
28	138	TEM	Suni	Dung	No	2016/11/14
29	144	TEM	Suni	Dung	Yes	2016/11/14
30	145	TEM	Suni	Dung	Yes	2016/11/14
31	149	TEM	Suni	Dung	Yes	2016/11/14
32	153	TSH	Suni	Dung	No	2016/11/15
33	155	TSH	Suni	Dung	No	2016/11/15
34	156	TSH	Suni	Dung	No	2016/11/15
35	157	TSH	Suni	Dung	No	2016/11/15
36	158	TEM	Suni	Dung	Yes	2016/11/16

37	159	TEM	Suni	Dung	Yes	2016/11/16
38	160	TEM	Suni	Dung	Yes	2016/11/16
39	162	TEM	Suni	Dung	Yes	2016/11/16
40	164	NDU	Suni	Dung	Yes	2016/11/17
41	165	NDU	Suni	Dung	No	2016/11/17
42	170	TSH	Suni	Dung	No	2016/11/15
43	176	TEM	Suni	Dung	Yes	2016/11/16
44	178	TEM	Suni	Dung	Yes	2016/11/16
45	180	NDU	Suni	Dung	Yes	2016/11/17
46	181	NDU	Suni	Dung	Yes	2016/11/17
47	183	NDU	Suni	Dung	Yes	2016/11/17
48	185	NDU	Suni	Dung	Yes	2016/11/17
49	187	NDU	Suni	Dung	Yes	2016/11/17
50	194	NDU	Suni	Dung	Yes	2016/11/17
51	197	NDU	Suni	Dung	Yes	2016/11/17
52	199	NDU	Suni	Dung	No	2016/11/17
53	203	NDU	Suni	Dung	No	2016/11/17
54	205	NDU	Suni	Dung	Yes	2016/11/17
55	208	NDU	Suni	Dung	Yes	2016/11/17
56	210	PHI	Suni	Dung	No	2016/09/03
57	216	PHI	Suni	Dung	No	2016/09/03
58	217	PHI	Suni	Dung	No	2016/09/03
59	219	PHI	Suni	Dung	No	2016/09/03
60	221	PHI	Suni	Dung	No	2016/09/03
61	225	PHI	Suni	Dung	No	2016/09/03
62	337	MNE	Suni	Dung	Yes	2016/11/02
63	338	MNE	Suni	Dung	Yes	2016/11/02
64	339	MNE	Suni	Dung	Yes	2016/11/02
65	340	MNE	Suni	Dung	Yes	2016/11/02
66	341	MNE	Suni	Dung	Yes	2016/11/02
67	342	MNE	Suni	Dung	Yes	2016/11/02
68	343	MNE	Suni	Dung	Yes	2016/11/03
69	344	MNE	Suni	Dung	Yes	2016/11/03
70	345	MNE	Suni	Dung	Yes	2016/11/03
71	346	MNE	Suni	Dung	Yes	2016/11/03
72	347	MNE	Suni	Dung	Yes	2016/11/04
73	348	MNE	Suni	Dung	Yes	2016/11/04
74	615	PHI	Suni	Dung	Yes	2017/08/24
75	616	PHI	Suni	Dung	Yes	2017/08/28
76	617	PHI	Suni	Dung	Yes	2017/08/28
77	621	PHI	Suni	Dung	No	2016/09/04

78	622	PHI	Suni	Dung	Yes	2017/09/05
79	633	PHI	Suni	Dung	Yes	2017/09/24
80	636	PHI	Suni	Dung	Yes	2017/09/25
81	639	PHI	Suni	Dung	Yes	2017/09/27
82	640	PHI	Suni	Dung	Yes	2017/09/27
83	648	PHI	Suni	Dung	Yes	2017/10/02
84	663	PHI	Suni	Dung	Yes	2017/10/05
85	671	PHI	Suni	Dung	Yes	2017/10/06
86	674	PHI	Suni	Dung	Yes	2017/10/06

Table S3.2 NCBI sequences and the codes used in the phylogenetic tree in Fig. 3.3

NCBI sequence (cytochrome *b* mitochondrial gene)	**NCBI accession number**	**Code used in the phylogenetic tree**
Neotragus moschatus livingstonianus	AF022069.1	*N. m. livingstonianus*
Neotragus moschatus from Tanzania	FJ959387.1	*N. m. moschatus*
Neotragus moschatus zuluensis	AF022051.1	*N. m. zuluensis*
Neotragus batesi	KJ193408.1	*N. batesi*
Neotragus pygmaeus	JF728777.1	*Neotragus pygmaeus*

Table S3.3 Pairwise distance matrix of cyt *b* gene sequences (1034 bp) in selected individuals representing each location along with NCBI reference sequences. *N. m. z* = *N. m. zuluensis, N. m. l* = *N. m. livingstonianus, N. m. m* = *N. m. moschatus* and *N. b* = *N. batesi*

	PHI.2	PHI.1	MOZ.1	MOZ.2	TEM	NDU	MNE	*N. m. z*	*N. m. l*	*N. m. m*	*N. b*
PHI.2											
PHI.1	0.000										
MOZ.1	0.012	0.010									
MOZ.2	0.011	0.008	0.004								
TEM	0.004	0.003	0.012	0.011							
NDU	0.001	0.000	0.010	0.008	0.003						
MNE	0.054	0.053	0.054	0.051	0.056	0.053					
N. m. z	0.004	0.004	0.018	0.017	0.008	0.005	0.059				
N. m. l	0.029	0.028	0.034	0.032	0.033	0.031	0.060	0.033			
N. m. m	0.048	0.046	0.048	0.046	0.050	0.047	0.013	0.049	0.054		
N. b	0.155	0.157	0.153	0.158	0.150	0.158	0.151	0.158	0.181	0.148	
Neotragus pygmaeus	0.117	0.122	0.132	0.127	0.122	0.123	0.128	0.138	0.149	0.148	0.222

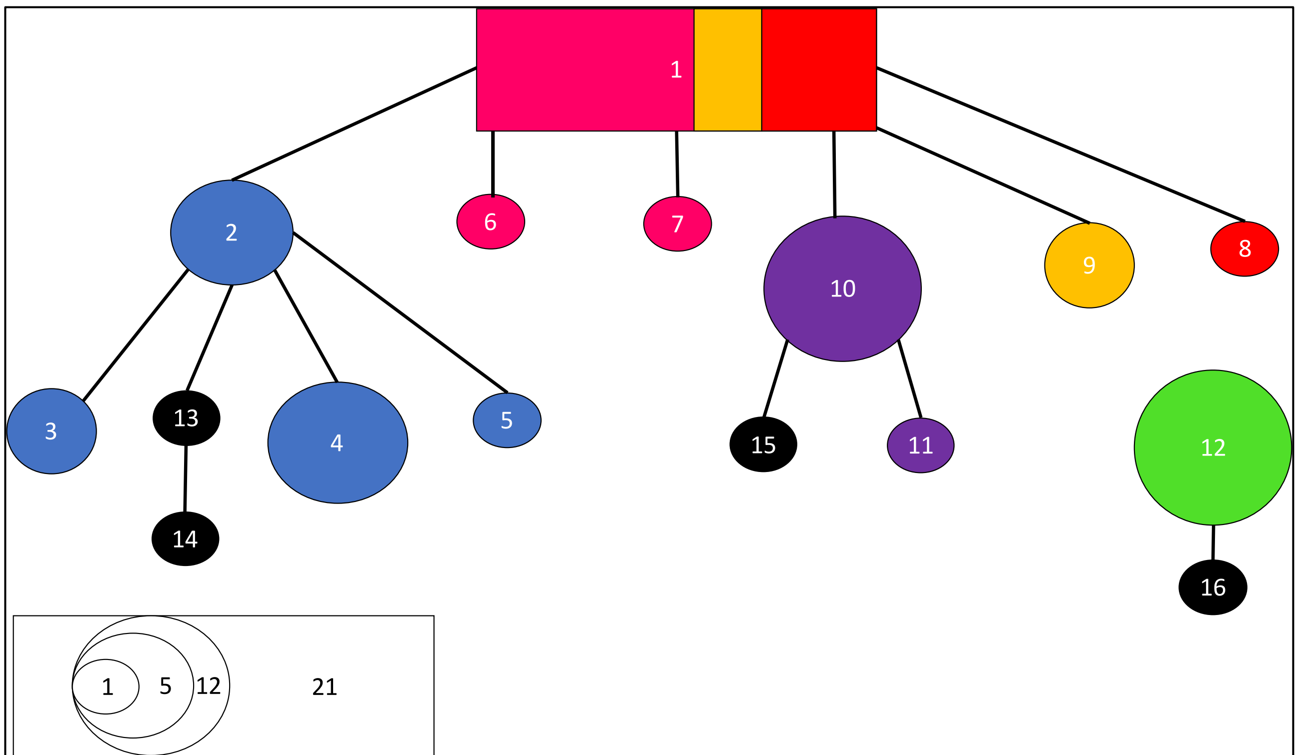


Fig. S3.1 Haplotype network of 74 suni individuals over six locations based on the cyt *b* gene. The white numbers inside each circle/ block corresponds to the haplotype number shown in Table S3.4, with the size of each circle proportional to the number of individuals displaying that specific haplotype. White block/circles = size scale with black numbers indicating number of individuals. Rectangle = haplotype with the highest frequency. Blue = MOZ, Pink = TEM, Red = TSH, Yellow = NDU, Purple = PHI, Green = MNE and Black = NCBI reference sequences

Table S3.4 Different haplotypes present over various locations, including NCBI sequences based on cyt *b* sequences of the suni samples

Haplotype	N	Location/ NCBI accession number
1	21	NDU, TEM, TSH
2	5	MOZ
3	4	MOZ
4	8	MOZ
5	1	MOZ
6	1	TEM
7	1	TEM
8	1	TSH
9	4	NDU
10	11	PHI
11	1	PHI
12	12	MNE
13	1	FJ959386.1 (*Neotragus moschatus* from Mozambique)
14	1	AF022069.1 (*Neotragus moschatus livingstonianus*)
15	1	AF022051.1 (*Neotragus moschatus zuluensis*)
16	1	FJ959387.1 (*Neotragus moschatus* from Tanzania)

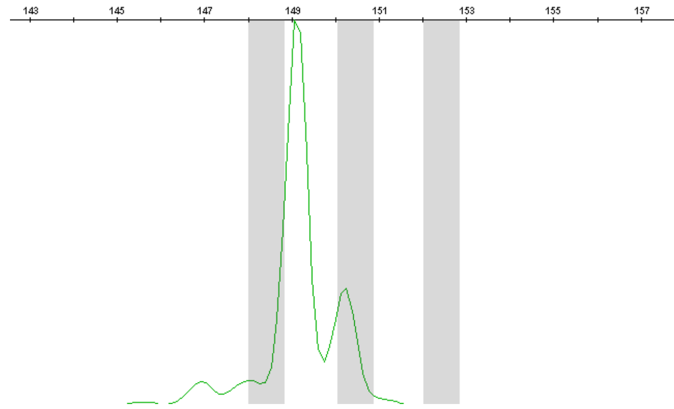


Fig. S3.2 Genotyping results presenting the unusual peak shape for the dinucleotide microsatellite marker SRCRSP24

Table S3.5 Genotypic data over six loci in all individuals genotyped. a = 5 loci analyses (all except SRCRSP24), b = 6 loci analyses (all markers)

Pop	Sample	Analyses used in	OARCP26		SRCRSP24		SPS113		SPS115		BM2113		TGLA122	
MOZ	4	a and b	0	0	150	150	147	147	253	266	131	149	155	161
MOZ	6	a and b	0	0	150	150	147	151	265	266	137	145	155	157
MOZ	11	a and b	133	133	150	150	147	151	253	266	145	153	155	155
MOZ	12	a and b	133	133	150	150	147	147	266	268	145	153	155	161
MOZ	13	a and b	133	133	150	150	147	151	253	266	145	153	155	155
MOZ	14	a and b	133	133	0	0	139	147	266	268	131	155	0	0
MOZ	21	a and b	133	139	150	150	147	147	253	266	131	149	155	161
MOZ	22	a and b	133	139	150	150	147	147	265	266	137	145	155	155
MOZ	24	a and b	129	129	150	150	151	151	253	266	153	153	155	155
MOZ	27	a and b	133	133	150	152	147	149	253	266	135	153	147	147
MOZ	32	a and b	0	0	150	150	143	147	266	267	131	145	155	161
MOZ	33	a and b	0	0	150	150	139	151	266	266	145	145	0	0
MOZ	34	a and b	0	0	150	150	143	151	266	266	145	145	155	155
MOZ	35	a and b	133	143	150	150	145	145	253	268	145	153	155	161
MOZ	36	a and b	133	141	0	0	149	151	266	266	0	0	155	155
MOZ	42	a and b	0	0	150	150	147	151	265	266	137	149	149	155
MOZ	43	a and b	133	133	150	150	147	147	266	266	131	155	0	0
MOZ	44	a and b	0	0	0	0	143	147	265	266	145	145	155	163
MOZ	46	a and b	133	139	150	150	147	151	253	266	137	145	0	0
MNE	337	a and b	133	133	0	0	143	151	260	260	141	141	0	0
MNE	338	a and b	133	133	0	0	145	149	0	0	141	141	149	149
MNE	339	a and b	135	135	148	150	145	145	260	260	141	141	0	0
MNE	340	a and b	133	133	148	150	145	149	259	260	141	141	0	0
MNE	341	a and b	133	133	150	150	145	149	0	0	141	141	149	159

MNE	342	a and b	133	133	148	150	145	145	0	0	141	141	149	149
MNE	343	a and b	133	141	150	150	145	149	260	260	141	141	155	155
MNE	344	a and b	133	141	0	0	147	149	260	260	141	141	0	0
MNE	345	a and b	135	135	0	0	145	145	259	260	141	141	0	0
MNE	346	a and b	135	135	148	150	145	145	260	260	141	141	0	0
MNE	347	a and b	141	141	0	0	145	145	259	259	141	141	159	163
MNE	348	none	129	129	0	0	0	0	0	0	141	141	159	159
TEM	128	a and b	127	145	152	152	139	147	0	0	135	147	159	159
TEM	129	a and b	127	127	0	0	0	0	268	270	151	151	157	161
TEM	134	a and b	131	145	0	0	137	147	0	0	147	151	159	161
TEM	136	a and b	127	127	150	152	0	0	254	254	151	151	159	161
TEM	137	a	0	0	0	0	147	147	0	0	151	151	159	161
TEM	144	a and b	131	133	0	0	143	143	254	266	151	153	155	155
TEM	145	none	143	145	0	0	0	0	0	0	0	0	0	0
TEM	149	a and b	129	145	0	0	149	149	262	262	133	153	159	159
TEM	158	a and b	127	133	0	0	143	143	0	0	151	151	159	159
TEM	159	a and b	131	131	0	0	135	145	254	254	133	133	161	161
TEM	160	a and b	131	131	0	0	0	0	254	254	135	145	157	161
TEM	162	a and b	127	145	0	0	137	149	268	268	151	151	161	161
TEM	176	a and b	127	127	0	0	137	137	262	262	133	133	159	161
TEM	178	a and b	145	145	0	0	0	0	270	270	151	151	161	161
NDU	164	none	129	129	148	152	0	0	0	0	0	0	159	161
NDU	180	a and b	129	129	150	150	143	145	0	0	149	151	159	161
NDU	181	a and b	129	129	0	0	149	149	262	262	143	151	159	161
NDU	183	a and b	129	129	148	148	137	137	268	269	151	151	161	161
NDU	185	a and b	129	143	152	152	147	147	269	270	137	137	159	161
NDU	187	a and b	129	129	0	0	145	147	0	0	137	149	159	159

NDU	194	a and b	129	129	0	0	145	145	0	0	151	151	157	159
NDU	197	a and b	129	129	152	152	137	137	268	268	133	133	159	159
NDU	205	a and b	129	129	152	152	137	141	0	0	133	133	153	159
NDU	208	a and b	129	147	0	0	145	145	0	0	151	151	159	161
Phi	615	a and b	145	145	152	152	137	149	0	0	133	133	159	161
PHI	616	a and b	145	147	152	152	137	149	0	0	133	133	159	161
PHI	617	a and b	145	145	152	152	137	149	262	270	133	133	159	161
PHI	622	a and b	127	127	152	152	137	143	262	270	133	151	159	159
PHI	633	a and b	129	129	152	152	137	137	270	270	133	151	159	161
PHI	636	a and b	0	0	152	152	137	137	0	0	151	151	159	161
PHI	639	a and b	0	0	152	152	149	149	0	0	133	133	159	159
PHI	640	a	0	0	0	0	137	137	0	0	133	133	159	159
PHI	648	a and b	129	145	152	152	149	149	0	0	133	133	159	159
PHI	663	a and b	127	145	152	152	137	137	270	270	133	145	159	161
PHI	671	a and b	0	0	152	152	137	137	0	0	133	133	155	159
PHI	674	a and b	127	145	152	152	137	137	0	0	133	151	161	161

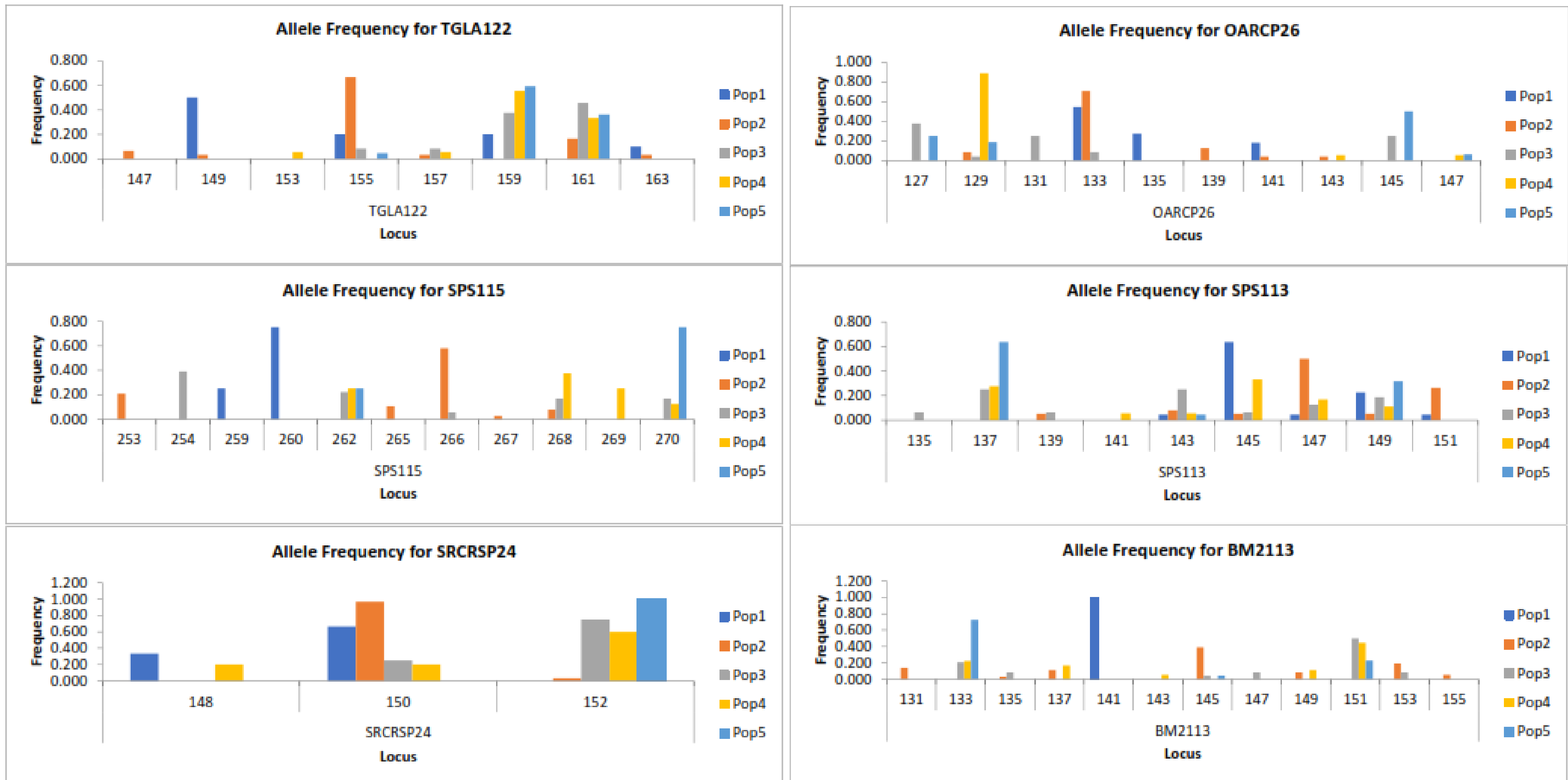


Fig. S3.3 Frequency and allele sizes of alleles present at each marker for all different localities. Pop1 = MNE, Pop2 = MOZ, Pop3 = TEM, Pop4 = NDU and Pop5 = PHI

Fig. S3.4 Structure harvester results. a) Represents the log probabilities of the possible number of clusters present in eastern and southern Africa, b) Delta *K* graph indicating two possible clusters in eastern and southern Africa, c) Represents the log probabilities of the possible number of clusters present in only southern Africa and d) Delta *K* graph indicating two possible clusters in only southern Africa.

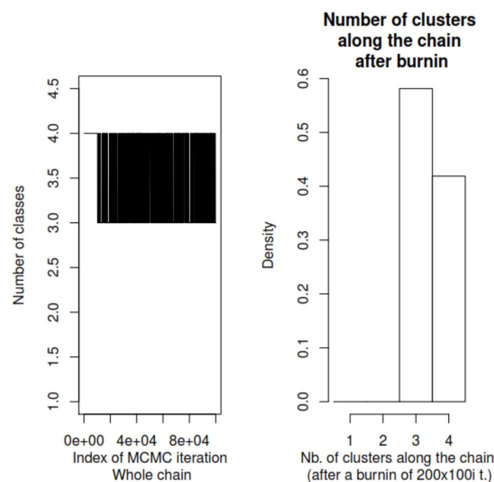


Fig. S3.5 Results generated in Geneland. a) Results from the MCMC iterations and b) Showing the most probable number of clusters *K* = 3, for 3 different South African populations from the MCMC iterations.

Chapter 4 Concluding remarks

Over the years the suni antelope (*Nesotragus moschatus*) had various subspecies described, each according to phenotypical features as well as the geographic regions in which they were found. However, to date uncertainty remains as to which subspecies exist, where they are distributed, and which ones should be recognised. According to the IUCN Red List there is approximately 365,000 suni (IUCN SSC Antelope Specialist Group 2016) with the current conservation status at of Least Concern (IUCN SSC Antelope Specialist Group 2016) and Endangered on the Red List of Mammals of South Africa, Lesotho and Swaziland (Hunnicutt et al. 2016). In addition, very little genetic research has been performed on suni; one study highlighted chromosomal variations being present between subspecies, raising the concern about possible hybridization between these subspecies in the wild, should there be lower fitness due to chromosomal differences (Kingswood et al. 1998).

In the present study, I aimed to characterize the differences between suni subspecies especially ones from south of the Zambezi River, as well as investigate the degree of gene flow taking place between the populations found in that region. In order to achieve these aims I had to make use of non-invasive methods (dung sampling) which enabled me to compare two extraction kits and consequently showing that the NucleoSpin DNA Stool kit (Macherey-Nagel) performed the best. This led to the development of an optimized extraction protocol for difficult samples. Besides this, this study allowed formulation of guidelines on how to optimise microsatellite markers not developed for the studied species, with the use of highly degraded faecal samples.

In conjunction with this, a key result is the identification of two Evolutionarily Significant Units (ESU) (Mnemba Island and southern Africa) and two possible Management Units (MU) (Mozambique and South Africa) in southern and eastern Africa with the use of population genetic approaches. Population structure within South Africa on the other hand was also examined and showed limited structure, suggesting that gene flow is taking place between Tembe Elephant Park (TEM), Ndumo Game Reserve (NDU) and Phinda Private Game Reserve (PHI). Apart from this, low genetic diversity within suni was also observed. However, this observation could be due to the fact that most of the locations included in the present study occurred on the edges of the species' distribution, where lower levels of diversity may be expected (Hardie and Hutchings 2010).

Although the results obtained in this study were ground breaking, some limitations were experienced. To elaborate one of these were the number of microsatellite markers used in this study. Initially 16 microsatellite markers were identified however only five of them were applied during the population genetics analyses. This was due to non-amplification, difficulty scoring and low polymorphism which could be explained by low quality samples, heterologous amplification and low numbers of samples. Hence, it is suggested that the suni genome be sequenced and mined for polymorphic markers and to apply those on more and high-quality samples.

Small sample sizes were another limitation. Making use of small sample sizes during population studies are not ideal since the data generated could give a false impression about what is actually

taking place within the species. Lower number of alleles, than what was present, might be identified which could create the impression that the population has low genetic diversity and possible inbreeding. Due to this it is advised that more samples from each locality and even sampling across localities should be included in future studies. In addition, sampling range as well as the distribution covered in this study were also very limited. From the results of this study it is advised that the sampling range be broadened, e.g. in Mozambique specifically in such a way that more samples are collected north of the Zambezi River. Coupled with this, samples from other localities in Africa should also be included since it could add additional insight into the suni antelope's current subspecies and distribution.

Lastly, as previously mentioned, sample quality also played an important role in this study's observations. Moreover, this factor was the reason why the number of samples that were evaluated in each location was so low. Initially several samples were evaluated however, these had to be reduced to the numbers shown in this dissertation due to amplification difficulty. In the end, only the individuals with complete genotypes were analysed to avoid biases and I recommend that only the freshest dung samples should be used.

Aside from these limitations, this study was able to achieve its goal and has impact in the broader context. In the first place, gene flow within South Africa was identified which means that good conservation management will be necessary to ensure that gene flow remains between the isolated reserves. However, other results obtained here could also be taken into consideration so that the current conservation management plans of this species could be improved. Additionally, this study also forms the first foundation layer of work on the population genetics of this species, which could only add to future studies as well as the limited current knowledge of this species. Moreover, the possible ESUs and MUs identified in this study can be investigated further to either confirm or disprove what was found. Hence adding information to the species current situation. Also, the microsatellite markers identified in this study created a starting point, which can be broadened by optimizing these markers further in suni as well as adding other markers which haven't been tested in this study. Future work on the latter could improve accuracy and add value to the conclusions drawn from the genetic data it provides.

With this said, the opportunity for more research on suni has opened up e.g. karyotypic studies can be performed on all of the suni subspecies present in the wild hence, providing more information which could aid to translocating and mixing of subspecies, especially if this become necessary for future breeding. Coupled with this, more population genetics studies on the confirmed subspecies or ESUs could also be done enabling one to evaluate genetic diversity, create better management plans and predict the species resilience to environmental changes. Future studies could also focus on a more geographical scale by doing a finer scale population genetics study along with camera trapping and dung counts. By implementing the above mentioned it will allow one to collect data over a broader context and finally obtain additional information about the suni's habits, abundance, social structure and connectivity on a finer scale. Performing research on these types of topics is really

important for suni reason being that the present knowledge of this species is very limited and by doing this, things such as species extinction as well as possible imbalance within the sand forest ecosystem can possibly be prevented.

References

Abdul-Muneer P (2014) Application of microsatellite markers in conservation genetics and fisheries management: recent advances in population structure analysis and conservation strategies. Genetics Research International 2014

Al-Samarai FR, Al-Kazaz AA (2015) Molecular markers: An introduction and applications. European Journal of Molecular Biotechnology 3:118-130

Allendorf FW (2017) Genetics and the conservation of natural populations: allozymes to genomes. Molecular Ecology 26:420-430

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nature Reviews Genetics 11:697-709

Allendorf FW, Luikart G (2009) Conservation and the genetics of populations. John Wiley and Sons

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215:403-410

Andrewartha HG, Birch LC (1986) The ecological web: more on the distribution and abundance of animals. University of Chicago Press

Ansell W (1971) Order Artiodactyla In: The mammals of Africa: an identification manual. Smithsonian Institute, Washington, pp 1–84

Avise J, Moritz C (1994) Molecular markers natural history and evolution. Trends in Ecology and Evolution 9:499-499

Barbara T, Palma-Silva C, Paggi GM, Bered F, Fay MF, Lexer C (2007) Cross-species transfer of nuclear microsatellite markers: potential and limitations. Molecular Ecology 16:3759-3767

Barendse W, Armitage S, Kossarek L, Shalom A, Kirkpatrick B, Ryan A, Clayton D, Li L, Neibergs H, Zhang N (1994) A genetic linkage map of the bovine genome. Nature Genetics 6:227

Bärmann EV, Schikora T (2014) The polyphyly of *Neotragus*–Results from genetic and morphometric analyses. Mammalian Biology 79:283-286

Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. Molecular Ecology Resources 9:1279-1301

Belton LE, Dalerum F, Van Rensburg BJ (2008) Factors associated with suni distribution in Tembe Elephant Park, South Africa: implications for management and conservation of a small antelope. African Journal of Ecology 46:631-636

Benavides FJ, Guénet J-L (2012) Mouse Genomics. In: Hedrich HJ (ed) The Laboratory Mouse, Second edn. Academic Press, Boston, pp 57-90

Benirschke K, Kumamoto A, Olsen J, Williams M, Oosterhuis J (1984) On the chromosomes of Gazella soemmeringi Cretzschmar, 1826. Zeitschrift für Säugetierkunde 49:368-373

Bhagavatula J, Singh L (2006) Genotyping faecal samples of Bengal tiger *Panthera tigris tigris* for population estimation: A pilot study. BMC Genetics 7:48

Bhebhe E, Kogi J, Holder D, Arevalo E, Derr J, Linn R, Ruvuna F, Davis S, Taylor J (1994) Caprine microsatellite dinucleotide repeat polymorphisms at the SR-CRSP-6, SR-CRSP-7, SR-CRSP-8, SR-CRSP-9 and SR-CRSP-10 loci. Animal Genetics 25:203-203

Bishop MD, Kappes SM, Keele JW, Stone RT, Sunden S, Hawkins GA, Toldo SS, Fries R, Grosz MD, Yoo J (1994) A genetic linkage map for cattle. Genetics 136:619-639

Boulet M, Couturier S, Cote SD, Otto RD, Bernatchez L (2007) Integrative use of spatial, genetic, and demographic analyses for investigating genetic connectivity between migratory, montane, and sedentary caribou herds. Molecular Ecology 16:4223-4240

Bourgeois S, Kaden J, Senn H, Bunnefeld N, Jeffery KJ, Akomo-Okoue EF, Ogden R, McEwing R (2019) Improving cost-efficiency of faecal genotyping: New tools for elephant species. PLoS ONE 14:e0210811

Bowkett AE, Plowman AB, Stevens JR, Davenport TR, van Vuuren BJ (2009) Genetic testing of dung identification for antelope surveys in the Udzungwa Mountains, Tanzania. Conservation Genetics 10:251-255

Bradburd GS, Coop GM, Ralph PL (2018) Inferring continuous and discrete population genetic structure across space. Genetics 210:33-52

Brezinsky L, Kemp SJ, Teale AJ (1993) ILSTS005: a polymorphic bovine microsatellite. Animal Genetics 24:73

Brooks TM, Pimm SL, Oyugi JO (1999) Time lag between deforestation and bird extinction in tropical forest fragments. Conservation Biology 13:1140-1150

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. Current Opinion in Genetics and Development 3:939-943

Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology and Evolution 18:249-256

Byrne B, Stone E, Mgoola WO, Uzal A (2019) First confirmed record of Suni (*Neotragus moschatus*) and Rusty-spotted genet (*Genetta maculata*) in Nyika National Park, Malawi. African Journal of Ecology 57:415-418

Caragiulo A, Pickles RSA, Smith JA, Smith O, Goodrich J, Amato G (2015) Tiger (*Panthera tigris*) scent DNA: a valuable conservation tool for individual identification and population monitoring. Conservation Genetics Resources 7:681-683

Caughley G (1994) Directions in conservation biology. Journal of Animal Ecology 63:215-244

Cerdeira JO, Pinto LS, Cabeza M, Gaston KJ (2010) Species specific connectivity in reserve-network design using graphs. Biological Conservation 143:408-415

Chapuis M, Estoup A (2006) FreeNA: microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution 24:621-631

Chiou KL, Bergey CM (2017) FecalSeq: methylation-based enrichment for noninvasive population genomics from feces. bioRxiv:032870

Chiou KL, Bergey CM (2018) Methylation-based enrichment facilitates low-cost, noninvasive genomic scale sequencing of populations from feces. Scientific Reports 8:1975

Chistiakov DA, Hellemans B, Volckaert FA (2006) Microsatellites and their genomic distribution, evolution, function and applications: a review with special reference to fish genetics. Aquaculture 255:1-29

Chytrý M, Mucina L, Rutherford M (2008) The vegetation of South Africa, Lessotho and Swaziland. South African National Biodiversity Institute, Pretoria

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657-1659

Coates DJ, Byrne M, Moritz C (2018) Genetic diversity and conservation units: Dealing with the species-population continuum in the age of genomics. Frontiers in Ecology and Evolution 6:165

Collier P, Conway G, Venables T (2008) Climate change and Africa. Oxford Review of Economic Policy 24:337-353

Costa V, Rosenbom S, Monteiro R, O'Rourke SM, Beja-Pereira A (2017) Improving DNA quality extracted from fecal samples—a method to improve DNA yield. European Journal of Wildlife Research 63:3

Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. Marine Science 1:443-466

Crandall KA, Bininda-Emonds OR, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. Trends in Ecology and Evolution 15:290-295

Creighton P, Eggen A, Fries R, Jordan S, Hetzel J, Cunningham E, Humphries P (1992) Mapping of bovine markers CYP21, PRL, and BOLA DRBP1 by genetic linkage analysis in reference pedigrees. Genomics 14:526-528

Cushman SA, Landguth EL, Flather CH (2013) Evaluating population connectivity for species of conservation concern in the American Great Plains. Biodiversity and Conservation 22:2583-2605

De Bondi N, White JG, Stevens M, Cooke R (2010) A comparison of the effectiveness of camera trapping and live trapping for sampling terrestrial small-mammal communities. Wildlife Research 37:456-465

De Queiroz K (2007) Species concepts and species delimitation. Systematic Biology 56:879-886

DeSalle R, Amato G (2004) The expansion of conservation genetics. Nature Reviews Genetics 5:702-712

Desjardins P, Conklin D (2010) NanoDrop microvolume quantitation of nucleic acids. Journal of Visualized Experiments 22:e2565

Dietz AB, Neibergs HL, Womack JE (1992) Assignment of eight loci to bovine syntenic groups by use of PCR: extension of a comparative gene map. Mammalian Genome 3:106-111

Dizon AE, Lockyer C, Perrin WF, Demaster DP, Sisson J (1992) Rethinking the stock concept: a phylogeographic approach. Conservation Biology 6:24-36

Dobzhansky T (1970) Genetics of the evolutionary process. Columbia University Press, New York

Dodge Y (2006) The Oxford dictionary of statistical terms. Oxford University Press

Du Y, Zou X, Xu Y, Guo X, Li S, Zhang X, Su M, Ma J, Guo S (2016) Microsatellite loci analysis reveals post-bottleneck recovery of genetic diversity in the tibetan antelope. Scientific Reports 6:35501

Dupuis JR, Roe AD, Sperling FAH (2012) Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. Molecular Ecology 21:4422-4436

Dyrendahl I, Gustavsson I (1979) Sexual functions, semen characteristics and fertility of bulls carrying the 1/29 chromosome translocation. Hereditas 90:281-289

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359-361

Eastwood EK, López EH, Drew JA (2016) Population Connectivity Measures of Fishery-Targeted Coral Reef Species to Inform Marine Reserve Network Design in Fiji. Scientific Reports 6:19318

Ede A, Pierson C, Crawford A (1995) Ovine microsatellites at the OarCP9, OarCP16, OarCP20, OarCP21, OarCP23 and OarCP26 loci. Animal Genetics 26:129-130

Eeley HA, Lawes MJ, Piper SE (1999) The influence of climate change on the distribution of indigenous forest in KwaZulu-Natal, South Africa. Journal of Biogeography 26:595-617

Ehrlich PR (1991) Biodiversity studies: science and policy. Science 253:758-762

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics 5:435-445

Engel SR, Linn RA, Taylor JF, Davis SK (1996) Conservation of microsatellite loci across species of artiodactyls: implications for population studies. Journal of Mammalogy:504-518

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620

Fernández MH, Vrba ES (2005) A complete estimate of the phylogenetic relationships in Ruminantia: a dated species-level supertree of the extant ruminants. Biological Reviews 80:269-302

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

Ferreira CM, Sabino-Marques H, Barbosa S, Costa P, Encarnação C, Alpizar-Jara R, Pita R, Beja P, Mira A, Searle JB (2018) Genetic non-invasive sampling (gNIS) as a cost-effective tool for monitoring elusive small mammals. European Journal of Wildlife Research 64:46

Festa-Bianchet M, Apollonio M (2003) Animal behavior and wildlife conservation. Island Press

Figueira WF, Crowder LB (2006) Defining patch contribution in source-sink metapopulations: the importance of including dispersal and its relevance to marine systems. Population Ecology 48:215-224

Fiske A (2011) The Aders's Duikers (*Cephalophus adersi*) of Mnemba Island, Zanzibar: A Study of the Behavior and Diet of a Critically Endangered Species. Digital Collections. https://digitalcollections.sit.edu/isp_collection/1023. Accessed 19 December 2018

Flagstad Ø, Røed K, Stacy J, Jakobsen K (1999) Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. Molecular Ecology 8:879-883

Forbes SH, Boyd DK (1997) Genetic structure and migration in native and reintroduced Rocky Mountain wolf populations. Conservation Biology 11:1226-1234

Frankham R (2010a) Challenges and opportunities of genetic approaches to biological conservation. Biological Conservation 143:1919-1927

Frankham R (2010b) Where are we in conservation genetics and where do we need to go? Conservation Genetics 11:661-663

Frankham R, Ballou JD, Eldridge MD, Lacy RC, Ralls K, Dudash MR, Fenster CB (2011) Predicting the probability of outbreeding depression. Conservation Biology 25:465-475

Frankham R, Briscoe DA, Ballou JD (2002) Introduction to conservation genetics. Cambridge University Press

Frankham R, Gilligan DM, Morris D, Briscoe DA (2001) Inbreeding and extinction: effects of purging. Conservation Genetics 2:279-284

Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Molecular Ecology 10:2741-2752

Frost W, Carnaby T (2015) The antelope of Africa. Jacana Media

Futuyma D (1998) Evolutionary biology, 3rd edn. Sinauer Associates, Sunderland Massachusetts

García-Dorado A (2012) Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection. Genetics 190:1461-1476

Gatesy J, Amato G, Vrba E, Schaller G, DeSalle R (1997) A cladistic analysis of mitochondrial ribosomal DNA from the Bovidae. Molecular Phylogenetics and Evolution 7:303-319

Gaubert P, Njiokou F, Olayemi A, Pagani P, Dufour S, Danquah E, Nutsuakor MEK, Ngua G, Missoup AD, Tedesco PA (2015) Bushmeat genetics: setting up a reference framework for the DNA typing of African forest bushmeat. Molecular Ecology Resources 15:633-651

Gaugris J, Matthews W, Van Rooyen M, Bothma JdP (2004) The vegetation of Tshanini Game Reserve and a comparison with equivalent units in the Tembe Elephant Park in Maputaland, South Africa. Koedoe 47:9-29

Gaugris J, Van Rooyen M (2007) The structure and harvesting potential of the sand forest in Tshanini Game Reserve, South Africa. South African Journal of Botany 73:611-622

Georges M, Massey J (1992) Polymorphic DNA markers in Bovidae. World Intellectual Property Organization, Geneva

Georgiadis NJ, Kat PW, Oketch H, Patton J (1990) Allozyme divergence within the Bovidae. Evolution 44:2135-2149

Gibson G, Reed LK (2008) Cryptic genetic variation. Current Biology 18:R989-R990

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

Gray TN, Vidya T, Potdar S, Bharti D, Sovanna P (2014) Population size estimation of an Asian elephant population in eastern Cambodia through non-invasive mark-recapture sampling. Conservation Genetics 15:803-810

Greenbaum G, Templeton AR, Bar-David S (2016) Inference and analysis of population structure using genetic data and network theory. Genetics 202:1299-1312

Grimaldi M-C, Crouau-Roy B (1997) Microsatellite allelic homoplasy due to variable flanking sequences. Journal of Molecular Evolution 44:336-340

Grover A, Sharma P (2016) Development and use of molecular markers: past and present. Critical Reviews in Biotechnology 36:290-302

Groves C, Grubb P (2011) Ungulate taxonomy. JHU Press, Baltimore

Guillot G, Mortier F, Estoup A (2005) GENELAND: a computer package for landscape genetics. Molecular Ecology Notes 5:712-715

Guschanski K, Vigilant L, McNeilage A, Gray M, Kagoda E, Robbins MM (2009) Counting elusive animals: comparing field and genetic census of the entire mountain gorilla population of Bwindi Impenetrable National Park, Uganda. Biological Conservation 142:290-300

Hájková P, Zemanová B, Bryja J, Hájek B, Roche K, Tkadlec E, Zima J (2006) Factors affecting success of PCR amplification of microsatellite loci from otter faeces. Molecular Ecology Notes 6:559-562

Hanski I (1998) Metapopulation dynamics. Nature 396:41

Hardie D, Hutchings J (2010) Evolutionary ecology at the extremes of species' ranges. Environmental Reviews 18:1-20

Harrison HB, Feldheim KA, Jones GP, Ma K, Mansour H, Perumal S, Williamson DH, Berumen ML (2014) Validation of microsatellite multiplexes for parentage analysis and species discrimination in two hybridizing species of coral reef fish (*Plectropomus spp.*, Serranidae). Ecology and Evolution 4:2046-2057

Harrison S, Taylor AD (1997) Empirical evidence for metapopulation dynamics. Metapopulation Biology: Ecology, Genetics, and Evolution:27-42

Hartl D, Clark A (1989) Principles of population genetics, 2nd edn. Sinauer Associates, Sunderland Massachusetts

Hedges S, Johnson A, Ahlering M, Tyson M, Eggert LS (2013) Accuracy, precision, and cost-effectiveness of conventional dung density and fecal DNA based survey methods to estimate Asian elephant (*Elephas maximus*) population size and structure. Biological Conservation 159:101-108

Hedrick P (2000) Genetics of populations, 2nd edn. Johns and Bartlett Publishers, Boston

Hedrick PW (1994) Purging inbreeding depression and the probability of extinction: full-sib mating. Heredity 73:363

Hedrick PW, Garcia-Dorado A (2016) Understanding inbreeding depression, purging, and genetic rescue. Trends in Ecology and Evolution 31:940-952

Hedrick PW, Hurt CR (2012) Conservation genetics and evolution in an endangered species: research in Sonoran topminnows*. Evolutionary Applications 5:806-819

Heinichen IG (1972) Preliminary Notes on the Suni, *Nesotragus Moschatus* and Red Duiker, *Cephalophus Natalensis.* Zoologica Africana 7:157-165

Hernández D, Casane D, Chevalier-Monteagudo P, Bernatchez L, García-Machado E (2016) Go west: a one way stepping-stone dispersion model for the cavefish *Lucifuga dentata* in western Cuba. PloS ONE 11:e0153545

Holder M, Lewis PO (2003) Phylogeny estimation: traditional and Bayesian approaches. Nature Reviews Genetics 4:275-284

Holleley CE, Geerts PG (2009) Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. Bio Techniques 46:511-517

Hoppe P, Van Hoven W, Von Engelhardt W, Prins R, Lankhorst A, Gwynne M (1983) Pregastric and caecal fermentation in dikdik (*Madoqua kirki*) and suni (*Nesotragus moschatus*). Comparative Biochemistry and Physiology Part A: Physiology 75:517-524

Hoppe PP (1977) Comparison of voluntary food and water consumption and digestion in Kirk's dikdik and suni. African Journal of Ecology 15:41-48

Höss M, Kohn M, Pääbo S, Knauer F, Schröder W (1992) Excrement analysis by PCR. Nature 359:199

Huebinger RM, Maar TWJD, Woodruff LH, Pomp D, Louis EE (2006a) Characterization of eight microsatellite loci in Grant's gazelle (*Gazella granti*). Molecular Ecology Notes 6:1150-1151

Huebinger RM, Maar TWJD, Woodruff LH, Pomp D, Louis EE (2006b) Characterization of nine microsatellite loci in impala (*Aepyceros melampus*). Molecular Ecology Notes 6:1152-1153

Huffaker C, Berryman A, Laing J (1984) Ecological Entomology. Academic Press, New York

Hunnicutt A, Mathews WS, Somers MJ, Child MF (2016) A conservation assessment of *Nesotragus moschatus zuluensis*. In: Child MF RL, Do Linh San E, Raimondo D, Davies-Mostert HT (ed) The Red List of Mammals of South Africa, Swaziland and Lesotho. South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa, pp 1-8

Inoue M, Hirano T, Nakane S, Watanabe T, Takeda H, Yamakuchi H, Morita M, Barendse W, Sugimoto Y (1995) Five bovine polymorphic dinucleotide microsatellite loci (DIK008, DIK010, DIK015, DIK016 and DIK020). Animal Genetics 26:447a-448

IUCN (2016) IUCN SSC Antelope Specialist Group 2016. *Nesotragus moschatus*. The IUCN Red List of Threatened Species 2016: e.T14604A50191073. https://www.iucnredlist.org/species/14604/50191073. Accessed 1 April 2019

Jarne P, Lagoda PJ (1996) Microsatellites, from molecules to populations and back. Trends in Ecology and Evolution 11:424-429

Jeffery KJ, Abernethy KA, Tutin CE, Bruford MW (2007) Biological and environmental degradation of gorilla hair and microsatellite amplification success. Biological Journal of the Linnean Society 91:281-294

Jetz W, Wilcove DS, Dobson AP (2007) Projected impacts of climate and land-use change on the global diversity of birds. PLoS Biology 5:e157

Jewitt D, Goodman PS, Erasmus BF, O'Connor TG, Witkowski ET (2015) Systematic land-cover change in KwaZulu-Natal, South Africa: Implications for biodiversity. South African Journal of Science 111:01-09

Kalinowski ST (2005) hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Molecular Ecology Notes 5:187-189

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16:1099-1106

Kappes SM, Keele JW, Stone RT, McGraw RA, Sonstegard TS, Smith T, Lopez-Corrales NL, Beattie CW (1997) A second-generation linkage map of the bovine genome. Genome Research 7:235-249

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. Trends in Ecology and Evolution 17:230-241

Kemp S, Hishida O, Wambugu J, Rink A, Teale A, Longeri M, Ma R, Da Y, Lewin H, Barendse W (1995) A panel of polymorphic bovine, ovine and caprine microsatellite markers. Animal Genetics 26:299-306

Kessler DA, Shnerb NM (2015) Generalized model of island biodiversity. Physical Review E 91:042705

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561

Kingdon J, Largen M (1997) The kingdom field guide to African mammals Zoological Journal of the Linnean Society 120:479

Kingswood S, Kumamoto A, Charter S, Jones M (1998) Cryptic chromosomal variation in suni *Neotragus moschatus* (Artiodactyla, Bovidae). Animal Conservation 1:95-100

Kirkwood D, Midgley J (1999) The floristics of sand forest in northern KwaZulu-Natal, South Africa. Bothalia 29:293-304

Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences 86:6196-6200

Kohn MH, Wayne RK (1997) Facts from feces revisited. Trends in Ecology and Evolution 12:223-227

Kolodziej K, Theissinger K, Brün J, Schulz H, Schulz R (2012) Determination of the minimum number of microsatellite markers for individual genotyping in wild boar (*Sus scrofa*) using a test with close relatives. European Journal of Wildlife Research 58:621-628

Kool JT, Moilanen A, Treml EA (2013) Population connectivity: recent advances and new perspectives. Landscape Ecology 28:165-185

Kool JT, Paris CB, Barber PH, Cowen RK (2011) Connectivity and the development of population genetic structure in Indo-West Pacific coral reef communities. Global Ecology and Biogeography 20:695-706

Kossarek L, Su X, Grosse W, Finlay O, Barendse W, Hetzel D, Dietz AB, Womack J, McGraw R (1993) Rapid communication: bovine dinucleotide repeat polymorphism RM067. Journal of Animal Science 71:3178-3178

Kovach AI, Litvaitis MK, Litvaitis JA (2003) Evaluation of fecal mtDNA analysis as a method to determine the geographic distribution of a rare lagomorph. Wildlife Society Bulletin 31:1061-1065

Krebs CJ (1994) Ecology: The experimental analysis of distribution and abundance. HarperCollins College Publishers, New York

Kumar S, Tamura K, Nei M (1994) MEGA: molecular evolutionary genetics analysis software for microcomputers. Bioinformatics 10:189-191

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

Lampa S, Gruber B, Henle K, Hoehn M (2007) An optimisation approach to increase DNA amplification success of otter faeces. Conservation Genetics 9:201

Landguth EL, Fedy BC, Oyler-McCance SJ, Garey AL, Emel SL, Mumma M, Wagner HH, Fortin MJ, Cushman SA (2012) Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. Molecular ecology resources 12:276-284

Lapedes D (1978) McGraw-Hill Dictionary of Scientific and Technical Terms, 2nd edn. McGraw-Hill, New York

Laurance WF, Lovejoy TE, Vasconcelos HL, Bruna EM, Didham RK, Stouffer PC, Gascon C, Bierregaard RO, Laurance SG, Sampaio E (2002) Ecosystem decay of Amazonian forest fragments: a 22-year investigation. Conservation Biology 16:605-618

Lawson D (1986) The ecology and conservation of suni in Natal. University of Natal, Pietermaritzburg

Lawson D (1989) The food habits of suni antelopes (*Neotragus moschatus*) (Mammalia: Artiodactyla). Journal of Zoology 217:441-448

Lebreton J-D, Burnham KP, Clobert J, Anderson DR (1992) Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs 62:67-118

Lillycrop KA, Burdge GC (2014) Environmental challenge, epigenetic plasticity and the induction of altered phenotypes in mammals. Epigenomics 6:623-636

Litvaitis MK, Litvaitis JA (1996) Using mitochondrial DNA to inventory the distribution of remnant populations of New England cottontails. Wildlife Society Bulletin 24:725-730

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? Molecular Ecology 19:3038-3051

Manne LL, Pimm SL (2001) Beyond eight forms of rarity: which species are threatened and which will be next? In: Animal Conservation Forum, pp. 221-229. Cambridge University Press.

Matthee CA, Davis SK (2001) Molecular insights into the evolution of the family Bovidae: a nuclear DNA perspective. Molecular Biology and Evolution 18:1220-1230

Matthee CA, Robinson TJ (1999) Cytochrome b Phylogeny of the Family Bovidae: Resolution within the Alcelaphini, Antilopini, Neotragini, and Tragelaphini. Molecular Phylogenetics and Evolution 12:31-46

Matthews W, Van Wyk A, Van Rooyen N, Botha G (2001) Vegetation of the Tembe Elephant Park, Maputaland, South Africa. South African Journal of Botany 67:573-594

Matthews WS (2005) Contributions to the ecology of Maputaland, southern Africa, with emphasis on Sand Forest. Dissertation, University of Pretoria South Africa

McElhany P, Ruckelshaus M, Ford M, Wainwright T, Bjorkstedt E (2000) Viable salmonid populations and the recovery of evolutionarily significant units. US Dept. Commer. NOAA Tech Memo NMFS-NWFSC 42:156

McMichael AJ, Woodruff RE, Hales S (2006) Climate change and human health: present and future risks. The Lancet 367:859-869

McNeely JA, Miller KR, Reid WV, Mittermeier RA, Werner TB (1990) Conserving the world's biological diversity. International Union for conservation of nature and natural resources

Miller SM, Clarke AB, Bloomer P, Guthrie AJ, Harper CK (2016) Evaluation of microsatellites for common ungulates in the South African wildlife industry. Conservation Genetics Resources 8:329-341

Mommens G, Van Zeveren A, Peelman L (1998) Effectiveness of bovine microsatellites in resolving paternity cases in American bison, *Bison bison* L. Animal Genetics 29:12-18

Mondol S, Karanth KU, Kumar NS, Gopalaswamy AM, Andheria A, Ramakrishnan U (2009) Evaluation of non-invasive genetic sampling methods for estimating tiger population size. Biological Conservation 142:2350-2360

Montgelard C, Zenboudji S, Ferchaud A-L, Arnal V, van Vuuren BJ (2014) Landscape genetics in mammals. Mammalia 78:139-157

Moore S, Byrne K, Berger K, Barendse W, McCarthy F, Womack J, Hetzel D (1994) Characterization of 65 bovine microsatellites. Mammalian Genome 5:84-90

Morin PA, Chambers KE, Boesch C, Vigilant L (2001) Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). Molecular Ecology 10:1835-1844

Morin PA, Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. Trends in Ecology and Evolution 19:208-216

Morin PA, Moore JJ, Chakraborty R, Jin L, Goodall J, Woodruff DS (1994) Kin selection, social structure, gene flow, and the evolution of chimpanzees. Science 265:1193-1201

Moritz C (1994a) Applications of mitochondrial DNA analysis in conservation: a critical review. Molecular Ecology 3:401-411

Moritz C (1994b) Defining'evolutionarily significant units' for conservation. Trends in Ecology and Evolution 9:373-374

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology 51:238-254

Mucina L, Rutherford MC (2006) The vegetation of South Africa, Lesotho and Swaziland. South African National Biodiversity Institute

Murphy MA, Waits LP, Kendall KC (2003) The influence of diet on faecal DNA amplification and sex identification in brown bears (*Ursus arctos*). Molecular Ecology 12:2261-2265

Narina SS, d'Orgeix CA, Sayre BL (2011) Optimization of PCR conditions to amplify microsatellite loci in the bunchgrass lizard (*Sceloporus slevini*) genomic DNA. BMC Research Notes 4:26-26

Narum S, Banks M, Beacham T, Bellinger M, Campbell M, Dekoning J, Elz A, Guthrie Iii C, Kozfkay C, Miller K (2008) Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. Molecular Ecology 17:3464-3477

National Research Council (1999) Perspectives on biodiversity: valuing its role in an everchanging world. National Academies Press, Washington

Ntie S, Johnston A, Mickala P, Bowkett AE, Van Vuuren BJ, Colyn M, Telfer P, Maisels F, Hymas O, Rouyer R (2010) A molecular diagnostic for identifying central African forest artiodactyls from faecal pellets. Animal Conservation 13:80-93

O'Connor T, Kuyler P (2009) Impact of land use on the biodiversity integrity of the moist sub-biome of the grassland biome, South Africa. Journal of Environmental Management 90:384-395

Ogliari JB, Boscariol RL, Camargo LE (2000) Optimization of PCR amplification of maize microsatellite loci. Genetics and Molecular Biology 23:395-398

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GV, Underwood EC, D'amico JA, Itoua I, Strand HE, Morrison JC (2001) Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. BioScience 51:933-938

Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma RK, Hedrick PW (2010) Conservation genetics in transition to conservation genomics. Trends in Genetics 26:177-187

Payseur BA, Cutter AD (2006) Integrating patterns of polymorphism at SNPs and STRs. Trends in Genetics 22:424-429

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

Petkova D, Novembre J, Stephens M (2016) Visualizing spatial population structure with estimated effective migration surfaces. Nature Genetics 48:94

Piggott MP (2004) Effect of sample age and season of collection on the reliability of microsatellite genotyping of faecal DNA. Wildlife Research 31:485-493

Pinto PV, Lopes S, Mourão S, Baptista S, Siegismund HR, Van Vuuren BJ, Beja P, Ferrand N, Godinho R (2015) First estimates of genetic diversity for the highly endangered giant sable antelope using a set of 57 microsatellites. European Journal of Wildlife Research 61:313-317

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. Nature Reviews Genetics 6:847

Prins HH, De Boer WF, Van Oeveren H, Correia A, Mafuca J, Olff H (2006) Co-existence and niche segregation of three small bovid species in southern Mozambique. African Journal of Ecology 44:186-198

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959

Proença-Ferreira A, Ferreira C, Leitão I, Paupério J, Sabino-Marques H, Barbosa S, Lambin X, Alves PC, Beja P, Moreira F (2019) Drivers of survival in a small mammal of conservation concern: An assessment using extensive genetic non-invasive sampling in fragmented farmland. Biological Conservation 230:131-140

Pruett CL, Winker K (2008) The effects of sample size on population genetic diversity estimates in song sparrows *Melospiza melodia*. Journal of Avian Biology 39:252-256

Puechmaille SJ, Mathy G, Petit EJ (2007) Good DNA from bat droppings. Acta Chiropterologica 9:269-276

Pusey A, Wolf M (1996) Inbreeding avoidance in animals. Trends in Ecology and Evolution 11:201-206

Radhika G, Aravindakshan T, Jinty S, Ramya K (2018) Evaluation of genetic diversity, population structure, and relationship between legendary Vechur cattle and crossbred cattle of Kerala state, India. Animal Biotechnology 29:50-58

Raeymaekers JA, Maes GE, Geldof S, Hontis I, Nackaerts K, Volckaert FA (2008) Modeling genetic connectivity in sticklebacks as a guideline for river restoration. Evolutionary Applications 1:475-488

Ramesh T, Kalle R, Rosenlund H, Downs CT (2016) Native habitat and protected area size matters: Preserving mammalian assemblages in the Maputaland Conservation Unit of South Africa. Forest Ecology and Management 360:20-29

Ramón-Laca A, Gleeson D, Yockney I, Perry M, Nugent G, Forsyth DM (2014) Reliable discrimination of 10 ungulate species using high resolution melting analysis of faecal DNA. PloS ONE 9:e92043

Ramón-Laca A, Soriano L, Gleeson D, Godoy JA (2015) A simple and effective method for obtaining mammal DNA from faeces. Wildlife Biology 21:195-204

Ramón-Laca A, White DJ, Weir JT, Robertson HA (2018) Extraction of DNA from captive-sourced feces and molted feathers provides a novel method for conservation management of New Zealand kiwi (*Apteryx* spp.). Ecology and Evolution 8:3119-3130

Raven PH, Johnson GB (1986) Biology. Mosby College Publishing, St. Louis

Raymond M (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248-249

Reddy PA, Bhavanishankar M, Bhagavatula J, Harika K, Mahla RS, Shivaji S (2012) Improved methods of carnivore faecal sample preservation, DNA extraction and quantification for accurate genotyping of wild tigers. PLoS ONE 7:e46732

Reitmann N (2008) Successful non-invasive genotyping of captive and wild suni *Neotragus moschatus* from South Africa. Honours dissertation, University of Pretoria

Renjifo LM (1999) Composition changes in a subandean avifauna after long-term forest fragmentation. Conservation Biology 13:1124-1139

Rinehart TA (2004) AFLP analysis using GeneMapper® software and an Excel® macro that aligns and converts output to binary. Biotechniques 37:186-188

Rivière-Dobigny T, Doan LP, Quang NL, Maillard JC, Michaux J (2009) Species identification, molecular sexing and genotyping using non-invasive approaches in two wild bovids species: *Bos gaurus* and *Bos javanicus*. Zoo Biology 28:127-136

Robinson T, Harrison W, de Leon AP, Elder F (1997) X chromosome evolution in the suni and eland antelope: detection of homologous regions by fluorescence in situ hybridization and G-banding. Cytogenetic and Genome Research 77:218-222

Røed KH, Ernest EM, Midthjell L, Msoffe PL (2011) Identification and characterization of 17 microsatellite loci in the blue wildebeest, *Connochaetes taurinus*. Conservation Genetics Resources 3:181-183

Roeder AD, Marshall RK, Mitchelson AJ, Visagathilagar T, Ritchie PA, Love DR, Pakai TJ, McPartlan HC, Murray ND, Robinson NA (2001) Gene flow on the ice: genetic differentiation among Adélie penguin colonies around Antarctica. Molecular Ecology 10:1645-1656

Roughgarden J, May RM, Levin SA (2014) Perspectives in ecological theory. Princeton University Press

Rovero F, Jones T, Sanderson J (2005) Notes on Abbott's duiker (*Cephalophus spadix* True 1890) and other forest antelopes of Mwanihana Forest, Udzungwa Mountains, Tanzania, as revealed by camera-trapping and direct observations. Tropical Zoology 18:13-23

Rutherford MC, Westfall RH (1994) Biomes of southern Africa: an objective categorization. National Botanical Institute

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. Trends in Ecology and Evolution 1:9-10

Ryman N, Palm S (2006) POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Molecular Ecology Notes 6:600-602

Sarukhán J, Whyte A, Hassan R, Scholes R, Ash N, Carpenter S, Pingali P, Bennett E, Zurek M, Chopra K (2005) Millenium Ecosystem Assessment: Ecosystems and human well-being.

Schlötterer C (2000) Evolutionary dynamics of microsatellite DNA. Chromosoma 109:365-371

Schlötterer C (2004) The evolution of molecular markers—just a matter of fashion? Nature Reviews Genetics 5:63-69

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. Trends in Ecology and Evolution 22:25-33

Schwartz MK, Monfort SL, Long R, MacKay P, Zielinski W, Ray J (2008) Genetic and endocrine tools for carnivore surveys. In: Noninvasive survey methods for carnivores. Island Press, pp 238-262

Schwerin M, Gallagher Jr D, Miller J, Thomsen P (1992) Mapping of repetitive bovine DNA sequences on cattle Y chromosomes. Cytogenetic and Genome Research 61:189-194

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology Letters 9:615-629

Seluja G, Kumamoto A, Benirschke K (1985) A cytotaxonomic approach in the comparative study of four Neotraginae species of different genera (Artiodactyla, Bovidae). Mammalian Chromosomes Newsletter 26:39

Shi Y, Chen J, Su J, Zhang T, Wasser SK (2019) Genetic Resilience of a once endangered species, tibetan antelope (*Pantholops hodgsonii*). bioRxiv:628727

Skinner JD, Chimimba CT (2005) The mammals of the southern African sub-region. Cambridge University Press

Smith RJ, Easton J, Nhancale BA, Armstrong AJ, Culverwell J, Dlamini SD, Goodman PS, Loffler L, Matthews WS, Monadjem A (2008) Designing a transfrontier conservation landscape for the Maputaland centre of endemism using biodiversity, economic and threat data. Biological Conservation 141:2127-2138

Smouse RPP, Peakall R (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537-2539

Snedecor G, Cochran W (1967) Statistical Methods, Sixth edn. Iowa State University Press, Iowa

Sokal R, Rohlf FJ (1969) Biometry. WH Freeman and Co, San Francisco

Somers M, Rasa OAE, Marat E (1990) Marking behaviour and dominance in Suni antelope (*Neotragus moschatus*). International Journal of Mammalian Biology 55:340-352

Sonsthagen SA, Wilson RE, Underwood JG (2017) Genetic implications of bottleneck effects of differing severities on genetic diversity in naturally recovering populations: An example from Hawaiian coot and Hawaiian gallinule. Ecology and Evolution 7:9925-9934

Spitzer R, Churski M, Felton A, Heurich M, Kuijper DP, Landman M, Rodriguez E, Singh NJ, Taberlet P, van Beeck Calkoen ST (2019) Doubting dung: eDNA reveals high rates of misidentification in diverse European ungulate communities. European Journal of Wildlife Research 65:28

Steffen P, Eggen A, Stranzinger G, Fries R, Dietz AB, Womack J (1993) Isolation and mapping of polymorphic microsatellites in cattle. Animal Genetics 24:121-124

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J (1996) Reliable Genotyping of Samples with Very Low DNA Quantities Using PCR. Nucleic Acids Research 24:3189-3194

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

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

Tallmon DA, Luikart G, Waples RS (2004) The alluring simplicity and complex reality of genetic rescue. Trends in Ecology and Evolution 19:489-496

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

Tarsi K, Tuff T (2012) Introduction to population demographics. Nature Education Knowledge 3:3

Toldo SS, Fries R, Steffen P, Neiberg HL, Barendse W, Womack JE, Hetzel DJS, Stranzinger G (1993) Physically mapped, cosmid-derived microsatellite markers as anchor loci on bovine chromosomes. Mammalian Genome 4:720-727

Treydte AC, Edwards PJ, Suter W (2005) Shifts in native ungulate communities on a former cattle ranch in Tanzania. African Journal of Ecology 43:302-311

Tuomainen U, Candolin U (2011) Behavioural responses to human-induced environmental change. Biological Reviews 86:640-657

Vaiman D, Imam-Ghali M, Moazami-Goudarzi K, Guérin G, Grohs C, Levéziel H, Saïdi-Mehtar N (1994a) Conservation of a syntenic group of microsatellite loci between cattle and sheep. Mammalian Genome 5:310-314

Vaiman D, Mercier D, Moazami-Goudarzi K, Eggen A, Ciampolini R, Lépingle A, Velmala R, Kaukinen J, Varvio SL, Martin P, Levéziel H, Guérin G (1994b) A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism. Mammalian Genome 5:288-297

Vaiman D, Osta R, Mercier D, Grohs C, Leveziel H (1992) Characterization of five new bovine dinucleotide repeats. Animal Genetics 23:537-541

Väli Ü, Dombrovski V, Treinys R, Bergmanis U, Daroczi SJ, Dravecky M, Ivanovski V, Lontkowski J, Maciorowski G, Meyburg B-U (2010) Widespread hybridization between the Greater spotted eagle *Aquila clanga* and the Lesser spotted eagle *Aquila pomarina* (Aves: Accipitriformes) in Europe. Biological Journal of the Linnean Society 100:725-736

Van Hooft W, Groen A, Prins H (2000) Microsatellite analysis of genetic diversity in African buffalo (*Syncerus caffer*) populations throughout Africa. Molecular Ecology 9:2017-2025

Van Hooft W, Hanotte O, Wenink P, Groen A, Sugimoto Y, Prins H, Teale A (1999) Applicability of bovine microsatellite markers for population genetic studies on African buffalo (*Syncerus caffer*). Animal Genetics 30:214-220

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

Van Vliet N, Zundel S, Miquel C, Taberlet P, Nasi R (2008) Distinguishing dung from blue, red and yellow-backed duikers through noninvasive genetic techniques. African Journal of Ecology 46:411-417

Van Wyk AM, Kotzé A, Grobler JP, Van Vuuren BJ, Barrow LN, Dalton DL (2018) Isolation and characterization of species-specific microsatellite markers for blue and black wildebeest (*Connochaetes taurinus* and *C. gnou*). Journal of Genetics 97:101-109

Vieira MLC, Santini L, Diniz AL, Munhoz CdF (2016) Microsatellite markers: what they mean and why they are so useful. Genetics and Molecular Biology 39:312-328

Vigilant L (2002) Technical challenges in the microsatellite genotyping of a wild chimpanzee population using feces. Evolutionary Anthropology: Issues, News, and Reviews 11:162-165

Vogler AP, Desalle R (1994) Diagnosing units of conservation management. Conservation Biology 8:354-363

Waples RS (1991) Pacific salmon, *Oncorhynchus* spp., and the definition of "species" under the Endangered Species Act. Marine Fisheries Review 53:11-22

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology 15:1419-1439

Wegmann M, Santini L, Leutner B, Safi K, Rocchini D, Bevanda M, Latifi H, Dech S, Rondinini C (2014) Role of African protected areas in maintaining connectivity for large mammals. Philosophical Transactions of the Royal Society of London B: Biological Sciences 369:20130193

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

Weikard R, Pitra C, Kühn C (2006) Amelogenin cross-amplification in the family Bovidae and its application for sex determination. Molecular Reproduction and Development 73:1333-1337

Williams G (1966) Adaptation and Natural Selection. Princeton University Press, New Jersey

Wilson DE, Reeder DM (2005) Mammal species of the world: a taxonomic and geographic reference. JHU Press

Wolffs P, Grage H, Hagberg O, Rådström P (2004) Impact of DNA polymerases and their buffer systems on quantitative real-time PCR. Journal of Clinical Microbiology 42:408-411

Woodruff DS (2001) Populations, species and conservation genetics. Encyclopedia of Biodiversity 4:811-829

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97

Wultsch C, Waits LP, Hallerman EM, Kelly MJ (2015) Optimizing collection methods for noninvasive genetic sampling of neotropical felids. Wildlife Society Bulletin 39:403-412

Yeb C, Kogi J, Holder M, Guerra T, Davis S, Taylor J (1997) Caprine microsatellite dinucleotide repeat polymorphisms at the SR-CRSP 21, 22, 23, 24, 25, 26 and 27 loci. Animal Genetics 28:370-371

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Molecular Ecology 12:563-584