MATERNAL LINEAGES AND DIVERSITY OF THE GROWTH HORMONE GENE OF SOUTH AFRICAN GOAT POPULATIONS

By

Keabetswe Tebogo Ncube

Submitted in accordance with the requirements for

The degree of

MASTER OF SCIENCE

In the subject

STATISTICAL GENOMICS

at the

UNIVERSITY OF SOUTH AFRICA

Supervisor: Dr. F. C. Muchadeyi

(SEPTEMBER 2016)

Student number: **53754166**

I Keabetswe Tebogo Ncube declare that **__Maternal Lineages And Diversity Of The Growth Hormone Gene Of South African Goat Populations___** is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

_____ _____

SIGNATURE DATE

Ms Ncube K.T (Student)

_____ _____

SIGNATURE DATE

Dr. F.C. Muchadeyi (Supervisor)

DEDICATION

To my dear mother, Mojaki Salmina Ncube.

ACKNOWLEDGEMENTS

Firstly, I would like to thank The Almighty God, the only Living God for hearing and answering all my prayers and giving me such a great opportunity to embark on a journey that yielded this dissertation, also for His unfailing love and constant guidance through His Word. All glory and honour and praise be to the Most High God, Majesty.

I am more than pleased to acknowledge those who played a role in the build-up of this thesis. Words cannot express my earnest gratitude to my supervisor Dr. Farai Catherine Muchadeyi for her support, explicit guidance, constructive criticism and dedication to my work. My greatest gratitude to Dr. Jasper Rees for seeing and believing in my potential and giving me a chance to join the Biotechnology Platform team. I would also like to acknowledge the Agricultural Research Council, Professional Development Programe (ARC-PDP) and National Research Foundation (NRF) for believing in me and taking care of my upkeep as I studied. My sincere gratitude goes to the Agricultural Research Council, Biotechnology Platform (ARC-BTP) for allowing me to access all their laboratory and other equipments especially the all fancy and awesomely cool Illumina Next Generation Sequencing Technology equipment in order to make this dissertation a success. I am grateful to the University of South Africa, Department of Life and Consumer Sciences for allowing me to be a part of their academic family.

Many thanks go to my mother, O' mommy how can I say thank you for your unending support, love, care, advice and constant encouragement and many other things that I did not mention I have come short of words to express my thanks to you. A great thanks to my grandmother Rachel N. Moerane for her words of wisdom throughout this journey. A big thanks to my sisters Seapei Lesejane and Tsholofelo Ncube for their support and wonderful advice throughtout. I thank you so much. Thank you Granddad L. Tabane for believing in me so much as to seek goat farming

advice from me and for attending a farmer's workshop where I was one of the speakers. I really appreciate this.

To my peers and sisters that I have made in the process; K. Mdladla, D.P. Malatji, and T. Nhlapo thank you for the support, inspiration. K Mdladla, your mentorship and motivation is highly appreciated. To Charles K. Wairuri, I am more than grateful for all your assistance, patience, guidance and great advice on the mitochondrial DNA experiment, your expert advice has really brought this chapter to light. Special thanks are due to all the Biotechnology Platform family for constructive critisism during journal club. I am also greatly indebted to farmers and villagers from Goodwill and van Tonder farms in the North West Province, Carnavon Research Station in Northern Cape, Hobeni village Pechelsdam village in Eastern Cape, Coniliva and Ngubo villages in Kwa-Zulu Natal, Mukovhabale village and Muila-muumone in Limpopo province and Pella village in North West province who allowed us to sample in their villages and farms without any hustles. My greatest gratitude to Chief G. Gasebone of the Bakwena ba Morare community in Pella village for his warm welcome in his village, smooth operation and for allowing us to hold a successful workshop for the farmers. Lastly, to all those that I did not mention by name – you are important!

SUMMARY

MATERNAL LINEAGES AND DIVERSITY OF THE GROWTH HORMONE GENE OF SOUTH AFRICAN GOAT POPULATIONS

By

Keabetswe Tebogo Ncube

Correspondence: NcubeK@arc.agric.za

Supervisor: Dr. F.C. Muchadeyi

Department: Life and Consumer Sciences

Summary

The maternal lineages and origins of the South African goat populations are unknown and hence pose challenges for breed characterization and conservation. This study investigated the maternal lineages of South African goats using complete mtDNA and ascertained the genetic diversity in the growth hormone gene within and between populations. Illumina MiSeq next generation sequencing was used to generate the full length of the mtDNA (16.64 kb) and growth hormone (2.54kb) genes in 50 goats of the commercial South African Boer (n =9), captive feral Tankwa (n =9), and SA village goat populations (n =32). The non-descript village populations were sampled from villages of the four major goat-producing provinces; (i) Hobeni village, Elliotdale municipality and Pechelsdam village, Inxubayethemba municipality in Eastern Cape (n=8), (ii) Coniliva and Ngubo villages in Msinga municipality Kwa-Zulu Natal (n=8), (iii) Mukovhabale village, Mutale municipality and Muila-muumone, Makhado municipality in Limpopo (n=8) and (iv) Pella village (n=6), Moses Kotane municipality North West (n=8) provinces of South Africa.

A total of 184 SNPs and 55 AA changes were observed across the complete mtDNA

genome. High within-population variation was observed in all the groups, ranging from 98.60 to 99.52%. A low F_{ST} (F_{ST} = 0.003-0.049) indicated close relatedness and possible gene flow between SA goat populations. Haplotypes and clades observed in the D-loop, COX1 and whole mtDNA network trees demonstrated relationships between South African goat populations. The South African goats clustered with Chinese goats from lineages A and B, suggesting common maternal lineages between the Chinese and South African goat populations. The results also suggested that the bezoar (*Capra aegagrus)* is a possible ancestor of South African domestic goats.

A range of 27 to 58 SNPs per population were observed on the growth hormone gene. Amino acid changes from glycine to serine, tyrosine to cysteine and arginine to glycine were observed at exon 2 and exon 5. Gene diversity ranged from 0.8268 ± 0.0410 to 0.9298 ± 0.0050. Higher within breed diversity (97.37%) was observed within the population category consisting of SA village ecotypes and the Tankwa goats. Highest pairwise F_{ST} values ranging from 0.148 to 0.356 were observed between the SA Boer and both the SA village and Tankwa feral goat populations. The maximum likelihood phylogenetic analysis indicated nine genetic clades, which reflected close relationships between the South African populations and the other international breeds. Results imply greater potential for within population selection programs particularly with SA village goats.

Key words: Genetic diversity, Goats mtDNA; maternal lineages; Growth Hormone Gene

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AA	Amino Acid (s)
AMOVA	Analysis Of Molecular Variance
Arg	Arginine
ARC	Agricultural Research Council
ATP	Adenosine triphosphate
BJG	Banjiau goat
BMP	Bone Morphogenetic Protein
BR	Broad Range
COX1	Cytochrome oxidase subunit 1
CYTB	Cytochrome B
Cys	Cysteine
DNA	Deoxyribonucleic Acid
dsDNA	Double stranded Deoxyribonucleic Acid
EC	Eastern Cape
EDTA	Ethylenediamine tetraacetic acid
FAOSTAT	Food and Agricultural Organization Statistics of the United Nations
F_{ST}	The amount of differentiation among subpopulations relative to the limiting amount under complete fixation
GHG	Growth Hormone Gene
GHG-F	Growth Hormone Gene forward primer
GHG-R	Growth Hormone Gene reverse primer
GHR	Growth Hormone Receptor

Gly	Glycine
GZW	Guizhou White goat
HG	Haplogroup
HHG	Huanghuai goat
HV1	Hyper-variable region 1
IGF-1	Insulin like growth factor I
KZN	Kwa-Zulu Natal
L	Limpopo
LEP	Leptin
LZG	Leizhou goat
MAS	Marker Assisted Selection
MSTN	Myostatin
MtDNA	Mitochondrial Deoxyribonucleic Acid
NAMC	National Agricultural Marketing Council
ND1-6	NADH dehydrogenase subunit 1-6
NGS	Next generation sequencing
NMC	Neimonggol Cashmere goat
NRF	National Research Foundation
NW	North West
PCR	Polymerase Chain Reaction
POU1F1	Pituitary specific transcription factor-1
RCA	Rolling circle amplification
RNA	Ribonucleic Acid

rRNA	Ribosomal Ribonucleic Acid
SA	South Africa
SAB	South African Boer
Ser	Serine
SNP(s)	Single Nucleotide Polymorphism (s)
SNW	Shaannan White goat
TIG	Tibetan goat
tRNA	Transfer Ribonucleic Acid
Tyr	Tyrosine
TWA	Tankwa
UV	Ultraviolet

CHAPTER 1

INTRODUCTION

1.1 Background and problem statements

Goats are produced worldwide with a total count of about 975.8 million (FAOSTAT, 2013). Sub-Saharan countries such as South Africa, Nigeria, Namibia and Botswana are the major goat producing African countries in the world. South Africa alone has about 6.2 million goats (FAOSTAT 2013), contributing about 3% of Africa's goat population (Moloko 2011). Africa contributes about 33.8% of goats to the world's goat population (FAOSTAT 2008). Most of the goats in Africa are indigenous goats that are owned by smallholder farmers in small communities and remote rural areas (Lahiff & Cousins 2005; Garine 2007; Moloko 2011).

Goats raised by communal farmers are used to improve food security and alleviate poverty through sales (Simela & Merkel, 2008; Gwaze 2009). In communal areas, goats are used in ceremonial and traditional functions (Masika *et al*. 2007), such as during burials and rituals by traditional healers (Moloko 2011). Goats provide manure that is used in communal home gardens (Gwaze 2009) and many communal farmers use goats as a form of investment and financial security (Moloko 2011). The commercial goat farming sector on the other hand uses goats for meat, milk and skins (Moloko 2011; Casey and Webb 2010) and to a less extent cheese, cashmere and mohair (Moloko 2011). Another area where goats are used in South Africa and globally is in shrub or bush management (Gwaze 2009).

South Africa is one of the few countries across the world with well-defined goat breeds such as the South African Boer goat, which was developed from the indigenous goat populations (Mdladla *et al*. 2016). Breeds such as the SA Boer, Savannah, and Kalahari Red are predominantly farmed under commercial farming systems for meat production (Visser *et al*. 2004). Breeds such as the Namibian Ovambo, Caprivi, Kunene, Kavango (Els *et al*. 2004), West African Dwarf, Sahel, and Red Sokoto (Fasae *et al*. 2012) are also farmed in other African countries.

In South Africa, the most dominant meat type goats are the SA Boer, Savannah and the Kalahari Red, which are recognized as commercial breeds. The SA village goats are farmed in villages across different sectors of the country (Pieters 2007). The village goats in most of these communal farming systems, are less characterized and are hence named according to their geographical locations e.g. the Venda goats from Vhembe District in Venda, Limpopo Province of South Africa (Visser *et al*. 2004), Namibian Ovambo from the Ovambo region in Namibia (Els *et al*. 2004), West African Dwarf from West Africa (Fasae *et al.* 2012), Tswana goats from Botswana and Malawi goats from Malawi (Gwaze 2009). However, these uncharacterized goat populations constitute an important genetic resource necessary to meet food security and improve livelihoods of marginalized communities

The South African village goat populations are considered highly adapted and suitable for smallholder farming systems, however, not much is known about their actual genetic potential (Visser *et al*. 2004). The history, origin and maternal lineages of these goat populations are unknown. It has been reported that the goats came to South Africa with the Nomadic nations as they migrated from North Africa to the southern parts of the continent (Morrison 2007). Even though South Africa has well defined goat meat breeds such as the SA Boer, the origin of these breeds is also unknown. Goat origin is important in terms of breed utilization and conservation. Knowledge of the origin of breeds will enable farmers to maximize the breeds potential, utilizing its characteristic such as meat, hair, skin and etc. optimally. Breed conservation is essential for the preservation of genetic potential and prevention of breed extinction. Extensive cross breeding may cause a reduction in the size of the genetic pool, changing breed characteristics and ultimately resulting in the loss of the breed (Morrison 2007). Studies on the mtDNA diversity will provide information about the genetic diversity and conservation of these indigenous animal genetic resources. This is essential for their conservation, while their characterization will also aid in ensuring the breeds are developed appropriately for the needs of future generations (Othman *et al*. 2015).

Over the years, livestock improvement has initiated the start of well-defined breeds that resulted in changes in livestock practices (Nomura *et al*. 2013). Farmers began to substitute local breeds with commercial high performance breeds, this led to the disappearance and endangering of livestock species such as goats (Othman *et al*. 2015). The use of mtDNA has played a significant role in maternal origin determination, haplogroup identifications and breed characterization (Liu *et al*. 2007; Luikart *et al*. 2001). The D-loop analysis has previously described various maternal lineages (Liu *et al*. 2007). Examples include determination of the origin and maternal lineage in Chinese (Liu *et al*. 2007), Coastal Indian (Rana *et al*. 2013) and Korean native (Odahara *et al*. 2006) goat populations where four lineages (A-D) were reported.

Village goats are known for their ability to use poor quality feed, resistance to prevalent diseases and parasites and their adaptation to harsh environmental and climatic conditions that is characteristic of their habitat (Webb & Mamabolo, 2004). The variations of coat colours, hair length and ear shape and length and body size characteristics of the village goats form the means by which varieties/individuals are distinguished from one another (Visser *et al*. 2004; Morrison 2007). Growth rates amongst goat breeds have been found to vary with an average of 0.13-0.23 kg/day (Little 2010). Breeds such as SA Boer that have been specially developed for meat production are fast growing (NAMC 2005) whereas the village goats are reported to have wide variation in growth rates (Kosgey *et al*. 2008; Gaddour *et al*. 2012; Hassen *et al*. 2012). Great variation in body weights and a below average mature body weight of about 38kg irrespective of sex have been observed (Visser *et al*. 2004). Since commercial goat farmers select for fast growing animals, culling against poor growth performance, these animals demonstrate a much faster growth rate when compared to the village goats of communal farms (Webb & Mamabolo 2004). It has been hypothesized, however, that because they are uncharacterized and have not been artificially selected for economically important traits, village goats should possess greater genetic potential for growth and body weight traits (Webb & Mamabolo 2004; Peacock *et al*. 2011). An improvement in growth in these village goats would result in increased chevon (goat meat) yields, which would contribute to improved food security amongst marginalized communities.

Growth is the primary factor influencing the amount and expense of meat produced at mature weight (Supakon, 2009). A number of genes affect growth. These include the growth hormone (*GH*), growth hormone receptor (*GHR*), insulin like growth factor I (*IGF-I*), leptin (*LEP*), caprine pituitary specific transcription factor-1 (*POU1F1*), caprine myostatin (*MSTN*) and bone morphogenetic protein (*BMP*) genes which are necessary for developments such as muscle and weaning growth (Supakon 2009). Variations in these genes have been associated with variability in growth traits in goats (Alakilli *et al*. 2012) with some of the genes currently being used in selection and breed improvement programs. A limited number of studies have however been performed on growth related genes in African goat breeds. In Africa and other developing countries, village goats pose great genetic potential with their variation in size and ability to use poor quality feed. Characterization of this genetic diversity could provide fundamental information for future selection programs.

1.2 Justification

In the genomics era, next generation sequencing (NGS) and targeted gene capture of the associated genes can be used to generate data on single nucleotide polymorphisms (SNPs) on specific genes of interest, providing information to make inferences on genetic diversity. Next generation sequencing is a modern day technology that allows one to do routine sequencing of either complete genomes or targeted sections in a faster and more effective manner (Myllykangas *et al*. 2011). Whole genome sequencing is a method whereby the entire genome of an organism or animal is sequenced. Targeted gene capture on the other hand sequences a specific gene or region of a gene (www.illumina.com) generating data that is exclusive to that gene and hence allowing for greater accuracy with greater specificity on selected regions of the genome. This study therefore used Next generation sequencing as well as targeted gene capture in order to analyze the mitochondrial DNA genome and the growth hormone gene.

Mitochondrial DNA is a 16.64kb genome (Jeon *et al*. 2004) that can be used to study the origin and maternal lineages of livestock (Liu *et al*. 2007). The sequencing of the complete mtDNA genome using next generation sequencing technologies will enable the identification of single nucleotide polymorphisms and amino acid changes, which

can be used to ascertain within and between breed genetic diversity and to infer breed history and maternal lineages. The demonstration of the divergence of breeds from their common ancestor may provide insight into the differences that are observed in production performance traits such as growth.

Growth is a quantitative trait that is influenced by several genes, of which the growth hormone gene is one of those that play major role in the growth of mammals, (Alakilli *et al*. 2012). This gene codes for the production and release of growth hormone (Supakorn 2009), playing an essential role in mammalian metabolism, growth and milk traits in goats and sheep (Alakilli *et al*. 2012). Growth hormone gene is plays a major role in the growth of mammals and observations on its effects on growth have been observed in the muscle, bone and adipose tissues of animals (Hua *et al*. 2009). The growth hormone gene also plays a role in lactation and milk production. Genetic association studies demonstrated exon 2 and exon 4 of this gene to be associated with milk yield in the two ecotypes of Serrana goats (Mousavizadeh *et al*. 2009). The high levels of polymorphism in exon 4 of the growth hormone gene indicate its potential to be used in marker-assisted selection (MAS) for milk producing goats (Mousavizadeh *et al*. 2009). Hua *et al*. (2009) identified 10 polymorphisms in the gene of which some were associated with milk production traits. This study also reports specific genotypes to correspond with heavier animals with a mass difference of 2kg. This gene plays a big role in the growth and production of goats. Studying this gene in South African goat populations will therefore assist in evaluating the potential of growth hormone as a candidate gene for use in selection programs and growth performance improvement.

1.3 Aim and Objectives

Aim

The overall aim of this study was to investigate the maternal lineages of indigenous South African goat populations using complete mtDNA, and to establish the genetic diversity in the growth hormone gene within and between populations under different farming systems.

Objectives

- To investigate genetic diversity and population genetic structure of South African indigenous goat populations using complete mitochondrial DNA sequences.
- To investigate maternal lineages and determine whether the populations are from the same ancestor and of the same genetic background.
- To investigate polymorphisms of the growth hormone gene using targeted gene capture and next generation sequencing.
- To use the sequence data to determine within and between breed variations and to ascertain the potential of using the growth hormone gene as a tool for selection and breed improvement programs.

1.4 References

Alakilli S. Y. M., Mahrous K. F., Salem L. M. & Ahmed E. S. (2012) Genetic polymorphism of five genes associated with growth traits in goat. *African Journal of Biotechnology* 11, 14738-14748.

Casey N. H., & Webb E. C. (2010) Managing goat production for meat quality. *Small Ruminant Research*. DOI: 10.1016/J.SMALLRUMRES.2009.12.047.

Fasae O. A., Sowande O.S. & Adewumi O.O. (2012) Ruminant animal production and husbandry. Department of Animal Production and Health. University of Agriculture, Abeokuta, Nigeria.

Els J.F., Kotze A. & Swart H. (2004) Genetic diversity of indigenous goats in Namibia using microsatellite markers: preliminary results. *South African Journal of Animal Science* 34, 65-67.

Gaddour A., Ouni A. & Najari S. (2012) Growth curve estimation in pure goat breeds and crosses of first and second generation in Tunisian oases. *Journal of Cell and Animal Biology* 6, 99-103.

Garrine C.M.L.P (2007) Genetic characterization of indigenous goat populations of Mozambique. Department of Production Animal Studies in Faculty of Veterinary Science, University of Pretoria, South Africa.

Gwaze F.R. (2009) Communal Production Systems of Goats Raised by Resource-Poor Farmers in the Eastern Cape Province of South Africa. Department of Livestock and Pasture Science, University of Fort Hare, South Africa.

Han L., Geng L.Y., Shen Z., Xu D.Q. & Yang L.G. (2009) Polymorphism of the growth hormone gene and its association with growth traits in Boer goat bucks. *Meat Science* 81, 391–395.

Hassen H., Baum M, Rischkowsky B. & Tibbo M. (2012) Phenotypic characterization of Ethiopian indigenous goat populations. *African Journal of Biotechnology* 11, 13838-13846.

Jeon G. J., Chung H. Y., Choi J. G., Lee M. S., Lee C. W., Park J. J., Ha J. M., Lee H. K. & Sung H. H. (2005) Relationship between genetic variants of mitochondrial DNA and growth traits in Hanwoo cattle. *Journal of Animal Science* 18, 301-307.

Kosgey I.S., Rowlands G.J., van Arendonk J.A.M. & Baker R.L. (2009) Small ruminant production in smallholder and pastoral or extensive farming systems in Kenya. *Small Ruminant Research* 77, 11-24.

Lahiff E. & Cousins B. (2005) Smallholder Agriculture and Land Reform in South Africa. *Institude of Development Studies Bulletin* 36,127-131.

Little K. (2010) Meat goats: Sustainable livestock farming. Nuffield Farming Scholarships Trust.

Liu R. Y, Lei C. Z, Liu S. H. & Yang G. S. (2007) Genetic diversity and origin of Chinese Domestic goats revealed by complete mtDNA D-loop sequence variation. *Asian Australian Journal of Animal Science* (20), 178-183.

Luikart G., Gielly L., Excoffier L., Vigne J., Bouvet J. & Taberlet P. (2001) Multiple maternal origins and weak phylogeographic structure in domestic goats. *PNAS* 98, 5927-5932.

Moloko B. (2011) A profile of the South African goat market value chain. *Department of agriculture, fishery and forestry*.

Morrison J. W (2007) A Guide to the identification of the Natural Indigenous Goats of Southern Africa. *South African journal of Animal Science* 4.

Mousavizadeh A., Abadi M.M., Torabi A., Nassiry M.R., Ghiasi H. & Koshkoieh A. (2009) Genetic polymorphism at the growth hormone locus in Iranian Talli goats by polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP). *Iranian Journal Of Biotechnology* 7, 51-53.

Myllykangas S, Natsoulis G, Bell J. M. & Ji H. P. (2011) Targeted sequencing library preparation by genomic DNA circularization. *BMC Biotechnology* 11, 122.

National Agricultural Marketing Council, 2005. Report on the investigation into the potential for the South African goat industry. Draft 1.

Nomura K., Yonezawa T., Mano S., Kawakam S., Shedlock A. M., Hasegawa M. & Amano T. (2013) Domestication process of the goat revealed by an analysis of the nearly complete mitochondrial protein-encoding genes. *PLOS ONE*, 8, e6775.

Odahara S, Chung H. J, Choi S. H , Yu S. L, Sasazaki S, Mannen H, Park C. S. & Lee J. H. (2006) Mitochondrial DNA Diversity of Korean Native Goats. *Asian Australian Journal of Animal Science* 19, 482-485.

Othman O., E., Pariset L, Balabel E. A. & Marioti Marco. (2015) Genetic characterization of Egyptian and Italian sheep breeds using mitochondrial DNA. *Journal of Genetic Engineering and Biotechnology* 13, 79–86.

Peacock C., Ahuya C.O., Ojango J.M.K. & Okeyo A.M. (2011) Practical crossbreeding for improved livelihoods in developing countries: The FARM Africa goat project. *Livestock Science* 136, 38–44.

Pieters A. (2007) Genetic characterization of commercial goat populations in South Africa. Department of Animal Breeding and Genetics. University of Pretoria, South Africa.

Rana S., Singh S., Dureja V., Joshi J., Banerjee P., Sharma U. & Viji R. K. (2013)

Mitochondrial D-loop analysis reveals single maternal lineage for coastal region goats of India. *International Journal Of Biomedical and Life Sciences* 4, 235-244.

Simela L. & Merkel R. (2008) The contribution of chevon from Africa to global meat production. *Meat Science* 80, 101–109.

SUPAKORN C (2009) The important candidate genes in goats - A Review. *Walailak Journal of Science & Technology* 6, 17-36.

Visser C., Hefer C.A., van Marle-Köster E. & Kotze A. (2004) Genetic variation of three commercial and three indigenous goat populations in South Africa. *South African Journal of Animal Science* 34, 24-27.

Webb E. C. & Mamabolo M. J. (2004) Production and reproduction characteristics of South African indigenous goats in communal farming systems. *South African Journal of Animal Science* 34, 236-239.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

A large percentage of the indigenous goats in South Africa are the village goats that are owned by smallholder communal farmers in the rural areas (Moloko 2011). These communal farming communities are characterized by sub-optimal management practices and lack of information on goat farming practices (Masika *et al*. 2007). Keeping fast growing commercial breeds in smallholder farms is a great challenge as it is costly, due to high feed requirements (NAMC 2005) that communal farmers cannot afford (Moloko 2011). The village goats that these farmers are presently keeping are unimproved, uncharacterized and have not been subjected to any selection methods (Moloko 2011). Village goats show great potential in genetic contribution due to their diversity in growth rates and mature body weights (Visser *et al*. 2004). Variation in the ability to use poor quality feed, disease resistance and adaptation to a wide variety of environmental conditions have also been reported in indigenous goat populations (Webb & Mamabolo, 2004; Visser *et al*. 2004). It is therefore hypothesized that village goats have good genetic potential and can be selected for optimum growth under the compromised production systems.

Goat farming is a growing industry in South Africa. The understanding of the breeds that can be used for goat farming and their growth performance is therefore of great importance. Mitochondrial DNA studies have been used to give insight into the origin, evolution and genetic relatedness of a number of livestock species. Targeted gene sequencing using next generation sequencing technology is a useful tool that will bring about insight and knowledge needed in order to understand growth related genes and their association with growth traits.

This chapter therefore aims to provide a review goat production; maternal origins and genetic structure of communal goat populations in Africa, as well as polymorphisms in genes conferring improved growth rates in smallholder goat production systems. The

chapter also reviews next generation sequencing based technologies and their potential application in the study of smallholder goat populations. A discussion on challenges and opportunities around genomics in smallholder livestock populations is given.

2.2 South African goat populations

There is a wide range of goat breeds developed for different production systems such as meat, milk and mohair (Nomura *et al*. 2013). South Africa is one of the few countries with well-defined meat type goats (Visser *et al*. 2004). The South African Boer, Savannah and Kalahari Red are such commercial meat breeds (Visser *et al*. 2004). These commercial meat breeds however, cannot be sustainably maintained in low input disease infested communal farming systems (Masika *et al.* 2007). The feral Tankwa goat is another South African goat breed kept as a conserved flock in the Northern Cape Province of South Africa (Kotze *et al*. 2014). In addition to these are a number of uncharacterized village goat populations found in the communal and smallholder sectors (Gwaze 2009; Casey & Webb, 2010). The South African village goats comprise non-descript indigenous ecotype populations that are adapted to low-input communal farming.

2.2.1 The SA Boer goat

The SA Boer goat is a commercial goat that was developed in South Africa in the early 1900s and was dubbed the "Boer" which means farmer in Dutch (http://www.boermeatgoats.com/Profit_Calculator.cfm). It was mainly developed for optimal production under the harsh African conditions (http://www.adga.org/breedinfo.html). There is limited information on breeds or populations that contributed to the Boer goat other than that it is a genetic combination of the African Indigenous goats, Indian goats, Angora goats and is influenced by some European dairy goats (http://studbook.co.za/boergoat/stand.html). Currently there are six types of Boer goats namely the ordinary Boer, long-hair, polled, white red-headed, bridle/briekwa, the mouse eared and short eared Boer goats (King 2009). The South African Boer goats are now widely distributed all over the world in counties such as New Zealand, Australia (http://www.boermeatgoats.com/Profit_Calculator.cfm) and USA (http://www.adga.org/breedinfo.html). The common commercial Boer goat is characterized by horns, short white haired body with red-brown colour on the head, a

solid white stripe on the forehead and long droopy ears (Griffith *et al*. 2013). It is one of the best goat breeds for meat production with an excellent body conformation, fast growth rate, good carcass quality, carcass yield, muscle to bone ratio and dressing percentage (55-60%) and early maturity (Lu 2013). Early maturity allows for the goat to be sold or mated at an earlier age thereby increasing economic gain (Mckenzie-Jakes 2007). The Boer goat is also very fertile (approx. 200% i.e produces kids twice a year) and reaches maturity between 4-6 months of age. At this age a buck weighs 110-135kg and a doe about 90-100 kg (Pieters 2007). They have an average daily gain of 0.4kg per day under good feeding conditions (Mckenzie-Jakes 2007). It is also important to note that this is a high performance breed requiring high and good quality maintenance such as good pen structure and high nutrient feed, but this alone makes them too expensive to keep (NAMC 2005).

2.2.2 The Kalahari Red goat

The Kalahari Red was selected from lop-eared animals for its uniform red coat colour by the South African and Namibian breeders and is hence characterized by long ears, a solid red coat colour together with an ability to strive in very hot conditions (Pieters 2007; Kotze *et al*. 2004). Initially identified as a landrace goat in 1998, the Kalahari Red is reported to comprise a mixture of the Savanna and Boer goats and was named from the red dunes of the Kalahari (Farmer's weekly 2010). It was genetically characterized by Kotze *et* a*l*. (2004). The Kalahari Red was initially restricted to the Eastern, Northern and Western Cape provinces and Namibia but has since spread all over South Africa and into other countries such as America (Farmer's weekly 2010). Management of this breed is as costly as the Boer as it also requires a high level of feed supplementation and minerals with accessible copper (Little 2010).

2.2.3 The Savannah goat

Messrs Cilliers and sons developed the Savannah goat from a mixture of indigenous white ewes and a white ram in 1957 (Visser *et al*. 2004). It is characterized by its solid white colour, long slightly curved face and long oval shaped ears said to be of Southern African origin (Pieters 2007), exact location of origin has been reported. The actual growth performance of the Savanna has not been reported. It is well known for its high

fertility and fast growth rate (http://www.extension.org/pages/19384/goat-breeds-savanna).

2.2.4 The S.A. Village goat

Village goats are kept by communal farmers in rural areas (Lahiff & Cousin 2005). The primary purpose of village goats is for meat production (Moloko 2011), they are also used for traditional ceremonies (Moloko 2011) and their manure is used for fertilization of soil in communal areas (Gwaze 2009). The village goats are characterized by loped ears and demonstrate a variety of coat colours, hair and ear lengths and body sizes (Morrison 2007), all of which are not well characterized (Visser *et al*. 2004). In addition the SA village goats are reported to demonstrate disease tolerance and an ability to utilize poor quality feed (Gwaze 2009). The mature live body weight of a village goat is approximately 38kg irrespective of sex (Lusweti 2000). They are believed to have come to South Africa as the black communities where migrating to different parts of Southern Africa (Campbell 2003). Their actual center of origin and evolution thereafter is not well known. Village goats are mostly owned by smallholder farmers in poor communities and are named based on their geographic location, for instance Nguni goats in Kwa-Zulu Natal and Venda goats in the Venda villages of the Limpopo Province of South Africa (NAMC 2005). It is, however unclear whether these SA village ecotypes represent genetically distinct populations.

Regardless of progress made in developing commercial breeds, the majority of SA's village goat populations remain uncharacterized with very little information available on their origin, production characteristics and genetic attributes. Village goats are mostly kept under smallholder farming systems (Moloko 2011), which are characterized by sub-optimal management, poor productivity, low weaning rate and lack of information and resources concerning livestock production systems (Masika *et al*. 2007). Goat performance in communal farms is said to be lower than that of commercial farming systems (Masika *et al*. 2007). This has been mainly due to the sub-optimal management practices of village farming systems (Masika *et al*. 2007). Since size and body weight are important traits in goat meat production, the improved commercial breeds are preferred over village breeds as they grow fast and yield higher mature weights (Kosgey *et al*. 2008). Village goats usually fend for their own feed in crop fields or around the

household and no supplement feed is given (Webb & Mamabolo, 2004). High level of diseases and parasites combined with poor nutrition leads to poor performance of village goats (Gwaze 2009). Village goats that are able to survive and thrive under such production challenges are then preferred for farming.

Even though some goat breeds such as the SA Boer have been developed in South Africa, more studies still need to be done on the South African goat populations. The SA village goats are considered to have great genetic potential due to the evidently high phenotypic variation within populations (Simela & Merkel, 2008). Their actual genetic potential is, however, yet to be fully explored and utilized. Further studies are required in order to reveal this genetic potential such that it can be utilized to benefit South African goat farmers. Morrison (2007) attempted to characterize the South African village goats to gain insight into how these goats can be grouped and better understood. This however was based on phenotypic characterization and needs to be supplemented with genetic characterization. Kotze *et al*. (2014) revealed the elevated genetic diversity in the feral Tankwa population. Genetic information regarding the growth potential of South African village goats is also necessary.

A study by Masika *et al*. (2007) reported poor performance and high within population variation in South African village goat populations. Additional studies are, however, required in order to better understand the growth performance and future growth potential of these goats. This will provide insight into what the optimal selection programs for these goats are and how to improve them for better growth performance.

2.3 Mitochondrial DNA and maternal origins of goats

The mitochondrial DNA (mtDNA) is a genome that is maternally inherited from generation to generation with no recombination effects (Naderi *et al*. 2007). MtDNA plays a major role in several cellular functions, which includes energy production by a process known as oxidative phosphorylation (Garrine 2007). Approximately 90% of the mammalian cell energy is thought to be produced by these organelles (Biase *et al*. 2007). The goat mtDNA is a circular double helix molecule that is approximately

16.64kb long (Jeon *et al*. 2004). The mtDNA is often used for maternal parentage verification and for the investigation of breed history and origin of domestic animals (Liu *et al*. 2007). The mitochondrial DNA has been used to study the origin and domestication of goats (Luikart *et al*. 2001; Odahara *et al*. 2006; Nomura *et al*. 2013; Naderi *et al*. 2007; Rana *et al*. 2013).

Previous mtDNA studies have suggested that goats were domesticated by humans over 10 000 years ago (Liu *et al*. 2007). Nomura *et al*. (2007) reported six mitochondrial DNA haplogroups named A, B, C, D, F and G [using nearly complete mtDNA protein coding genes (ATP6, COX1, COX2, COX3, ND1, ND2, ND4L, ND4, ND5, cytochrome b)] of populations from Japan, Korea, Mongolia, Indonesia, Philippines and Bangladesh. The same study demonstrated that haplogroup A was the most frequent and prevalent in the Asian goat populations. A study that was conducted in 13 Chinese goat breeds, demonstrated four maternal lineages in all the populations with lineage A, being common across all breeds (Liu *et al*. 2007). Liu *et al.* (2007) demonstrated Chinese goats to all be related and from similar maternal origins.

The Tibetan breed of goats demonstrated only two distinct dam lines/maternal lines, much like that of the Indian goats (Liu *et al*. 2007). The presence of these two lineages suggested that the Tibetan goats might have originated from India as Tibet is close to India (Liu *et al*. 2007). In Burkina Faso, mtDNA has been used to characterize domestic goats and ascertain the genetic diversity of the West African goat with no inferences made on goat domestication (Royo *et al*. 2008). Although mtDNA has been used extensively to study the genetic structure of European and Asian goats, very little information regarding mtDNA diversity and maternal lineages of African goats is available (Royo *et al*. 2008). Even the origins and genetic structure of breeds such as the South African Boer, Kalahari Red and Savanna is scarce, although these breeds have been well-studied and characterized for production performance (http://studbook.co.za/boergoat/stand.html). In South Africa, mtDNA has primarily been used as a tool for forensic research especially in the management of livestock theft and its usefulness has been demonstrated even in old samples (Zwane *et al*. 2013).

2.4 The different genes of the mtDNA and their use in genetic studies

The mitochondrial DNA circular genome is an important tool used in genomic studies of goats (Wu *et al*. 2012) and other species. It contains 22 tRNAs, 2 rRNAs, encodes 13 polypeptides together with genes such as the D-loop, cytochrome oxidase sub-unit 1-3, ND 1-4, ATP6 and 8 (Wu *et al*. 2012). Some of these genes have been used in various studies to better understand maternal origins, diversity and domestication of goats (*Capra hircus)* as shown in Table 2.1.

Though most studies suggest that the D-loop is the most effective tool to use in maternal lineage studies, the whole mtDNA genome has been reported to be the most reliable tool that provides a better phylogeographic and population structure and brings about a better understanding on the evolution and maternal origins of goats (Wu *et al*. 2012).

2.5 Goat domestication and maternal lineages

Goats were domesticated more than ten thousand years ago primarily because of their usefulness to humans in providing easily accessible meat and milk (Nomura *et al*. 2013). Studies suggest that goats were domesticated in the Fertile Crescent from the wild bezoar (*Capra aegagrus*) as the human beings developed hunting skills and the ability to manipulate wild animals to adapt to their civil way of living (Naderi *et al*. 2007). Archaeological studies coupled with some molecular studies have shown evidence that the domestic goat (*Capra hircus*) is of West Asian origins, however their domestication process and selection pressures remain unknown (Nomura *et al.* 2013).

Table 2.1 Mitochondrial DNA genes and studies where they have been used as tools for ascertaining domestication, maternal lineages and origin of goats.

Gene	**Populations**	**Haplo types**	**Maternal lineages/Haplogroup**	**Reference**
D-loop	Chinese	135	A, B, C, D	Liu *et al* 2007
HV1	West African	83	A, C	Royo *et al*. 2008
Protein encoding genes	Japan, Korea, Mongolia, Indonesia, Philippines and Bangladesh	-	A, B, B1, C, D, F, G	Nomura *et al*. 2013
D-loop, cytochrome *b*	Chinese	327	A, B, C, D	Liu *et al*. 2008
HV1	42 countries	1540	A, B1, B2, C, D, G	Naderi *et al*. 2007
D-loop	Southern Sardinian	419	A, C	Pirras *et al*. 2012
D-loop	Indian coastal	68	A	Rana *et al*. 2013
HV1	Korean Natives	6	A	Odahara *et* al. 2006
D-loop	Chinese	10	A, B, C, D, B1, B2	Han *et al*. 2010
D-loop and cytochrome *b*	Laos Natives	19	-	Mannen *et al*. 2007
D-loop	Chinese	326	A, B, C, D	Zhao *et al*. 2014

D-loop and Cytochrome *b*	Pakistani	38	A, B, C, D	Sultana *et al*. 2003
Complete genome	Sardinian	-	A	Doro *et al* 2014
Complete genome	Chinese	-	A, B, C, D	Wu *et al*. 2012

*L=Lineage, *HG=Haplogroup

Naderi *et al*. (2007) has reported the Fertile Crescent as the place of origin for the domestic goat based on mitochondrial DNA studies. Divergence time studies however indicate goat genetic diversity to have existed before goat domestication (Naderi *et al*. 2007; Nomura *et al*. 2013) although no clear understanding of lineages and possible domestication centers were established. To further understand the mystery behind goat domestication, Naderi *at al*. (2008) did a comparison between the genetic diversity of the domestic goat and that of the wild bezoar to clarify the domestication process. Figure 1 below shows two maps adopted from Naderi *et al*. (2008) explaining the possible domestication centers as well as the distribution of haplogroups (A, B, C, D, F and G).

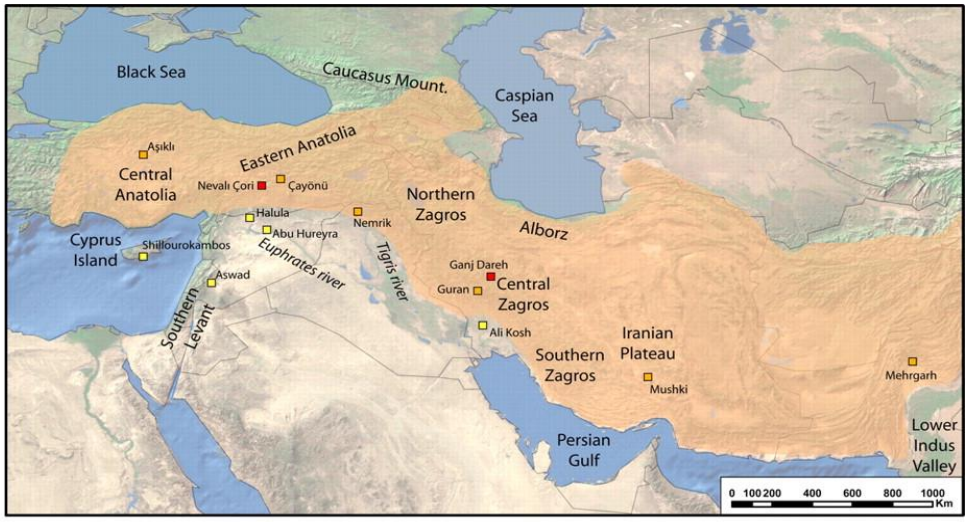
Although Naderi *et al*. (2008) proposed the possible domestication centers and haplogroup diversity as illustrated in Figure 1, the domestication and migration routes of South African goats from the Fertile Crescent to Africa, further down to South Africa is unclear. Campbell (2003) described the migration of the nomadic black and coloured (Khoikhoi) people who possessed domestic animals including goats from North Africa to the Southern parts of Africa through the narrow tsetse-free paths during the 5th century. The nomadic human populations have been reported to be of Chinese origins who migrated from their region in China (Wang *et al*. 2008) to North Africa and then to the Southern African parts of the continent as a result of poverty (Campbell 2003).

2.6 Goat growth genes

Growth is a quantitative trait that is affected by many genes at many loci. Goats have about 271 genes that can be categorized according to economic traits such as milk, fiber and meat production, disease resistance, reproduction and growth to which they are associated (Supakorn 2009). Seven of these genes are associated with growth and are presented in Table 2.2 (Supakorn 2009).

Some of these genes show potential for use in optimizing animal growth for production purposes. For example, the leptin and caprine myostatin genes can be used in meat production improvement. The leptin gene plays a role in carcass traits, with

A


Black Sea
Caucasus Mount.
Caspian Sea
Eastern Anatolia
Central Anatolia
Aşıklı
Nevalı Çori
Çayönü
Halula
Nemrik
Abu Hureyra
Northern Zagros
Alborz
Cyprus Island
Shillourokambos
Euphrates river
Tigris river
Aswad
Ganj Dareh
Guran
Central Zagros
Southern Levant
Ali Kosh
Iranian Plateau
Southern Zagros
Mushki
Mehrgarh
Persian Gulf
Lower Indus Valley
0 100 200 400 600 800 1000 Km

B

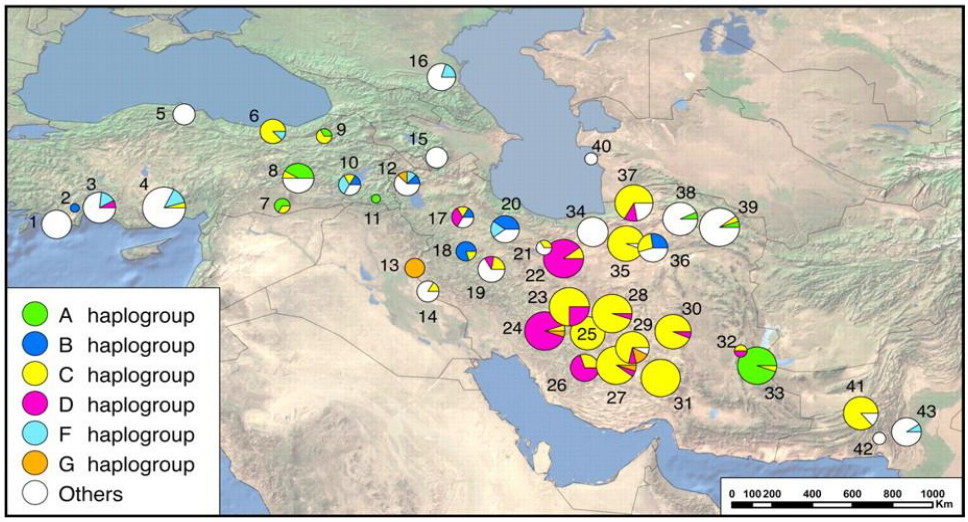

A haplogroup
B haplogroup
C haplogroup
D haplogroup
F haplogroup
G haplogroup
Others
0 100 200 400 600 800 1000 Km

Figure 2.1 Study area and geographic distribution of the mtDNA haplogroups in the bezoar. (A) Natural distribution of the bezoar according to Uerpmann (1987). This distribution may not have changed since the beginning of goat management/domestication, and stops at the eastern limit of the map. The archaeological sites that give evidence of local pre-Neolithic goat domestication are represented in red. The sites that suggest either local goat domestication or early pre-pottery Neolithic transfer of domesticated goat are represented in orange. Finally, the sites that provide evidence of transfer of domestic goats out of the original geographic range of the bezoar before the middle of the 10th millennium cal. B.P. are represented in yellow. The northern Zagros comprises the Iranian Provinces of Azerbaijan Gharbi, Zanjan and Kurdistan; the Central Zagros comprises Kermanshah, Lorestan, Khuzestan, and Isfahan Provinces. The Southern Zagros mainly comprises the Fars Province. (B) Geographic distribution of the mtDNA haplogroups in the bezoar. The size of the circles is proportional to the number of individuals analyzed (Naderi et al. 2008).

studies suggesting the use of this gene as a marker for weight improvement (Javanmard *et al*. 2009). The caprine myostatin on the other hand, has been reported for its potential use in the improvement of meat to develop desirable traits such as meat tenderness (Jain *et al*. 2012). The growth hormone gene is also one such important gene playing a big role in the growth of goats.

Growth hormone gene is a 2.544kb gene and consists of 5 exons and 4 introns (Alakilli *et al*. 2012). It is responsible for the production of the growth hormone (Supakorn 2009) and is essential for metabolism, post-natal growth, and milk traits in goats, sheep and other mammals (Alakilli *et al*. 2012). This gene has been observed to have an effect growth in the muscle, bone and adipose tissues of animals (Hua *et al*. 2009). Genetic association studies have observed that polymorphisms in exon 2 and exon 4 of this gene were associated with milk yield in the two ecotypes of Serrana goats (Mousavizadeh *et al*. 2009).

The study by Mousavizadeh *et al*. (2009) demonstrated that animals with the genotype A/B and FF on exon 4 produced more milk than others. It was also reported

that a high level of polymorphism in exon 4 of the growth hormone gene has the potential to be used in MAS for milk producing goats (Mousavizadeh *et al*. 2009). Hua *et al*. (2009) identified 10 polymorphisms in the gene, of which some genotypes were associated with milk production traits and others with heavier body weight with a mass difference of 2kg. The polymorphisms observed in this gene suggest the gene can be used as a candidate gene in selection programs (Marques *et al*. 2003; Gupta *et al*. 2009; Mousavizadeh *et al*. 2009).

2.7 Goat genomic tools in the next generation sequencing era

There are several tools that can be used in goat genomics studies. One of such tools is the use of Single Nucleotide Polymorphisms (SNPs) that have the ability to explain most of the genetic discrepancies evident between animals (Salem *et al*. 2012) and is associated with protein functions (Van Marle-Köster & Nel, 2003). The next generation era offers a range of tools that can generate high quality data in a short period of time (www.illumina.com).

Table 2.2 Growth hormone genes and their role in growth performance and other growth traits.

Gene	Populations	Key findings	References
Growth hormone	Chinese Boer	Association with growth traits	Hua *et al*. 2009
Growth hormone receptor	Alpine, Polish Fawn Improved, Saanen, Polish White Improved and Boer	Associated with some metabolic processes in the liver and may have an effect on the growth of goats.	Strzelec and Niżnikowsi, 2011
Insulin like growth factor I	Egyptian and Saudi	Major role in protein synthesis, growth of mammals, skeleton and hair growth. Plays a role in reproduction, ageing and growth development	Alakilli *et al*. 2012
Leptin	Boer	Positive regulator of the carcass, growth and production traits	Javanmard *et al*. 2009
Caprine pituitary specific transcription factor-1	Chinese indigenous	Polymorphisms of the coding gene of POU1F1 have a positive influence on animal weight, milk performance and reproduction	Lan *et al*. (2007)
Caprine Myostatin	*Capra hircus*	The use of gene knock out of the caprine myostatin gene to improve goat meat and develop desirable traits such as high protein content, soft and tenderness and less fat	Jain *et al*. 2012
Bone morphogenetic protein	Egyptian and Saudi	i) Influences goat growth ii) Influence on chest circumference; body height and trunk index whereas the sub-division	Alakilli *et al*. 2012, Suparkon 2009

This data can then be used to study informative genes that can assist in the generation of useful information and hence allow for a better understanding of the South African goats and more specifically the SA village goats. Their potential can hence be ascertained and developed for their improvement.

2.7.1. Single Nucleotide Polymorphisms

Single Nucleotide Polymorphisms (SNPs) are single base pair modifications within the genome (Dough 2010), comprising of non-synonymous and synonymous (Sharma *et al*. 2012) as well as coding and non-coding variations. Coding SNPs are found near the coding regions of various genes such as the mitochondrial DNA (Salem *et al*. 2012). A number of techniques have been developed to study SNPs, which include SNP arrays, whole genome sequencing (WGS) and targeted gene sequencing (Sharma *et al*. 2012).

2.7.2 The Illumina goat SNP50K

The Illumina goat SNP50K chip is a high density SNP chip consisting of markers evenly spaced across the goat genome, that has been found to be a very useful tool for population genetics, genetic association studies and the study of genetic disorders in goats (Tosser-Klopp 2014; Lashmar *et al*. 2015; Mdladla *et al*. 2016). It was developed using six goat breeds and has been validated to be suitable for use in a variety of goat breeds (Tosser-Klopp 2014). This technology has been used in various studies in South African goats. One such study was performed on Angora goats where the suitability of the SNP50 bead chip for specific breeds was explored (Lashmar *et al*. 2015). Lashmar and co workers (2015) also suggest that the chip can be used as a tool for genome-wide association, genetic variation, signatures of selection and genomic selection studies as well as for parentage verification. The 50K SNP chip has also been used in the analysis of population structure of South African goat populations (Mdladla *et al*. 2016). Mdladla *et al*. (2016) reported that, the feral Tankwa population is a genetically distinct population as compared to other South African goat populations. This study by Mdladla and co-workers (2016) further showed the usefulness of the SNP50K in the investigation of population structure, genetic diversity and relationshionships between the SA village goats and the feral

Tankwa populations. The SNP50K chip also has the ability to reveal valuable information on the demographic history of South African goat populations (Mdladla *et al*. 2016).

2.7.3 Whole genome sequencing

Whole genome sequencing (WGS) is a method whereby the whole genetic makeup of an organism or animal is sequenced (www.illumina.com). Whole genome sequencing in practice does, however, not necessarily cover all components of the genome (www.acmg.net). To ensure sufficiency in data generated, a large number of sequence reads is needed (approx. 6.3 million) and multiplexing is also a challenge (Boon & Faas 2013). These alone make sequence cost per sample very high, thereby placing limitations on sample numbers etc. (Boon & Faas 2013). WGS studies done on the South African feral Tankwa goat has revealed variants associated with cellular pathways, defense and immunity. SNPs that show a potential role in goat performance, diseases and adaptation processes have also been identified (Mohlatlole *et al*. 2015). The present study will generate additional information that can be used for the development and implementation of suitable breeding programs in South African goats, particularly for the uncharacterized and undeveloped Tankwa and SA village goats.

2.7.4 Targeted gene capture

Targeted gene capture is a technology that is simplified for NGS (Myllykangas *et al*. 2011). In this method, instead of sequencing a whole genome, a specific region in the genome is enriched and targeted by circularization and only that region is sequenced (Myllykangas *et al*. 2011). From this method one can generate data that is relevant to the specific region, allowing focus on that gene and thus more accurate and targeted analysis. A small gene fragment and fewer reads are required, hence making the cost per sample lower than that of whole genome sequencing (Boon & Faas, 2013). Targeted gene capture has been widely used in human studies (Myllykangas *et al*. 2011). One such study was for the identification of C90r75 as the mutated gene in non-syndromic deafness (Rehman *et al*. 2010). The first *Anaplasma phagocytophilum*, an obligate zoonotic bacterium transmitted by ticks, was isolated

from cattle in France using targeted gene capture (Dugat *et al*. 2014). Studies in livestock are, however, less common. No work has been reported on the use of targeted gene capture in goats.

2.8 Conclusion and recommendations

Very little is known about the genetic background and relatedness of SA goats. Studies have been conducted on goat growth genes and their associations with growth traits however, no such studies have been performed in South African goat breeds. SA goat breeds have been demonstrated to be genetically diverse but no report about growth genes, their genetic variation or the maternal lineage of these animals has been published. A number of tools are now available to characterize the genome, targeted regions of the genome and the mitochondrial genome of South African goat breeds. There is a need for such studies to be performed, as these will enable the optimization and development of breeding programs that may be relevant to communal farmers in South Africa. The characterization of these breeds will also play a vital role in their conservation.

2.9 References

Alakilli S. Y. M, Mahrous K. F, Salem L. M. & Ahmed E. S. (2012) Genetic polymorphism of five genes associated with growth traits in goat. *African Journal of Biotechnology* 11, 14738-14748.

American Boer Goat Association http://www.adga.org/breedinfo.html, accessed 20 January 2014.

Boer goat profit guide. (2010) http://www.boermeatgoats.com/Profit_Calculator.cfm. Accessed 20 January 2014.

Boon E. M. J. & Faas B. H. W. (2013) Benefits and limitations of whole genome versus targeted approaches for noninvasive prenatal testing for fetal aneuploidies. *Prenatal Diagnosis* 33, 563-568.

Campbell Q. P. (2003) The origin and description of Southern Africa's indigenous goats. *South African Animal Science* 4, 18-22.

Casey N. H. & Webb E. C. (2010) Managing goat production for meat quality. *Small Ruminant Research*. DOI: 10.1016/J.SMALLRUMRES.2009.12.047.

Dough B. (2010) Single nucleotide polymorphisms (SNPs). Computational Molecular Biology Biochemistry 218. *BioMedical Informatics* 231. (http://biochem218.stanford.edu/) accessed 20 January 2015.

Dugat T., Loux V., Marthey S., Moroldo M., Lagrée A., Boulouis H., Haddad N. & Maillard R. (2014) Comparative genomics of first available bovine Anaplasma phagocytophilum genome obtained with targeted sequence capture. *BMC Genomics* 15, 1471-2164.

Garrine C.M.L.P. (2007) Genetic characterization of indigenous goat populations of Mozambique. Department of Production Animal Studies in Faculty of Veterinary Science, University of Pretoria, South Africa.

Gwaze F.R. (2009) Communal production systems of goats raised by resource-poor farmers in the Eastern Cape Province of South Africa. Department of Livestock and Pasture Science, University of Fort Hare, South Africa.

Gupta N., Pandey A., Malik G. & Gupta S.C. (2009) Single nucleotide polymorphism (SNP) in growth hormone gene of Jakhrana, a prominent milk goat breed in India. *Small Ruminant Research* 81, 35–41.

Hua G.H., Chen S.L., Yu J.N., Cai K.L., Wu C.J., Li Q.L., Zhang C.Y., Liang A.X., Han L., Geng L.Y., Shen Z., Xu D.Q. & Yang L.G. (2009) Polymorphism of the growth hormone gene and its association with growth traits in Boer goat bucks. *Meat Science* 81, 391-395.

Jain H., Singh S. & Jain S. K. (2012) Designing and Development of Anti-Myostatin Short Hairpin RNA (Shrna) Constructs for *Capra hircus. Advances in Biological Research* 6, 37-43.

Javanmard A., Panandam J. M., Sugnaseelan S. & Yusoff K. (2009) Allele frequencies at six candidate genes associated with growth and carcass quality traits in the Boer goats. *African Journal of Biotechnology* 9, 7236-7238.

Jeon G. J., Chung H. Y., Choi J. G., Lee M. S., Lee C. W., Park J. J., Ha J. M., Lee H. K. & Sung H. H. (2005) Relationship between genetic variants of mitochondrial DNA and growth traits in Hanwoo cattle. *Journal of Animal Science* 18, 301-307.

King F. J. M. (2009) Production Parameters For Boer Goats In South Africa. Department of Animal, Wildlife and Grassland Sciences, University of the Free State, South Africa.

Kosgey I.S., Rowlands G.J., van Arendonk J.A.M. & Baker R.L. (2009) Small ruminant production in smallholder and pastoral/ extensive farming systems in Kenya. *Small Ruminant Research* 77, 11-24.

Kotzé A., Grobler J.P., van Marle-Köster E., Jonker T. & Dalton D.L. (2014) The Tankwa Karoo National Park feral goat population: A unique genetic resource. *South African Journal of Animal Science* 44, 43-48.

Lahiff E. & Cousins B (2005) Smallholder agriculture and land reform in south africa. *IDS Bulletin* 36, 127-131.

Lan X. Y., Pan C. Y., Chen H., Lei C. Z., Hua L. S., Yang X. B. & Qiu G. Y. (2007). Ddel Polymorphism in coding region of goat POU1F gene and its association with production traits. *Asian Australian Journal of Animal Science* 20, 1342 – 1348.

Lashmar S.F., Visser C. & E. van Marle-Köster (2015) Validation of the 50k Illumina goat SNP chip in the South African Angora goat. *South African Journal of Animal Science* 45, 56-59.

Little K. (2010) Meat goats: Sustainable livestock farming. Nuffield Farming Scholarships Trust.

Liu R. Y, Lei C. Z, Liu S. H. & Yang G. S. (2007) Genetic diversity and origin of Chinese Domestic goats revealed by complete mtDNA D-loop sequence variation. Asian-Aust. J. Anim. Sci. Vol. 20, No. 2: 178-183.

Luikart G., Gielly L., Excoffier L., Vigne J., Bouvet J. & Taberlet P. (2001) Multiple maternal origins and weak phylogeographic structure in domestic goats. *PNAS* 98, 5927-5932.

Lusweti E. C. (2000) A survey of goat production in the developing areas of the North West province of South Africa. South African Journal of Animal Science 30, 34-35.

Marques, P.X., Pereira, M., Marques, M.R., Santos, I.C., Belo, C.C., Renaville, R. and Cravador, A. (2003) Association of milk traits with SSCP polymorphisms at the growth hormone gene in the Serrana goat. *Small Ruminant Research* 50, 177-185.

Masika PJ, Mafu JV, Goqwana MW, Mbutiand C. & Raats J. (2007) A comparison of goat growth performance in a communal and commercial farming system in the Central Eastern Cape Province, South Africa.

Mckenzie-Jakes A (2007) Selecting and Evaluating Goats for Meat Production. Getting started in the meat goat business. Bulletin I, Vol. VII.

Mdladla, K., Dzomba, E., Huson, H., Muchadeyi F.C., 2015. Population structure and admixture of South African goat breeds investigated using genome-wide SNP data. Animal Genetics. doi: 10.1111/age.12442.

Mohlatlole R.P., Muchadeyi F.C. & Dzomba E.F. (2015) Whole-genome sequence and genetic variant analysis of the Tankwa feral goat of South Africa. 48th South African Society for Animal Science Congress. Poster.

Moloko B. (2011) A profile of the South African goat market value chain. Department of agriculture, fishery and forestry.

Mousavizadeh A., Abadi M.M., Torabi A., Nassiry M.R., Ghiasi H. & Koshkoieh A. (2009) Genetic polymorphism at the growth hormone locus in Iranian Talli goats by

polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP). *Iranian Journal Of Biotechnology* 7, 51-53.

Morrison J. W. (2007) A Guide to the identification of the Natural Indigenous Goats of Southern Africa . *South African Journal of Animal Science* 4.

Myllykangas S, Natsoulis G, Bell J. M. & Ji H. P. (2011) Targeted sequencing library preparation by genomic DNA circularization. *BMC Biotechnology* 11, 122.

Naderi S., Rezaei H.R., Taberlet P., Zundel S., Rafat S.A. et al (2007) Large-Scale Mitochondrial DNA Analysis of the Domestic Goat Reveals Six Haplogroups with High Diversity. *PLoS ONE* 2(10): e1012.

National Agricultural Marketing Council, (2005) Report on the investigation into the potential for the South African goat industry. Draft 1.

Nomura K., Yonezawa T., Mano S., Kawakam S., Shedlock A. M., Hasegawa M., & Amano T. (2013) Domestication process of the goat revealed by an analysis of the nearly complete mitochondrial protein-encoding genes. *PLOS ONE*, 8, Issue 8.

Odahara S, Chung H. J, Choi S. H , Yu S. L, Sasazaki S, Mannen H, Park C. S, & Lee J. H. (2006) Mitochondrial DNA Diversity of Korean Native Goats. Asian Australian Journal of Animal Science 19, 482-485.

Pieters A (2007) Genetic characterization of commercial goat populations in South Africa. Department of Animal Breeding and Genetics, University of Pretoria, Pretoria, South Africa.

Rana S., Singh S., Dureja V., Joshi J., Banerjee P., Sharma U. & Viji R. K. (2013) Mitochondrial D-loop analysis reveals single maternal lineage for coastal region goats of India. *International Journal Of Biomedical and Life Sciences* . 4, 235-244.

Rehman A.U., Robert J.M., Inna A.B., Shahid Y.K., Erich T. B., Mohsin S., Zubair M.A., Saima R., Shaheen N.K., Sheikh R. & Thomas B.F. (2010) Targeted capture and next-generation sequencing identifies C9orf75, encoding taperin, as the mutated gene in nonsyndromic deafness DFNB79. *The American Journal of Human Genetics* 86, 378-388.

Royo L. J., Traore A., Tamboura´H. H., lvarez I. A´., Kabore´A., Ferna´ ndez I ., Oue´ draogo- Sanou G., Toguyeni A., Sawadogo L. and Goyache F. (2009) Analysis of mitochondrial DNA diversity in Burkina Faso populations confirms the maternal genetic homogeneity of the West African goat. *International Society for Animal Genetics, Animal Genetics* 40, 344-347.

Salem M., Vallejo R. L., Leeds T. D., Palti Y., Liu S., Sabbagh A., Rexroad III C. E. & Yao J. (2012) RNA-Seq Identifies SNP Markers for Growth Traits in Rainbow Trout. *PLoS ONE* 7, e36264.

Sharma U, Banerjee P, Joshi J. & Kumar R. (2012) Identification of SNPs in Goats (*Capra hircus*) using RNA-Seq Analysis. *International Journal of Animal and Veterinary Advances* 4, 272-283.

Simela L. & Merkel R. (2008) The contribution of chevon from Africa to global meat production. *Meat Science* 80, 101-109.

SUPAKORN China (2009) The Important Candidate Genes in Goats - A Review. *Walailak J Science & Technology* 6,17-36.

South African Boer Goat Breeders' Association http://studbook.co.za/boergoat/stand.html. Accessed 20 January 2013.

Tosser-Klopp, G., Bardou, P., Bouchez, O., Cabau, C., Crooijmans, R., Dong, Y., Donnadieu-Tonon, C., Eggen, A., Heuven, H.C.M., Jamli, S., Jiken, A.J., Klopp, C., Lawley, C.T., McEwan, J., Martin, P., Moreno, C.R., Mulsant, P., Nabihoudine, I., Pailhoux, E., Palhiere, E., Rupp, R., Sarry, J., Sayre, B.L., Tircazes, A., Wang, J., Wang, W., Zhang, W. & the International Goat Genome Consortium. (2014). Design and Characterization of a 52K SNP Chip for Goats. PLoS ONE 9, e86227

Van Marle-Köster, E.V. & Nel, L.H. (2003). Genetic Markers and their Application in Livestock Breeding in South Africa: A review. Pretoria, South Africa.

Visser C., Hefer C.A., van Marle-Köster E. & Kotze A. (2004) Genetic variation of three commercial and three indigenous goat populations in South Africa. *South African Journal of Animal Science* 34, 24-27.

Wang J., Chen Y.L., Wang X.L., and Yang Z.X. (2008) The genetic diversity of seven indigenous Chinese goat breeds. *Small Ruminant Research* 74, 231–237.

Webb E. C. & Mamabolo M. J. (2004) Production and reproduction characteristics of South African indigenous goats in communal farming systems. *South African Journal of Animal Science* 34, 236-239.

Wu Y., Huo J. & Gua W. (2012) Sequence analysis of four caprine mitochondrial

DNA lineages. *Italian Journal of Animal Science* 11, 378-383.

Zwane A.A., van Marle-Köster E., Greyling B.J., & Mapholi N. (2013) A review: Forensic DNA technology to meet the stock theft challenges in South Africa. *Applied Animal Husbandry and Rural Development* 6, 36-47.

CHAPTER 3

Investigation of genetic diversity and maternal lineages of SA goat populations using complete mitochondrial DNA Sequences

Abstract

South Africa is one of the leading goat producing countries and has some well-defined goat breeds. The origins and maternal lineages of these goats is however unknown and most populations still remain uncharacterized. Complete mitochondrial DNA was used to investigate genetic diversity and maternal lineages of 50 South African goat breeds in conjunction with mtDNA goat sequences from GenBank (n=13). Rolling Circle Amplification together with Illumina MiSeq next generation sequencing were used to generate the full length (16.64kb) of the mtDNA of South African commercially developed Boer (n=9), the captive feral Tankwa (n=9), and undeveloped village (n=33) goat populations. The reconstructed mtDNA consisted of two ribosomal RNAs (12S and the 16S), 22 transfer RNA genes, COX1-COX3, CYTB, ND1-ND6 ND4L, ATP6, ATP8 and the D-loop. The mtDNA was highly polymorphic in the South African goat populations. A total of 184 SNPs and 55 AA changes were observed. High within-population variation was observed in all the groups ranging from 98.60-99.52%. Low F_{ST} (F_{ST} = 0.003-0.049) indicated close relatedness between SA goat populations and suggested gene flow between populations. The 50 sequences yielded forty-two, nineteen and twenty-six haplotypes for the D-loop, COX1 and complete mtDNA respectively. Both the D-loop and the complete mtDNA haplotypes clustered into six clades whilst the 19 COX1 haplotypes were distributed amongst two clades. Some of the South African D-loop haplotypes clustered together with Chinese haplotypes belonging to lineages A and B. The presence of the two lineages in the South African goat populations suggest that they either share maternal lineages with the Chinese goat populations or they are of the same origin. The results further confirmed that the Bezoar (*Capra aegagrus)* goat is a possible ancestor of the domestic goats.

Key words: Complete mtDNA, SNPs, genetic diversity, maternal lineages, goat domestication

3.1. Introduction

Goats are important farm animals that provide easily accessible meat and milk (Nomura *et al*. 2013). Goats were domesticated over 10 000 years ago (Han *et al*. 2010) and have since then been farmed and distributed all over the world. South Africa is one of the few countries across the world with well-defined goat breeds. The four South African indigenous goat breeds are categorized into meat type goats, namely (i) SA Boer (ii) Savannah and (iii) Kalahari Red (Visser *et al*. 2004), and (iv) SA village goats (Morrison 2007). The village goats are mostly kept and raised by communal farmers in the marginalized areas. Goats are used commercially and for subsistence needs that include meat, milk and skins (Casey & Webb, 2010; Moloko 2011). Secondary uses mainly by commercial farmers are for cheese, cashmere and mohair (Moloko 2011). Communal farmers use goats as a method to alleviate poverty through sales (Simela & Merkel, 2008; Gwaze 2009) and also as an investment for financial security (Moloko 2011). Meat provision is a primary role of goats and this meat provides easily accessible and high quality proteins to poor village families (Gwaze 2009).

Goat faming in South Africa is predominated by village goats and over 60% of the goats are found in the communal areas (Moloko 2011). Regardless of their numbers, not much is known about the village goat populations. Although some populations have been phenotypically characterized (Morrison 2007), their genetic characteristic and potential is unknown. The origin of South African goats including the well-developed and characterized commercial breeds is unknown. Studies by Campbell (2003) suggested that the nomadic people brought goats into South Africa from North Africa. This, however, does not clarify their origin neither does it give information of their maternal lineages. Knowledge on goat origins or any other livestock species assist on understanding their genetic background and can be used as a tool for breed characterization, improvement and conservation. The lack of information on South African village goats hinders genetic improvement and conservation strategies.

The mitochondrial DNA has been used in molecular studies to study and identify the origins and movement of goat populations. Using the D-loop region of the mtDNA, Han *et al*. (2010) implied that the Chinese goats had multiple maternal origins

observed as lineages (A, B, and D) with conclusions that maternal lineage B1 and B2 were of Chinese origins. Whilst other studies report multiple maternal origins and are uncertain about goat origins, suggestions that the domestic goat may have been domesticated from the bezoar in the Fertile Crescent was investigated and confirmed using mtDNA (Zhao *et al*. 2014).

Various regions of the mtDNA have been used in phylo-geographic studies to investigate genetic diversity, domestication and migration of goat populations (Liu *et al*. 2007; Naderi *et al*. 2007; Nomura *et al*. 2013). The goat mtDNA is a circular molecule which is the approximately 16.64kb in size (Jeon *et al*. 2004) that have roles in several cellular functions such as energy production. About 90% of the mammalian cell energy is produced by these organelles through oxidative phosphorylation (Biase *et al*. 2007). The maternal inheritance of the mtDNA as a single haplotype with no recombination makes it an ideal marker to study domestication and maternal lineages as well as breed history (Garrine 2007). The D-loop has been the most studied fragment (Corte´s *et al*. 2008; Garrine 2007; Liu *et al*. 2007; Rana *et al*. 2013). Four D-loop based maternal lineages (A-D) were identified in Chinese goat populations (Liu *et al*. 2007). Rana and co-workers also used the D-loop to confirm single maternal lineages of goat populations from the Indian coastal region. Recent advances in next generation sequencing have allowed easy generation of complete mtDNA sequences in genetic diversity and livestock domestication studies (Wu *et al*. 2012). In a study on Chinese goats, complete mtDNA sequences revealed more information on molecular evolution as more SNPs (124) were identified, and other new data from the tRNA and rRNA genes that were previously not used in analysis was observed (Wu *et al*. 2012). The study by Wu *et al*. (2012) demonstrated the usefulness of complete mtDNA analysis as compared to that of a single mtDNA genome. The current study investigated the genetic diversity and maternal lineages and inferred on genetic origin of South African goat breeds using complete mtDNA sequences.

3.2. Materials and methods

3.2.1 Goat populations

Six ml of blood were collected from the jugular vein of 50 South African indigenous goats representing commercially developed Boer (n=9), feral/captive Tankwa (n=9) and unimproved village goat populations (n=32) from Eastern Cape (n =8), Kwa-Zulu Natal (n=8), Limpopo (n =8) and North West Provinces (n=8). The Boer goats were sampled from commercial farms in the North West province. One goat per household was sampled from 32 village households from the four provinces as indicated in Table 3.1 (Fig 3.1). The non-descript village populations were sampled from villages of the four major goat-producing provinces; (i) Hobeni village, Elliotdale municipality and Pechelsdam village, Inxubayethemba municipality in Eastern Cape, (ii) Coniliva and Ngubo villages in Msinga municipality Kwa-Zulu Natal, (iii) Mukovhabale village, Mutale municipality and Muila-muumone, Makhado municipality in Limpopo and (iv) Pella village, Moses Kotane municipality North West provinces of South Africa. The Tankwa goats were sampled from Carnavon Research Station. Blood samples were stored in an icebox and transported to the Agricultural Research Council, Biotechnology Platform laboratories on ice at the end of every sampling period at which they were stored at -20°C freezer till further use.

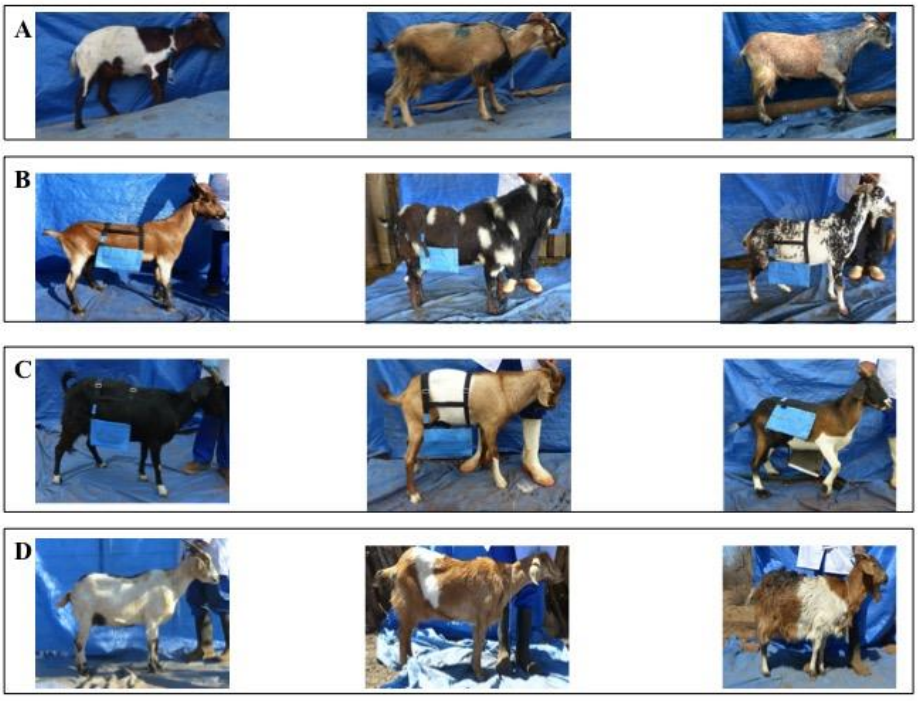


Figure 3.1 Spectrum of morphological variations of the 32 non-descript South African Village goats raised in communal farms. A: Eastern Cape village goats, B: Kwa-Zulu Natal village goats, C: Limpopo village goats and D: North West village goats.

3.2.2 DNA isolation and quantification

DNA was isolated using the optimized Qiagen DNeasy blood and tissue kit (www.qiagen.com) according to manufacturer's instruction with slight modifications. The optimization included pipetting higher volume of 200µl of goat blood and an increased incubation period of 120 minutes after adding 20µl proteinase K and 100µl PBS into the blood. Furthermore, 200µl lysis and 200µl of absolute ethanol (100%) were added to the mixture after incubation. The mixture was transferred into a spin column tube and centrifuged (Spectrafuge 24D Labnet. Labnet International, Inc.). Followed by two wash steps with wash buffers containing guanidine hydrochloride. The DNA was eluted with 100µl of buffer AE (elution buffer).

DNA quantification was performed on the Qubit® 2.0 Fluorometer using the Invitrogen's Qubit™ dsDNA BR Assay Kit (Invitrogen, Life Technologies). The quality of the DNA template was also investigated electrophoretically on 1% agarose gel with 4µl of ethidium bromide at 80V for 30min. Gels were examined under UV light in a BIORAD Imaging System (BIORAD GelDoc XR) (www.bio-rad.com).

3.2.3 Primer designing

The Intergraded DNA Technologies company designed and synthesized twenty-two goat primers (forward and reverse) (Appendix Table A1) on the basis of a known goat mtDNA sequence of Vietnamese origin (NC_005044) extracted from the GenBank (www.ncbi.nhl.gov) database.

3.2.4 Rolling circle amplification

A Qiagen REPLI-g Mitochondrial DNA Kit was used and the protocol was optimized for the amplification of complete mtDNA. Briefly, the RCA reaction mix contained 10µl of genomic DNA of variable concentrations ranging from 55.6-83.8 ng/µl, 7.5µl of nuclease free water and 29µl REPLI-g mt reaction buffer. The mixture was briefly vortexed and centrifuged. The samples were incubated for 5min at 75°C in a heating

block (Accublock ™ Digital Dry bath Labnet. Labnet International, Inc.), and cooled to room temperature (approx. 30min). After cooling, 1µl of REPLI-g Midi polymerase was added to the mixture followed by brief centrifugation.

Table 3.1 The distribution of goats sampled from both commercial and village goats from Eastern Cape, Kwa-Zulu Natal, Limpopo, North West and Northern Cape provinces of South Africa

Goat breed/ecotype	**Acronym**	**N**	**Production system**	**Location**
Boer goat	SAB	9	Intensive	Stella farm (n=2) and van Tonder farm (n=7) – North West
Tankwa	TWA	9	Captive feral	Carnavon Research Station – Northern Cape
Xhosa	EC	8	Extensive	Eastern Cape Province
Zulu	KZN	8	Extensive	Kwa-Zulu Natal Province
Venda	L	8	Extensive	Limpopo Province
Tswana	NW	8	Extensive	North West Province

A primer mix was made by adding 2µl of each primer (forward and reverse) to make a 44µl mixture, to which 56µl of nuclease free water was added to make up a 100µl primer mix. The primer mix (2.5µl) containing a set of 11 primer sets was added to

the mixture. Samples were incubated overnight (approx. 17hrs) at 33^{o}C in a heating block (Accublock ™ Digital Dry bath, Labnet International, Inc.). After incubation the polymerase was inactivated by heating the samples for 3min at 65^{o}C, and cooled to room temperature for 30min. The quality of the amplicons (2µl) was checked on a 1% agarose gel with 4µl of ethidium bromide at 80V for 30min. The gel was examined under UV light in a BIORAD Imaging System (BIORAD GelDoc XR) (www.bio-rad.com). Amplicons were then stored at -20^{o}C till further use. The RCA amplicons were quantified (ng/µl) using the Invitrogen's Qubit^{TM} dsDNA BR Assay Kit (Invitrogen, Life Technologies). Samples that were above 30ng/µl were submitted for sequencing at the Agricultural Research Council, Biotechnology Platform Core Facility.

3.2.5 MtDNA sequencing

Sample preparation was performed using the Nextera® v.2 Sample prep kit for the following a standard Illumina protocol. The prepared samples were run on the Illumina MiSeq Bench-top Sequencer (Illumina Inc., San Diego, CA, USA) using the MiSeq Reagent Kit v3 at a 2X 300bp read length configuration. Approximately 0.5GB of data consisting of 1.5 million reads was generated for each animal.

3.2.6 Data analysis

3.2.6.1 Quality control

Sequence reads were checked for quality based on quality score (QC >30) and number of ambiguous nucleotides (0) and trimmed to remove adaptor sequences using CLC Genomics Workbench v8.1.0 (CLC Bio, Aarhus, Denmark). Trimmed reads were mapped to a reference genome (Accession NC_005044) obtained from NCBI (http://www.ncbi.nlm.nih.gov.za) using the *"Map reads to a reference"* tool, at a similarity factor of 0.9. Consensus sequences were extracted for each animal sequenced and used for downstream analysis. Annotation of consensus sequences was done using the *"Annotation"* tool from CLC genomics workbench v8.1.0 (CLC Bio, Aarhus, Denmark).

3.2.6.2 MtDNA SNP analysis

Data was analyzed for single nucleotide polymorphisms (SNPs) per mtDNA gene using the *Probalistic Variant Detection* tool, which also detects mutations that result in Amino Acid (AA) changes. SNPs were called at a variant probability of 95% and minimum coverage of 10 and variants were screened on both the forward and reverse strands of the gene.

3.2.6.3 Genetic diversity and analysis of molecular variance (AMOVA)

Gene diversity per population for South African goat populations was estimated using Arlequin v3.5 (Schneider *et al.* 2000). Pairwise F_{ST} was used to measure genetic distances between pairs of South African goat populations using Arlequin v3.5 (Schneider *et al.* 2000). Analysis of Molecular Variance (AMOVA) for the estimation of population sub-structure was done for different population categories consisting of (i) SA village goats only; (ii) SA village and Boer goats; (iii) village and Tankwa; (iv) SA village, Boer and Tankwa goats and (v) all six sub populations.

3.2.6.4 Phylo-geographic structure of SA goat breeds

Median Joining Network version 4.6.1.1 (Bandelt *et al*. 1999) was used for the construction of Network trees from the mtDNA haplotypes of the 50 SA goats sequenced. Three median joining networks were constructed from (i) D-loop, (ii) COX1 and (iii) complete mtDNA genome sequences. Clades were defined, as a group of haplotypes clustering together was resolved and the distribution of goats from different breeds or population in each clade counted.

Sequences representing maternal lineages based on the D-loop from studies by Liu *et al*. (2007), COX1 by Nomura *et al*. (2013) and complete genome by (Hassannin *et al* 2006; Wu *et al* 2012; Niu *et al* 2014) were drawn from the GenBank (Appendix Table A2) database and together with the SA haplotypes used to draw a median joining network. The maternal lineages and probable origins of the SA goats were inferred

based on the clustering of SA haplotypes in relation to those from literature (Liu *et al*. 2007; Wu *et al*. 2012 and Nomura *et al.* 2013).

3.3. Results

3.3.1 Whole Mitochondrial DNA genes and Single Nucleotide Polymorphisms

A total of 1 246 777 to 2 857 386 reads with an overall average read length of 278,2 nucleotides (nt) after trimming were generated per animal. Mapping and annotation resulted in a consensus sequence of 16.64kb with the expected mtDNA genes (Figure 2), which was used for further analysis. The consensus mtDNA sequences consisted of two ribosomal RNAs (12S and the 16S), 22 transfer RNA genes, which were distributed along the mtDNA and not only in one region. It also contained the ND1-ND6 including ND4L, COX1-3 and CYTB ATP6 and ATP8, and the control region D-loop (Figure 3.2).

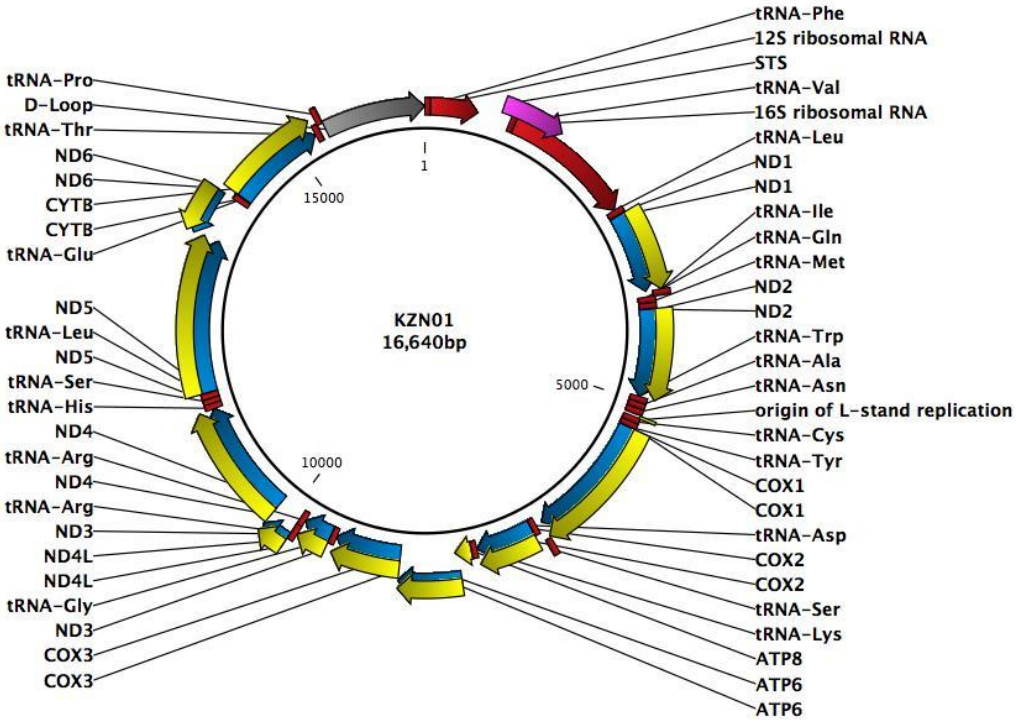


Figure 3.2 An annotated complete goat mtDNA from a South African goat (KZN01) constructed using CLC genomics workbench version 8.1.

3.3.2.1 Single Nucleotide Polymorphism (SNP) analysis analysis

A total of 184 SNPs were observed in the 50 complete mtDNA sequences of the South African goat breeds. The percentage of SNPs per mtDNA gene ranged from 0.010% in tRNA-Ala to 33% in the ND3 gene (Figure 3.3).

3.3.2.2 Amino Acid changes

Some of the mutations resulted in amino acid changes (AA) and a total of 55 AA were observed in the mtDNA ND1-ND6, COX1-COX3, CYTB, ATP6, and ATP8. There were no AAs observed on the 12S and 16S rRNA as well as the tRNA regions.

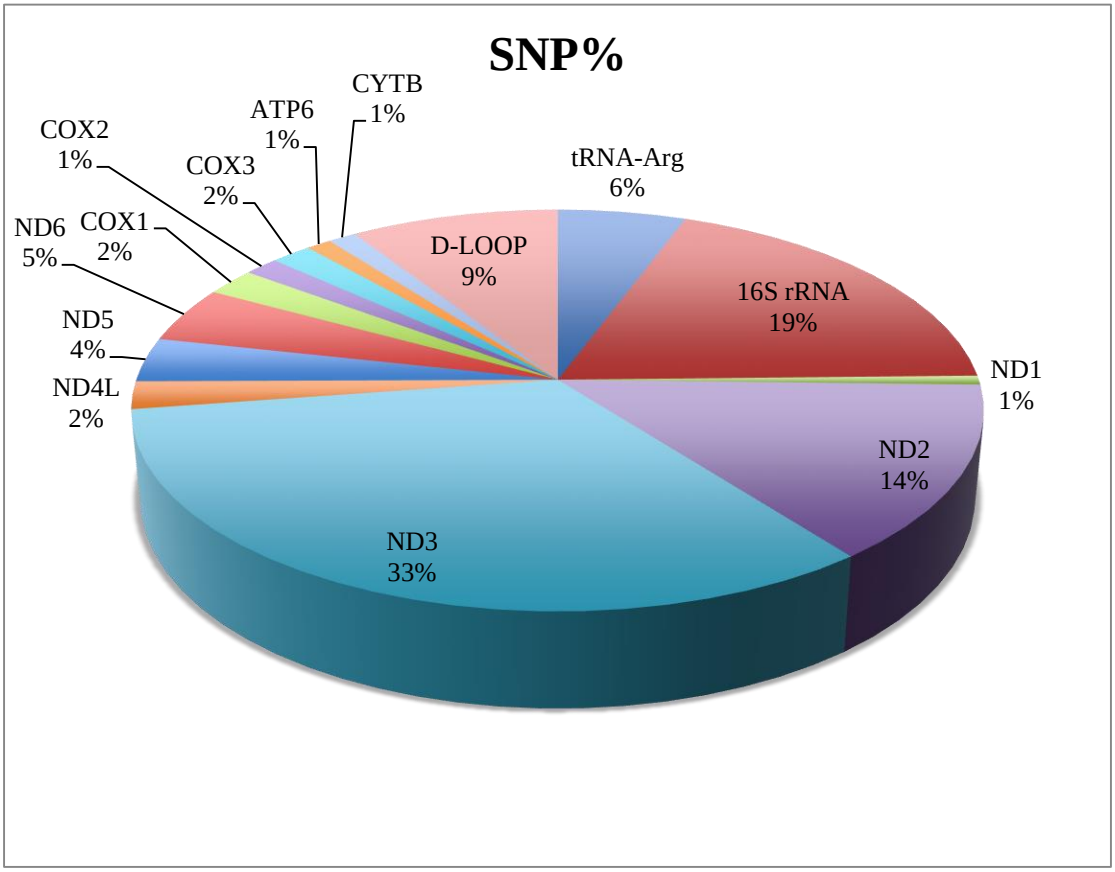


Figure 3.3 A pie chat representation of the SNP distribution across the different mtDNA genes of the South African goat populations showing SNP contribution in the mtDNA genome.

3.3.3 Pairwise F_{ST} genetic distances

Pairwise genetic distances (F_{ST}) between pairs of populations ranged from 0.003 to 0.036 between the SAB and village goat populations, and from 0.010 to 0.031 between pairs of village goat populations (Table 3.2). The NW village goat population demonstrated a very slight near zero genetic distance when compared with the SAB and the TWA with F_{ST} values of (F_{ST} =0.003 and F_{ST} =0.006) respectively.

Table 3.2 FST pairwise genetic distances of the South African goat population based on complete mtDNA sequences.

*Population	EC	KZN	L	NW	SAB
EC					
KZN	0.017				
L	-0.031	0.021			
NW	-0.017	-0.010	-0.036		
SAB	0.016	0.032	0.015	-0.003	
TWA	-0.005	0.049	0.037	0.034	-0.020

*EC = village goats from Eastern Cape province; KZN = village goats from KwaZulu-Natal; LIM = village goats from Limpopo province; NW = village goats from North West province; SAB = South African Boer goats; TWA = Tankwa feral goats.

3.3.4 Analysis of Molecular Variance and Population substructure

High within-population variation was observed across all population categories investigated (Table 3.3). A slightly elevated between-population variation was observed in the categories consisting of (i) village and Boer goats; (ii) village and Tankwa and (iii) village, Boer and Tankwa goats (Table 3.3).

Table 3.3 Analysis of molecular variance and population sub-structure within and between South African goat populations using complete mtDNA sequences.

	Percentage (%) of variation		
Populations compared	**Among groups**	**Among populations within groups**	**Within population**
SA Village only	0	0.92	99.08
SA Village goats, Tankwa and Boer	1.68	1.20	99.52
SA Village and Boer	2.90	1.90	98.20
SA Village and Tankwa	1.03	1.08	99.37
All six groups	0	1.4	98.60

3.3.5 Phylogeographic structure and maternal lineages of SA goat breeds

3.3.5.1 mtDNA D-loop analysis

A total of 42 D-loop haplotypes were observed out of the 50 individual South African goats sequenced and two of these haplotypes were shared between populations (Table 3.4). Haplotype EC02 was the most predominant at a frequency of 7% and was shared between four populations namely of Eastern Cape, KwaZulu-Natal, and North West. Haplotypes NW03 was the least frequent (2%). The median network joining for the SA haplotypes produced six clades as illustrated in Figure 3.4 and Table 3.4.

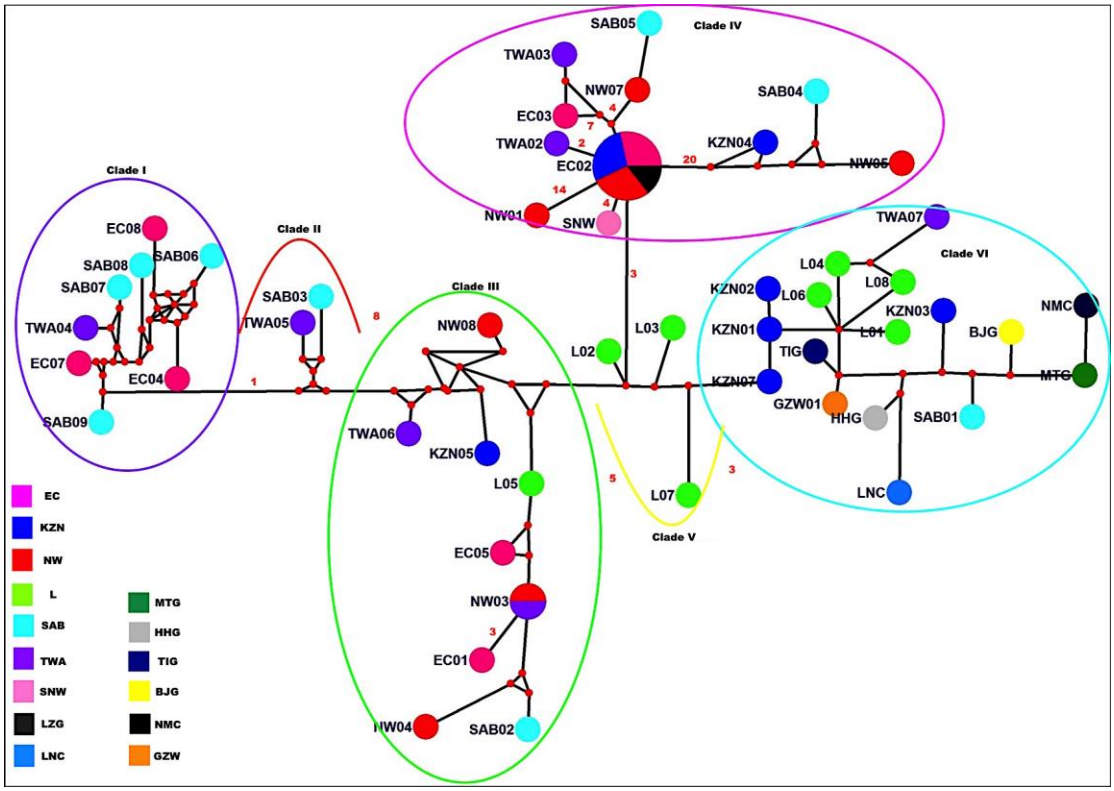


Figure 3.4 Median network of South African goat D-loop sequences as well as the Chinese lineages A-D from GenBank (Appendix table A2). Haplotypes are represented by a circle, and frequency of each haplotype by the size of the circle. Number of mutations between haplotypes and clades has been represented in red. The EC, KZN, L, NW, SAB and TWA are South African goat populations from Eastern Cape (EC), Kwa-Zulu Natal (KZN), Limpopo (L), North West (NW), South African Boer (SAB) and Tankwa feral (TWA) goats. The LZG and the SNW are the Chinese goat populations from maternal lineage A, and LNC, MTG, HHG, TIG, BJG, NMC and GZW are maternal lineage B populations (Liu et al. 2007).

Clade I consisted of 8 haplotypes from three populations of EC, SAB and TWA. The SAB was the most frequent population (n =4) in this clade (table 3.4). Clade II was the smallest clades with only two haplotypes from SAB and TWA. Clade III had 10 haplotypes and one of the haplotypes (haplotype NW03) was shared between the NW and the TWA goats whilst the other haplotypes were from EC, KZN, L, NW, SAB and the TWA. The most dominant population in this clade was NW (n =3). Clade IV had the highest number of haplotypes (n =15) and the NW population was the most dominant population. Haplotype EC02 was shared among populations and was observed in Clade IV; three of the South African goat populations shared this

haplotype and it was the most predominant haplotype. Clade V consisted of haplotypes from one population, of the village goats from Limpopo province.

Table 3.4 D-loop clade representations showing relationships between South African goat populations.

*Population	Clade I	Clade II	Clade III	Clade IV	Clade V	Clade VI	Total
EC	3	0	2	3	0	0	8
KZN	0	0	1	3	0	4	8
L	0	0	1	0	3	4	8
NW	0	0	3	5	0	0	8
SAB	4	1	1	2	0	1	9
TWA	1	1	2	2	0	1	9
Total	8	2	10	15	3	10	50

*EC = village goats from Eastern Cape province; KZN = village goats from KwaZulu-Natal; LIM = village goats from Limpopo province; NW = village goats from North West province; SAB = South African Boer goats; TWA = Tankwa feral goats.

Clade IV and VI of the South African goat populations clustered with some of the Chinese goat populations from LZG and the SNW populations representing lineages A and B respectively. Clade VI on the other had consisted of six populations from Chinese lineage B and four of the South African goat populations, L (n=4), KZN (n=3), SAB (n=1) and TWA (n=1). Although this clade consisted of many goats from the Chinese lineage B, no haplotypes were shared between these populations and the South African goats. Haplotype NW03 was shared between the South African goat populations only and had no link with GenBank sequences. Four of the six clades produced by the D-loop network were unique to the South African goats.

The two clades that were shared between the South African goats and Chinese goat populations from lineage A and lineage B were used to better describe maternal lineages of the South African goat populations. Clade IV was then defined as Lineage A as two of the Chinese goat populations (LZG and SNW) observed in this clade were from lineage A, while Clade VI consisted of six goats (LNC, MTG, HHG, BJG, NMC and GZW) from the Chinese Lineage B and was therefore defined as Lineage B clade.

3.3.5.2 MtDNA COX1 analysis

The 50 SA goats sequenced produced 19 COX1 haplotypes (Table 3.5). Haplotype EC01 was highly frequent (24%) and was observed in goats from EC, KZN, L, NW, SAB and TWA populations. Haplotype AB02 was the second most frequent (9%) as illustrated in Figure 3.5. This network produced only two clades as shown in table 3.5.

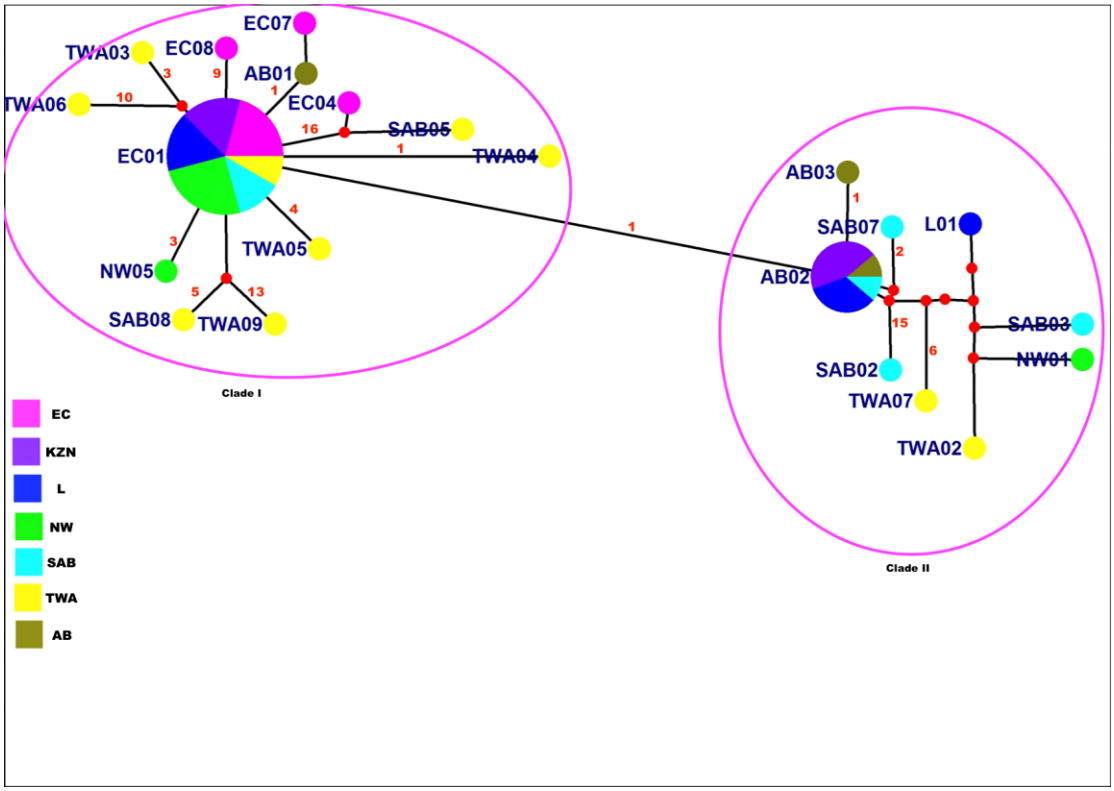


Figure 3.5 Median network joining of the cytochrome oxidase sub-unit 1 (COX1) region of the mtDNA for South African as well as sequences from the GenBank. Haplotypes are represented by a circle, and frequency of each haplotype by the size of the circle. Number of mutations between haplotypes and clades has been represented in red. The EC, KZN, L, NW, SAB and TWA are South African goat populations from Eastern Cape (EC), Kwa-Zulu Natal (KZN), Limpopo (L), North West (NW), South African Boer (SAB) and Tankwa feral (TWA) goats. AB01is the GenBank sequences from Bangladesh Black Bengal goats while AB02 is from the Wild Bezoar and AB03 is a domestic goat of Mongolian native).

Clade I was the most predominant with highest number of haplotypes (n=35) from all the populations under study. This clade also contained the highly frequent haplotype EC01, which was shared between 6 South African goat populations from EC, KZN, L, NW, SAB, and TWA populations. NW population was the most dominant population among the populations observed in haplotype EC01.

Table 3.5 Clade representation of the COX1 showing relationships between South African goat populations.

*Population	Clade I	Clade II	Total
EC	8	0	8
KZN	4	4	8
L	4	4	8
NW	7	1	8
SAB	5	4	9
TWA	7	2	9
Total	35	15	50

*EC = village goats from Eastern Cape province; KZN = village goats from KwaZulu-Natal; LIM = village goats from Limpopo province; NW = village goats from North West province; SAB = South African Boer goats; TWA = Tankwa feral goats.

The second clade contained only 15 haplotypes whereby haplotype AB02 had a frequency of 20% and was shared between three South African goat populations from (KZN, L and SAB). The EC populations did not share this haplotype and were neither part Clade II. Clade I clustered with AB01 sequence from the Bangladesh Black Bengal goat. Clade II on the other hand clustered with the wild bezoar and some Mongolian native goats. The haplotype EC01 was the most abuntant haplotype that was unique to South African goat.

3.3.5.3 Complete mtDNA sequence analysis

The 50 goat sequences produced 26 complete haplotypes (Accession: SAMN04978812, SAMN04998821, SAMN04994848, SAMN04994861, SAMN04994863, SAMN04994866, SAMN04994868-69, SAMN04994884-85,

SAMN04994913-14, SAMN04994916, SAMN04994918-19, SAMN04995027, SAMN04995075-76, SAMN04998558, SAMN04998774-79, SAMN0499889) (Figure 3.6; Table 3.6). Haplotype EC02 was the most predominant at a frequency of 21% and was observed in EC, KZN, L, NW, SAB and TWA populations. Haplotype KZN04 was the second most frequent (8%), observed in KZN, L, NW and SAB populations

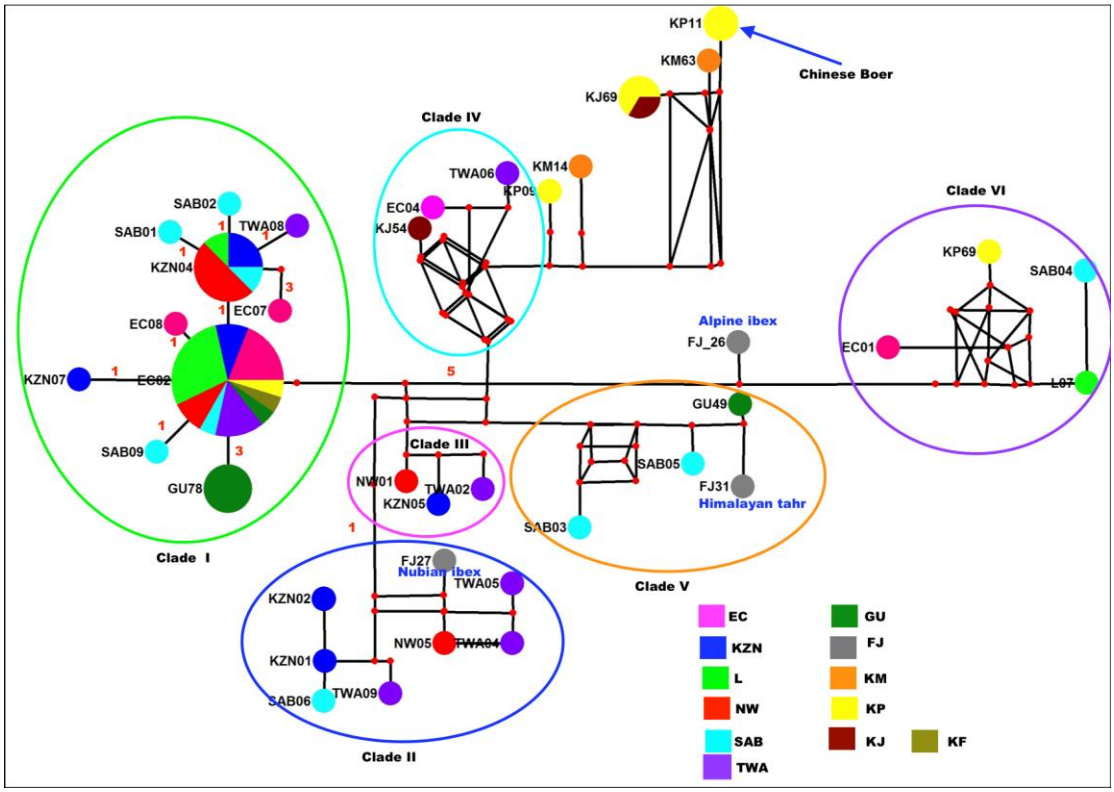


Figure 3.6 Median network joining of the whole mtDNA region of the SA and international goat breeds from the GenBank. Haplotypes are represented by a circle, and frequency of each haplotype by the size of the circle. Number of mutations between haplotypes and clades has been represented in red. The EC, KZN, L, NW, SAB and TWA are South African goat populations from Eastern Cape (EC), Kwa-Zulu Natal (KZN), Limpopo (L), North West (NW), South African Boer (SAB) and Tankwa feral (TWA) goats. The FJ are the wild goat populations FJ27 is the Nubian ibex found in North African and Arabian mountains, FJ31 is the Himalayan tahr found in the Asian mountains and FJ_26 is the Alpine ibex found in European mountains.

Clade I was the most dominant clade consisting of 33 haplotypes, two (EC02 and KZN04) were major haplotypes. The most predominant haplotype was EC02 with a

frequency of 21% and observed in 5 of the South African goat populations (EC, KZN, L, SAB, and TWA), with the L population being the most dominant in the haplotype. Haplotype KZN04 was the second dominant haplotype with a frequency of 8% and shared by 4 South African goat populations (KZN, L, NW and SAB). Clade II consisted of 7 haplotypes, which were from KZN, NW, SAB and the TWA goats while clade III on the other hand had three haplotypes from three populations (KZN, NW and TWA). Clade IV had two populations of the EC and TWA while Clade V was observed in SAB goats only. Clade VI contained of EC, L and SAB goat populations.

Table 3.6 Clade representation of the complete mtDNA genome showing relationships between South African goat populations.

*Population	Clade I	Clade II	Clade III	Clade IV	Clade V	Clade VI	Total
EC	6	0	0	1	0	1	8
KZN	5	2	1	0	0	0	8
L	7	0	0	0	0	1	8
NW	6	1	1	0	0	0	8
SAB	5	1	0	0	2	1	9
TWA	4	3	1	1	0	0	9
Total	33	7	3	2	2	3	50

*EC = village goats from Eastern Cape province; KZN = village goats from KwaZulu-Natal; LIM = village goats from Limpopo province; NW = village goats from North West province; SAB = South African Boer goats; TWA = Tankwa feral goats.

Clade I consisted of several GenBank sequences and 3 shared haplotypes. Haplotype EC02 from Clade I was shared with the Chinese goat (KF) which are from the Shannan White population of Chinese maternal lineage A. Haplotype KZN04 in this

same clade on the other hand was shared only between the South African goat populations whereas haplotype GU78 was shared only amongst the Chinese goat populations. Clade II clustered with the Nubian ibex (FJ27), whereas clade III contained of goats from the South African populations with no association with the GenBank sequences. Clade IV consisted of two South African populations and one Chinese goat of. Clade V on the other hand clustered with the wild species (Hamalayan tahr, FJ31) and the Mongolian native GU49. Clade VI was shared between South African and the Chinese goat populations. Unexpectedly, the Chinese Boer did not cluster with the SAB. The Alpine Ibex (FJ_26) was observed to be on its own and not with other populations whether from South Africa on GenBank. Clade III had only the South African populations showing no link with the GenBank sequences.

Clade I was defined as Chinese lineage A. The remaining clades however couldn't be grouped or described under any lineage since they did not cluster with any of the populations from the studied Chinese lineages.

3.4. Discussion

This study utilized a next generation sequencing strategy to generate complete mtDNA sequences and investigate population genetic structure and maternal lineages of SA goat populations. As described by Moloko (2011), a large percentage of the South African goats are farmed in communal areas under extensive production systems. South Africa also consists largely of the intensive production system, which is a mainly practiced in commercial farms using developed high performing breeds such as the SA Boer goat (Visser *et al*. 2004). The Tankwa population is kept as a feral population (Kotze *et al*. 2014). It has been reported that he South African commercial meat type goats have been developed from the indigenous village goat breeds (Campbell 2003, Mdlala *et al*. 2015). This study therefore hypothesized that the South African goat populations should be sharing maternal lineages even though their origin is unknown. The purpose of this study was therefore to investigate maternal lineages and origins of these goats as well as the sharing of maternal lineages between South African goat populations. Maternal lineages play a significant role in growth, production and reproduction traits as well as breed characterization

(Malau-Aduli *et al*. 2004). Knowledge on the genetic diversity and maternal lineages provides insight on important historic events such as domestication processes, which is crucial information assisting in breed characterization and differentiation (Doro *et al*. 2014).

The complete mtDNA was successfully captured with 22 tRNA genes and two rRNA that were distributed along the sequence, which was comparable to observations by Wu and coworkers (2012). A total of 184 SNPs were observed suggesting that the South African goat mtDNA was diverse. Wu *et al*. (2012) observed 124 SNPs in the Chinese goats, which is lower than was observed in this study. Large numbers of SNPs suggest a higher diversity and the South African goats were therefore more genetically diverse than the Chinese ones. As described by Masika and co-workers (2008), SA village goats are raised under challenging extensive environments and greater diversity as observed by the high number of SNPs and amino acids changes observed could be beneficial to the populations' ability to adapt and produce optimally. The tRNA genes were less polymorphic ranging from 0.01-8.00%. tRNA-Arg was highly polymorphic amongst the tRNA genes. The 12S rRNA was least polymorphic (0.7%) compared to 16S rRNA. Both rRNA and16S rRNA genes have not been reported to have any association with growth traits. ND3 had the highest number of SNPs compared to the rest of the genes (33%). It was however expected that the D-loop would be the most polymorphic as it has been reported to be hyper variable in other genetic diversity studies (Royo *et al*. 2008). In this study only 13% of the observed SNPs were on the D-loop suggesting that the D-loop was not as polymorphic in the South African goat populations. The D-loop on the South African goat populations was however more polymorphic than observed in previous studies such as by Luikart *et al*. (2001), Mannen *et al*. (2001) and Wu *et al*. (2012) who reported only 7% of SNPs to be on the D-loop. High diversity of the D-loop was observed in the Chinese goats in which 144 polymorphic sites were identified (Liu *et al*. 2007). The D-loop of Indian goat populations was also shown to be more diverse with 92 polymorphisms (Rana *et al*. 2013) compared to the 42 SNPs observed in South African goat populations. The variations in the D-loop of the Japanese Black cattle have been associated with carcass traits and some of the substitutions have been associated with meat quality in cattle (Malau-Aduli *et al*. 2004). The association of the mtDNA genome with growth traits has not been reported in goats.

Gene flow and genetic distances are important parameters that give insight into genetic variations within or between populations. The F_{ST} is one of the methods widely used to measure genetic variations and population structure (Rana *et al*. 2013). The F_{ST} values between pairs of populations in the South African goat populations suggested close relations between the village goats with F_{ST} values ranging from 0.010 to 0.036 with a significance P value of $p \leq 0.05$. These F_{ST} values implied that these populations shared some genetic variants. Phenotypic characterization has been observed in the village goat population and Morrison (2007) in the study of indigenous veld goats had characterized them according to their various phenotypic characteristics. The observation by Morrison (2007) suggested these populations would be genetically different. The geographic distance between these goat populations also brought about an expectation that the village goats from geographically distant provinces would be genetically distinct. The distance between Kwa-Zulu Natal and North West provinces for example was approximately 730km, which makes chances for the goats in these two provinces to share the same gene pool are very slim. However, the F_{ST} values observed suggested close relation and possible gene flow despite the geographic location and phenotypic diversity between populations. The results in this study suggested that the SA goat breeds were genetically similar at the mtDNA level regardless of where they are raised and irrespective of their production systems. These findings are supported by observations by Mdladla *et al*. (2016) based on genome-wide population structure analysis, and Campbell (2003) who described the origin of South African goats and suggested that the commercial breeds were developed from the indigenous ecotype (village) goats and therefore shared their gene pool. It was however unexpected that the Tankwa feral goat population would share clade with the farmed goats as it has been reported to be a genetically distinct population (Kotze et al., 2015; Mdladla *et al*. 2016).

Great genetic distance was expected, especially between the meat breeds of SAB and the undeveloped TWA and SA village goats due to different production systems and the fact that the village goats had not been selected and developed for any specific traits. Great genetic distance was also expected among the SA village goats due to geographical distance. The results in this study however prove otherwise and might be

explained by the nature of the mtDNA markers used to characterize this diversity. Since the mtDNA is maternally inherited as a single haplotype with no recombination, it tends to stay in populations even after breed divergence as signatures of maternal origins (Garrine 2007). Although the Tankwa population is a feral population that is found in completely different production system to that of the farmed goats, it was also found to have lower F_{ST} values between them and the Boer and village goats, which indicated similar genetic background. Other studies had shown the genetic distance between populations to correspond to the geographic distance as was observed in the Mehsana and the Odisha Coast where F_{ST} of 0.190 was observed between populations (Rana *et al*. 2013).

The degree of population sub-structuring analyzed using the analysis of molecular variance revealed high diversity within the South African goat populations with most of this variation due to within population diversity. This finding relates to high haplotype diversity, which was brought about by high number of SNPs and implies wide spread distribution and diversity to favor for selection within populations (Rana *et al*. 2013). The SA village goats together with the TWA populations have been reported to be highly diverse (Simela & Merkel 2008; Kotze *et al*. 2014) and the findings of this study correspond to suggestions from literature. Within population variation was higher in the South African goat populations as compared to populations from Nigeria, Iceland, Mongolia and Malaysia, and Burkina Faso (Rana *et al*. 2013; Luikart *et al*. (2001) and Royo *et al*. (2009). Very low population sub-structuring was observed (<3.00%) in all the categories under study.

The median network joining trees in this study presented larger/predominant haplotypes with smaller haplotypes surrounding them. A branch in the median network-joining tree represented a single nucleotide change (Rana *et al*. 2013). The larger predominant haplotypes such as haplotype EC02 were considered maternal lineages with new mutations surrounding them. The D-loop and the whole mtDNA networks both produced six clades and the COX1 tree produced only two clades. There had also been some populations sharing these clades irrespective of the gene from which the tree was derived. For instance, the EC population was consistent with clade I in all the genes and was not observed in clade II of all the genes and KZN was

not observed in clade II of the D-loop tree but was consistent in COX1 and complete mtDNA clade II. These clades can therefore be hypothesized as haplogroups or maternal lineages.

The COX1 gene is a highly conserved gene commonly used for barcoding and species identification and can accurately distinguish closely related species (Herbet *et al*. 2003). It will be expected that the COX1 gene will, based on its mutation rate show an ancient population structure. Hence, the South African goat populations that shared a haplotype and a clade with the ancient wild bezoar are believed to be representative of the ancestral goat populations (Nomura *et al*. 2013). The D-loop on the other hand will show a more recent population structure, as it is highly polymorphic. High polymorphism was demonstrated by the high number of singletons observed in the D-loop network tree (Liu *et al*. 2007; Rana *et al*. 2013) and the same was observed in the current study.

The mtDNA D-loop is one of the main tools that have been used to study the origins and genetic diversity of goats (Liu *et al*. 2007). It has also been reported as a region where most of the phylogenetic information is concentrated (Benjelloun *et al*. 2015). The median joining network analysis indicated multiple clades that were shared between populations, implying highly diverse maternal background that is shared amongst the goat populations of South Africa. This was expected especially for the commercially developed SA Boer based on literature confirming that the SA Boer goat was developed from the indigenous goats in the 20th century (Morrison 2007). Therefore, it does make sense for this breed to form clades with the SA village goats as seen on the network trees. For the feral Tankwa population, it was expected that it would form its own cluster based on the findings that it is a genetically unique population by Mdladla *et al*. (2016). However, in this study, this population does not come out as having separate maternal lineages as it lies within the clade of the farmed South African goats.

Liu *et al*. (2007) reported four mtDNA lineages (A-D), which either suggested four

maternal origins or one origin from a very high population that has four highly divergent lineages. D-loop haplotype EC02 was shared between the South African goat populations as well as Liu's Chinese LZG (Leizhou goat), which was found in haplotype H1of lineage A. The Chinese SNW (Shaannan White goat) from lineage A clustered as a mutation from haplotype EC02 in D-loop's clade I. This observation suggests that the South African goat populations share the same maternal lineages with the Chinese populations of lineage A. Lineage A has been reported to be the most common lineage found in all continents (Liu *et al*. 2007), making its presence in South African goats to be expected. It has been reported that Lineage A was probably derived from the Tibetan founders and was further domesticated in North China while other goats dispersed to South China (Zhao *et al*. 2013). Findings from this study therefore suggest that the South African goats may have originated from Tibet. Several domestication centers have been reported and the domestic goat is believed to have been domesticated in the Middle East (Cañón *et al* 2006). The Chinese goats have however been domesticated in West Asia then migrated east into China with the domestic goat originating from south west China and mid-Asia (Li *et al*. 2002). Li and co-workers (2002) have reported that the Tibetan goats are divided into two sub groups of which one group originates from south China and the other group from North China. The nomadic and semi-nomadic populations were known to live in the northern and northwestern parts of China before they migrated to other parts of the world and some dispersed and settled in the Northern parts of Africa (Wang *et al*. 2008). Campbell (2003) described the migration of the nomadic black and coloured (Khoikhoi) people who possessed domestic animals including goats from North Africa to the Southern parts of Africa through the narrow tsetse-free paths during the 5^{th} century. Human migration and commercial trading is an activity that has been happening over thousands of years ago (Benjelloun *et al*. 2015), this therefore adds as a supporting factor that goats would be transported from one region to the other, from continent to continent also though colonization of African counties therefore making it possible for goats to share maternal lineages irrespective of geographical location. A report on human migration by (Benjelloun *et al*. 2015) shows the possibility of goats being transported from Middle East to China, to Tibet, to North Africa and furthers down to South Africa, hence the relationship between the Tibetan and the South African goat populations. The D-loop clade IV consisted of both the South African goat populations and that of lineage B of Liu *et al*. (2007). Lineage B on the other hand is mostly found in the Indian subcontinent, Mongolia and Southeast Asia

(Liu *et al*. 2007). Tibet is very close to India (Zhao *et al*. 2013) suggesting possible gene flow between the Indian and the Tibetan goat populations. Lineage C and D were not observed in South African goat populations, which was expected as these have been reported to be rare lineages Liu *et al*. (2007).

Studies have suggested that goats were domesticated from a wild goat bezoar (*Capra aegagrus*) in west Asia (Nomura *et al*. 2013). The COX1 network in this study did show ancient population structure as well as haplotype sharing and close relationships with the ancient wild bezoar, the suggested ancestor of the domestic goat. The EC01 haplotype did however show very close relations with the Bangladesh Black Bengal (AB01). The clustering of the South African haplotype with the bezoar may be supporting evidence that the bezoar is the ancestor of the domestic goats (Naderi *et al*. 2007).

It was expected that the TWA (Tankwa feral) goat population would either share a haplotype or form a large part of clade II of the wild bezoar, as it is also a wild goat. However, only few Tankwa goats were observed in this clade. Clade II in this study also revealed close relationships between the South African goat population and the Mongolian native goats, which was expected due to the D-loop results in this study and those of Liu *et al*. (2007) that lineage B, is also found in Mongolia.

This study assumed that the complete mtDNA would produce a more accurate population structure based on both the highly conserved genes such as COX1 and the highly polymorphic ones like D-loop which would defines better phylogeny. The increased number of haplotypes it produced in this study suggested the accuracy of the complete mtDNA usage in maternal lineage studies. The current study is supported by that of Wu *et al*. (2012) that the complete mtDNA would result in a more accurate and better population structure.

Haplotype EC02 of the complete mtDNA consisted a Chinese Shaannan White goat

described by Liu and co-workers (2007) as haplotype H1 from lineage A. On the D-loop analysis, two of the Chinese lineage A populations were found to have clustered and shared D-loop haplotype H1 with the South African goat populations, stipulating that the South African goat populations share maternal lineage A with the Chinese goat population; this result is further confirmed by the complete mtDNA sequences. The other important thing to note is that, in the D-loop network tree, the SAB population did not share a haplotype with any of the populations, though some of the individuals were found to have clustered with the lineage A population, however in the complete genome analysis this population shares haplotypes EC02 with other populations as well as that from lineage A. Wu and co-workers (2012) sampled and analyzed four populations that represented four maternal lineages defined by Liu *et* al. (2007). It was observed in this study that all the four populations shared the same haplotype (GU78), which shows close relations with EC02 suggesting that GU78 mutated from EC02, as this is the major haplotype in this tree. Whilst the D-loop analysis shows the non-existence of lineage C and D, complete mtDNA genome analysis shows traces of these two lineages and possible mutations from lineage A. This further confirms the suggestions by Wu *et al*. (2012) that the complete mtDNA sequence analysis is a much better and more accurate way to study and characterize maternal lineages. However, more studies would have to be performed to confirm these lineage traces as well as mutations perhaps by increasing the sample size of the whole genome lineage C and D, which is currently a challenge as maternal lineage studies using complete mtDNA genome are rare.

The whole mtDNA analysis shows close relations between the North Eastern Arabia and African (*Nubian ibex* > FJ27), suggesting a possible genetic influence by this species as it formed part of cluster II with the South African goat populations, including the feral Tankwa population. This result is supported by the migration route of the Nomadic people who settled in North Africa before coming to the Southern parts of Africa and lastly South Africa (Campbell 2003).

The D-loop together with the complete mtDNA genome networks brings a strong suggestion that the South African goat populations originate from maternal lineage A

from Tibet. The origin of the highly commercially developed SA Boer has been in question over the years (Campbell 2003; Morrison 2007). In this study, the SAB population was observed in maternal lineage A, clade IV of the D-loop as well as maternal lineage A, haplotype EC02 and clade I of the whole mtDNA which means that the SAB though developed in South Africa from the SA village (indigenous) goat populations, share the same maternal lineages with the village goat populations suggesting the same origins of lineage A. Singleton haplotypes that did not cluster with any maternal lineages in literature are novel and unique to South Africa.

3.5 Conclusion

The complete (16.64kb) mtDNA was successfully generated by the use of rolling circle amplification and Illumina next generation sequencing technology. The mtDNA was highly polymorphic in South African goat populations. Low F_{ST} values suggest very close relations and possible gene flow. High within and low between population variation was observed in all the populations under study. Analysis of molecular variance results further suggests that there is no population sub-structure in the South African goat populations. The presence of lineage A and B in South African goat populations suggest that the South African goats share maternal lineages with the Chinese goat populations. The median network joining of the COX1 further suggests close relations with the wild bezoar, which may mean that the South African goats also originate from there. The complete mtDNA network shows a better structure and as compared to the D-loop and the COX1 trees. The D-loop and then complete mtDNA genome networks suggest that the South African goat populations originate from Tibet. Unique and novel clades were also observed in this study. This study therefore suggests that the complete mtDNA sequence is a better way to study maternal lineages and population structure.

Acknowledgements

The Agricultural Research Council Biotechnology Platform and the National Research Foundation funded this study. Ms. Ncube was funded through the NRF Innovative Scholarship and ARC-Professional Development Program, University of South Africa.

3.6. References

Bandelt H.J., Forster P. & Röhl, A. (1999) Median-joining networks for infer- ring intraspecific phylogenies. *Molecular Biology and Evolution*. 16, 37-48.

Benjelloun B., Alberto FJ., Streeter I, Boyer F., Coissac E., Stucki S., BenBati B., Ibnelbachy M., Chentouf M., Bechchari A., Leempoel K., Aalberti A., Engelen S., Chikhi A., Clarke L., Flicek P., Joost S., Taberlet P., Pompanon F., & NextGen Consortium (2015) Characterizing neutral genomic diversity and selection signatures in indigenous populations of Moroccan goats (*Capra hircus*) using WGS. *Frontiers in Genetics* 6, 107.

Biase H. F., Meirelles F. V., Gunski R., Vozzi P. A., Bezerra L. A. F, Vila R. A., Rosa A. J. M., Lôbo R. B. & Martelli L. (2007) Mitochondrial DNA single nucleotide polymorphism associated with weight estimated breeding values in Nelore cattle (Bos indicus). *Genetics and Molecular Biology* 30, 1058-1063.

Campbell Q. P. (2003) The origin and description of Southern Africa's indigenous goats. *South African Animal Science* 4, 18-22.

Can̰ón J., Garćıa D., Garćıa-Atance M. A., Obexer-Ruff G., Lenstra J. A., Ajmone-Marsan P., Dunner S. & The ECONOGENE Consortium. (2006). Geographical partitioning of goat diversity in Europe and the Middle East. *International Society for Animal Genetics, Animal Genetics*. doi:10.1111/j.1365-2052.2006.01461.

Casey N. H., & Webb E. C. (2010) Managing goat production for meat quality. *Small Ruminant Research*. DOI: 10.1016/J.SMALLRUMRES.2009.12.047.

Corte´s O., Tupac-Yupanqui I., Dunner S., Garc´ıa-Atance M. A., Garc´ıa D., Ferna´ndez J. & Can˜o´n J. (2008) Ancestral matrilineages and mitochondrial DNA diversity of the Lidia cattle breed. *International Society for Animal Genetics, Animal Genetics* 10.1365-2052.

Doro M. G., Piras D., Leoni G. G., Casu G., Vaccargiu S., Parracciani D., Naitana S., Pirastu M. & Novelletto A. (2014) Phylogeny and patterns of diversity of goat mtdna haplogroup A revealed by resequencing complete mitogenomes. *PLOS ONE* 9, e95969.

Garrine C.M.L.P. (2007) Genetic characterization of indigenous goat populations of Mozambique. Department of Production Animal Studies in Faculty of Veterinary Science, University of Pretoria, South Africa.

Gwaze F.R. (2009) Communal production systems of goats raised by resource-poor farmers in the Eastern Cape Province of South Africa. Department of Livestock and Pasture Science, University of Fort Hare, South Africa.

Han L., Yu H., Cai D., Shi H., Zhu H. & Zhou H. (2010) Mitochondrial DNA analysis provides new insights into the origin of the Chinese domestic goat. *Small Ruminant Research* 90, 41–46.

Herbert PDN., Cywinska A., Ball S. L. & deWaard J.R. (2003) Biological identifications through DNA barcodes. *The Royal Society* 270, 313-321.

Jeon G. J., Chung H. Y., Choi J. G., Lee M. S., Lee C. W., Park J. J., Ha J. M., Lee H. K. & Sung H. H. (2005) Relationship between genetic variants of mitochondrial DNA and growth traits in Hanwoo cattle. *Journal of Animal Science* 18, 301-307.

Kotzé A., Grobler J.P., van Marle-Köster E., Jonker T. & Dalton D.L. (2014) The Tankwa Karoo National Park feral goat population: A unique genetic resource. *South African Journal of Animal Science* 44, 43-48.

Li M., Zhao S., Bian C., Wang H., Wei H., Liu B., Yu Mei., Fan B., Chen S., Zhu M., Li S., Xiong T. & Li K. (2002) Genetic relationships among twelve Chinese indigenous goat populations based on microsatellite analysis. *Genetic Selection and Evolution* 34, 729–744.

Liu R. Y., Lei C. Z., Liu S. H. & Yang G. S. (2007) Genetic diversity and origin of Chinese domestic goats revealed by complete mtDNA D-loop sequence variation. *Asian-Australian Journal of Animal Science* 20, 178-83.

Luikart G., Gielly L., Excoffier L., Vigne J., Bouvet J., & Taberlet P. (2001) Multiple maternal origins and weak phylogeographic structure in domestic goats. *PNAS* 98, 5927-5932.

Malau-Aduli A. E. O.,. Nishimura-Abe A., Niibayashi T., Yasuda Y., Kojima T., Abe S., Oshima K., Hasegawa K. & Komatsu M. (2004) Mitochondrial DNA polymorphism, maternal lineage and correlations with postnatal growth of japanese black beef cattle to yearling age. *Asian-Australian Journal of Animal Science 2004* 17, 1484-1490.

Mannen H., Nagata Y., & Tsuji Soichi (2001) Mitochondrial DNA reveal that domestic goat (*Capra hircus*) are genetically affected by two subspecies of bezoar (*Capra aegagurus*). *Biochemical Genetics* 39, 149-154.

Masika PJ., Mafu JV., Goqwana MW., Mbutiand C. & Raats J. (2007) A comparison of goat growth performance in a communal and commercial farming system in the Central Eastern Cape Province, South Africa.

Mdladla, K., Dzomba, E., Huson, H., Muchadeyi F.C., 2015. Population structure and admixture of South African goat breeds investigated using genome-wide SNP data. Animal Genetics. doi: 10.1111/age.12442.

Moloko B. (2011) A profile of the South African goat market value chain. department of griculture, fishery and forestry.

Morrison J. W. (2007) A Guide to the identification of the natural indigenous goats of Southern Africa. *South African Journal of Animal Science* 2003, 4.

Naderi S., Rezaei H-R., Taberlet P., Zundel S., Rafat S-A., Naghash H., El- Barody M.A. A., Ertugrul O., Pompanon Francois. for Econogene Consortium (2007) Large-scale mitochondrial dna analysis of the domestic goat reveals six haplogroups with high diversity. *PLoS ONE* 2: e1012.

Nomura K., Yonezawa T., Mano S., Kawakam S., Shedlock A. M., Hasegawa M. & Amano T. (2013) Domestication process of the goat revealed by an analysis of the nearly complete mitochondrial protein-encoding genes. *PLOS ONE*, 8, Issue 8.

Odahara S., Chung H. J., Choi S. H, Yu S. L., Sasazaki S., nn H., Park C. S., & Lee J. H. (2006.) Mitochondrial DNA diversity of Korean native goats. *Asian-Australian Journal of Animal Science* 19, 482-485.

Rana S., Singh S., Dureja V., Joshi J., Banerjee P., Sharma U., Viji R. K. (2013) Mitochondrial D-loop analysis reveals single maternal lineage for coastal region goats of India. *International Journal Of Biomedical and Life Sciences.* 4, 235-244.

Royo L. J., Traore´ A., Tamboura´H. H., lvarez I. A´., Kabore´A., Ferna´ ndez I ., Oue´ draogo- Sanou G., Toguyeni A., Sawadogo L. & Goyache F. (2009) Analysis of mitochondrial DNA diversity in Burkina Faso populations confirms the maternal genetic homogeneity of the West African goat. *International Society for Animal Genetics, Animal Genetics* 40, 344-347.

Simela L. & Merkel R. (2008) The contribution of chevon from Africa to global meat production. *Meat Science* 80, 101-109.

Schneider S., Roessli D. & Excoffier L. (2000) ARLEQUIN Ver. 2.0: A software for population genetics data analysis Department of Anthropology, University of Geneva, Geneva, Switzerland.

Visser C., Hefer C.A., van Marle-Köster E. & Kotze A. (2004) Genetic variation of three commercial and three indigenous goat populations in South Africa. *South African Journal of Animal Science* 34, 24-27.

Wu Y., Huo J., & Gua W. (2012) Sequence analysis of four caprine mitochondrial DNA lineages. *Italian Journal of Animal Science* 11, 378-383.

Zhao Y., Zhao R., Zhao Z., Xu H., Zhao E., & Zhang J. (2014) Genetic diversity and molecular phylogeography of Chinese domestic goats by large-scale mitochondrial DNA analysis. *Molecular Biology and Reproduction* 41, 3695-3704.

CHAPTER 4

Targeted high throughput growth hormone gene sequencing reveals high within breed genetic diversity in South African goats*

Abstract

Genetic variations in the growth hormone gene have been observed in several livestock species and have potential use in growth improvement and selection programs. This study assessed the genetic diversity in the growth hormone gene within and between South African (SA) goat breeds. Polymerase chain reaction targeted gene amplification together with Illumina MiSeq next generation sequencing were used to generate the full length (2.54kb) of the growth hormone gene and screen for SNPs in the South African Boer (n=17), Tankwa (n=15), and village (n=35) goat populations. A range of 27 to 58 SNPs per population were observed. Mutations resulting in amino acid changes from glycine to serine, tyrosine to cysteine and arginine to glycine were observed at exon 2 on position 781bp and exon 5 positions 2012bp and 2017bp respectively. Gene diversity ranged from 0.8268 ± 0.0410 to 0.9298 ± 0.0050. Higher within breed diversity of 97.37% was observed within the population category consisting of SA village ecotypes and the Tankwa goats. Highest pairwise F_{ST} values ranging from 0.148 to 0.356 were observed between the SA Boer and both the SA village and Tankwa feral goat populations. Seven additional sequences were extracted from GenBank for phylogenetic analysis. The maximum likelihood phylogenetic analysis indicated nine genetic clusters which, reflected close relationships between the South African populations and the other international breeds with the exception of the Italian Sarda breeds. Results imply great potential for within population selection programs particularly with SA village goats.

Key words: Growth, Genetic diversity, Growth Hormone Gene, Goats

*This chapter has been published in the Journal of Animal Genetics under the title "Targeted high throughput growth hormone gene sequencing reveals high within breed genetic diversity in South African goats"

4.1 Introduction

South Africa is one of the major goat-producing countries, having approximately 6.2 million goats (FAOSTATS 2013), which make up 3% of African goats and 1% of the world's goat populations (Moloko 2011). Goat farming is predominantly (63%) in the communal farms that are distributed in all provinces of South Africa (Moloko 2011; Lahiff & Cousins 2005). Goats raised by communal farmers improve food security and alleviate poverty through sales (Simela & Merkel 2008; Gwaze 2009). In communal areas, goats are also used for ceremonial and traditional functions (Masika *et al*. 2007) such as marriage, burials and other rituals (Moloko 2011) and also provide manure for home gardens (Gwaze 2009). Goat meat is of great importance as it provides easily accessible high quality protein to the poor families (Gwaze 2009). It contains less fat and cholesterol and has high levels of protein and iron (NAMC 2005).

South Africa has three indigenous goat breeds namely the South African Boer, Savannah and Kalahari Red that are recognized meat breeds used by commercial farmers nationally and around the world (Visser *et al*. 2004). South African village goats are non-descript indigenous ecotype populations that are kept and raised by communal farmers in villages (Morrison 2007). The Tankwa is a feral goat population that is kept as a conserved flock in the Northern Cape Province of South Africa (Kotze *et al*. 2014). The commercial goats are bred for high growth rates and meat yields (NAMC 2005), and contributes good quality meat and income through the chevon industry. Regardless of their desirable traits, commercial breeds cannot be sustainably grown in communal farming systems. NAMC (2005) reported that fast growing breeds such as the SA Boer goat require more feed and nutrients for maintenance, which makes them expensive and unaffordable particularly for resource-poor communal farmers. Communal goat production is characterized by low input management practices often associated with suboptimal nutrition and housing, inadequate veterinary services and lack of breed/animal improvement programs (Masika *et al.* 2007; Gwaze 2009). As observed by Gwaze (2009), diseases such as anemia, diarrhea and helminth parasites infections are major constraints to goat productivity in communal farming systems. Unlike the commercial goats, the SA village goats kept by most of the communal farmers are well adapted to travelling long distances in search of feed and water and are able to utilize low quality feed

(Gwaze 2009). The majority of the communal farmers do well with these village goats that are well adapted to the local environments and production conditions (Pieters 2007). The village goats are also known for their resistance to diseases and adaption to a wide range of climatic conditions (Webb & Mamabolo 2004). Despite their advantages in the smallholder sector, village goats are generally characterized by poor feed conversion efficiency and growth rates. It has been suggested that improving the production performance such as growth and other economic traits of the adapted SA village goats will be a feasible option for the resource limited communal farmers.

Although considered highly adapted and suitable for communal farming systems, nothing is known about the actual genetic potential of the South African village goat populations. The village goats are rather categorized according to their geographical locations (Morrison 2007). They have variable mature body sizes although mostly characterized as medium sized breeds (Campbell 2003). According to Simela & Merkel (2008), there is high variability observed among and within SA village goat populations that makes it feasible to improve growth and meat production through within population selection and breeding programs.

Growth is a quantitative trait that is influenced by many genes such as the growth hormone, caprine myostatin, caprine pituitary specific transcription factor-1 and bone morphogenic genes (Supakon 2009). Polymorphisms in genes such as the growth hormone have been associated with pre- and post-natal growth in Egyptian and Saudi goat populations (Alakilli 2012). Hua *et al*. (2009) has demonstrated that some growth hormone genotypes are responsible for the changes in growth traits in Chinese Boer goats. Studies by Hua *et al*. (2009) and Wickramaratne *et al*. (2010), suggested that informative genes such as growth hormone gene could be used as markers for growth associated studies and selection for high growth performance.

Targeted gene and next generation sequencing technologies have recently been used for studying gene variations and identifying genotypes and genes in animals (Crosby & Criddle 2007). The growth hormone gene is a 2.544kb fragment that consists of 5

exons and 4 introns (Alakilli *et al*. 2012). It is responsible for the production of growth hormone from the anterior pituitary, which plays a role in the metabolism and growth of mammals (Supakon 2009). This study used targeted gene capture by polymerase chain reaction (PCR) and Illumina-based next generation sequencing technology to generate the full length growth hormone gene sequence data and analyze genetic diversity within and between the different South African goat breeds. Information generated was used to infer on diversity in growth potential of the South African goat breeds as well as feasibility of selection programs.

4.2 Materials and methods

4.2.1 Goat populations and blood collection

Blood samples (6ml) were collected from the jugular vein of 67 South African goats into EDTA tubes and placed in a cooler box with ice packs and transported to the Agricultural Research Council, Biotechnology Platform laboratories were it was kept at -20°C till further use. The goats were sampled from three indigenous goat populations of South Africa, namely the Boer (SAB, n = 17), Tankwa (TWA, n = 15) and non-descript-village goats (n = 35). The Boer goats are commercial meat type goats and were sampled from commercial farms in the North West province of the country. The Tankwa is a captive feral goat population that was sampled from Carnavon Research Station in the Northern Cape province. The non-descript village goats were sampled from the Eastern Cape (EC, n = 10), KwaZulu-Natal (KZN, n = 9), Limpopo (LIM, n = 10), and North West (NW, n = 6) provinces as shown in Table 2.1. Only one goat was sampled per communal household.

4.2.2 DNA isolation and quantification

DNA was isolated using an optimized Qiagen DNeasy blood and tissue kit (www.qiagen.com) extraction protocol. DNA quantification was performed using the Invitrogen's Qubit™ dsDNA BR Assay Kit (Invitrogen, Life Technologies). The quality of the bands was determined on 1% agarose gel with 4µl of ethidium bromide run at 80V for 30min. The gel was examined under UV light in a BIORAD Imaging System (BIORAD GelDoc XR) (www.bio-rad.com).

Table 4.1 The distribution of goat sampled from both commercial and village goats from Eastern Cape, Kwa-Zulu Natal, Limpopo, North West and Northern Cape provinces of South Africa.

Goat breed	**Sample ID**	**No. Per breed**	**Production system**	**Location**
Boer goat	*SAB	17	Intensive	Stella farm (n=5) and van Tonder farm (n=12) – North West
Tankwa	*TWA	15	Captive feral	Carnavon Research Station – Northern Cape
Village	*EC	10	Extensive	Hobeni village (n=5), Elliotdale municipality and Pechelsdam village (n=5), Inxubayethemba municipality - Eastern Cape Province
	*KZN	9	Extensive	Coniliva (n=4) and Ngubo (n=4) villages, Msinga municipality - Kwa-Zulu Natal Province
	*L	10	Extensive	Mukovhabale village (n=4), Mutale municipality, Muila-muumone village (n=5) - Limpopo Province
	*NW	6	Extensive	Pella village, Moses Kotane municipality, North West Province

*EC = village goats from Eastern Cape province; KZN = village goats from KwaZulu-Natal; L = village goats from Limpopo province; NW = village goats from North West province; SAB = South African Boer goats; TWA = Tankwa feral goats.

4.2.3 Growth hormone gene sequencing

4.2.3.1 Primer designing and gene amplification and sequencing

A reference growth hormone sequence from GenBank (Accession: D00476.1) was used to design primers using the Primer3 tool on (http://www.frodo.wi.mit.edu/). The primers *GHG*-F: 5'-*GATGATGATGAGCCTGGGGG*-3' and *GHG*-R: 5'-*GGGTGTGTCACAGAGAAGGG*-3', validated using PRIMER Blast tool on the NCBI website (http://www.ncbi.nlm.nih.gov.za), were 20bp long and had a melting temperature of 59.96°C. Polymerase chain reactions (PCR) were performed in 50µl total reaction volumes containing 2.0µl of 50ng genomic DNA, 1.0µl 10pmol of each forward and reverse primers, PCR multiplex master mix and nuclease free water. The amplification was carried out on an ABI GeneAmp 9700 thermocycler using optimized cycling conditions that involved an initial denaturation step at 93°C for 3 minutes, followed by 35 cycles at 93°for 15 seconds, annealing at 62°C for 30 seconds, 72°C for 2 minutes and 30 seconds, and a final extension of 10 minutes at 72°C. Amplicons were electrophoretically separated on 2% agarose gel and examined under UV light in a BIORAD Imaging System (BIORAD GelDoc XR) (www.bio-rad.com). Amplicons were purified using the MinElute Qiagen PCR purification kit (www.qiagen.com) according to manufacturer's instructions. Sequencing preparations were performed using the Nextera® v.2 Sample Prep for the MiSeq from Illumina (www.illumina.com/miseq). The products were sequenced on the Illumina MiSeq Bench-top Sequencer (Illumina Inc., San Diego, CA, USA) using the MiSeq Reagent Kit v3 at 2 x 300bp read length configuration (www.illumina.com/miseq).

4.3.3.2 Sequence read quality control

Sequence reads were checked for quality based on quality score (QC >30) and number of ambiguous nucleotides (0.00); and trimmed to remove adaptor sequences using the CLC Genomics Workbench v8.0.1 (CLC Bio, Aarhus, Denmark). Trimmed reads were mapped to the reference gene (Accession: D00476.1) obtained from NCBI (http://www.ncbi.nlm.nih.gov.za) using the "*Map reads to a reference"* tool of CLCBio genomics workbench v.8.0.1 (CLC Bio, Aarhus, Denmark) at a similarity factor of 0.9. Mapped reads were selected for further analysis.

4.3.4 Data analysis

4.3.4.1 Single Nucleotide Polymorphisms

The *Probalistic Variant Detection* tool in CLCBio genomics workbench v.8.0.1 (CLC Bio, Aarhus, Denmark) was used to screen for single nucleotide polymorphisms (SNPs) as well as detect SNPs that would result in amino acid changes, using the same reference gene used for mapping. MEGA version 6.0 (Tamura *et al*. 2011) was used to analyze for distribution and frequencies of polymorphic sites on sequences generated.

4.3.4.2 Within population diversity

Gene and haplotype diversity were calculated using Arlequin version 3.5 (Schneider *et al.* 2000). Gene/Haplotype diversity ($\hat{H}$) was defined as:

$$\hat{H} = \frac{n}{n-1}\left(1 - \sum_{i=1}^{k} p_{i^2}\right)$$

where n is the number of gene copies in the sample, k is the number of haplotypes, and p_i is the sample frequency of the i-th haplotype (Nei 1987).

4.3.4.3 Between population diversity and population substructure

Variation between pairs of sub-populations was measured using pairwise F_{ST} estimates (Weir and Cockerham 1984). Analysis of molecular variance (AMOVA) implemented in Arlequin v3.5 (Schneider *et al.* 2000) was used to determine diversity within and among goat sub-populations. AMOVA was run for different population categories of (i) the four SA Village goats; (ii) the four SA Village goats and the South African Boer; (iii) the four SA Village goats and the Tankwa goat and (iv) all six subpopulations.

4.3.4.4 Phylogenetic analysis

Seven additional sequences (Table 4.2) were extracted from the GenBank database for phylogenetic analysis.

Table 4.2 A representation of sequences downloaded from NCBI (http://www.ncbi.nlm.nih.gov.za) for comparison in phylogenetic analysis.

Accession no.	**Breed**	**Origin**
D00476.1	Unknown	Japan
GU355689.1	Sarda	Italy
GU355688.1	Sarda	Italy
GU355687.1	Sarda	Italy
GU355686.1	Sarda	Italy
DQ531712.1	Beetal	Pakistan
EU651859.1	Jamunapari	India

A maximum likelihood phylogenetic tree was constructed using MEGA v6.0 (Tamura *et al*. 2011) to investigate relatedness and divergence within and between populations, the Tamura-Nei substitution model was used for nucleotide data with InDels. The model was selected on the analysis settings of MEGA v6.0 with 1000X bootstraps for this analysis. The sheep (*Ovis aries* accession: EF077162.1) growth hormone gene was used as an out-group in the analysis.

4.3. Results

4.3.1 Growth hormone gene capture and mapping

In total, 1 455 937 to 2 354 484 million reads of sequence data were generated for each animal. The average trimmed read length was 188 nucleotides (nts). For each

animal, a consensus sequence of 2.54kb was generated (Figure 4.1) and used for further analysis.

Figure 4.1 A figure of the growth hormone gene read mappings showing the reference gene and the 2.54kb consensus sequence generated from read mappings with exons and introns.

4.3.2 Single nucleotide polymorphism (SNP) analysis

A total of 142 SNPs were observed in the 67 sequences of the South Africa goat populations. A number of insertions, deletions, replacements and amino acid (*AA*) changes were observed. The number of SNPs ranged from 27 in the Eastern Cape goats to 58 in goats from Limpopo as illustrated in Table 8. The type and distribution of SNPs is illustrated in Appendix Table A8.

The majority of the SNPs were observed on exons. SNPs were also observed on introns. Exon 5 had the highest number of SNPs (n = 7) as compared to other exons whilst intron 4 had the most SNPs (n = 9). There were no SNPs observed on exon 1.

Table 4.3 Distribution of SNPs, InDels and replacements on the growth hormone of SA goat populations.

*Population	No. SNPs	Insertions	Deletions	Replacements
EC	31.8	2.3	0.6	1
KZN	27.4	1.3	0.25	0.5
L	57.9	2.5	1.2	2.8
NW	49.5	1.5	0.83	2.5
SAB	39.1	2.2	1.18	0.4
TWA	43.9	2.5	0.8	1.47

*EC = village goats from Eastern Cape province; KZN = village goats from KwaZulu-Natal; L = village goats from Limpopo province; NW = village goats from North West province; SAB = South African Boer goats; TWA = Tankwa feral goats.

Mutations resulting in amino acid changes (*Tyrosine* to *Cysteine, Arginine* to *Glycine* and *Glycine to Serine)* were observed on exon 5 positions 2012 and 2017, and exon 2 at position 781. The EC population had the highest number of AA changes with 3 changes of *Tyr>Cys, Arg>Gly and Gly>Ser* observed in the same population. The most common AA change observed was *Gly>Ser* on exon 2 position 781 which was observed in all the populations that had AA changes. There were no AA changes observed in the SA Boer goat population. The EC population is the only population that had AA changes on exon 5.

An average of 26.83 polymorphic sites were identified in the South African goat populations. The Eastern Cape population was more polymorphic with 40 sites while the KwaZulu-Natal having the least (n = 20) polymorphic sites Table 4.4.

Table 4.4 A summary of the diversity of the growth hormone gene shown by polymorphic sites, haplotypes and gene diversity within South African goat breeds.

*Population	N	Number of polymorphic sites	Mean gene diversity ± Standard Deviation
EC	10	40	0.889 ± 0.014
KZN	09	20	0.879 ± 0.018
L	10	24	0.889 ± 0.014
NW	06	21	0.827 ± 0.041
SAB	17	27	0.930± 0.005
TWA	15	29	0.930 ± 0.006

*EC = village goats from Eastern Cape province; KZN = village goats from KwaZulu-Natal; L = village goats from Limpopo province; NW = village goats from North West province; SAB = South African Boer goats; TWA = Tankwa feral goats.

4.3.3 Within population haplotype diversity

The 67 sequences of the South African goat population yielded 41 haplotypes (Table 4.5),10 of the 41 samples were shared within and between populations and 31 were unique (Appendix Table A9). The sequences of the 10-shared haplotypes were deposited in GenBank (KU288603-KU288612). The South African Boer (SAB) population had the highest number of shared haplotypes (n = 10) whilst NW ecotypes had the least (n = 2). SAB and TWA populations had an equal number of unique haplotypes (7), which was higher than all the populations under study and KZN was the lowest (3). Gene diversity within populations ranged from 0.879 for KwaZulu-Natal to 0.930 for South African Boer goats (Table 4.5).

Table 4.5 A representation of haplotypes within and between South African goat populations.

	*Population						
Haplotype	**EC**	**KZN**	**L**	**NW**	**SAB**	**TWA**	**Total**
GH1	2	0	1	1	0	0	4
GH2	1	3	0	0	0	0	3
GH3	1	0	0	0	0	1	2
GH4	1	0	0	0	2	0	3
GH5	0	2	0	0	0	0	2
GH6	0	1	3	0	0	0	4
GH7	0	0	1	1	0	0	2
GH8	0	0	0	0	4	0	4
GH9	0	0	0	0	4	4	8
GH10	0	0	0	0	0	4	4
Total	5	6	5	2	10	9	36

*EC = village goats from Eastern Cape province; KZN = village goats from KwaZulu-Natal; LIM = village goats from Limpopo province; NW = village goats from North West province; SAB = South African Boer goats; TWA = Tankwa feral goats.

Haplotype sharing amongst populations are illustrated in Table 4.5. There were no haplotypes that were shared amongst all the populations under study. Haplotype GH1 was shared among SA village goats only and not found in the commercial and Tankwa goats. Haplotype GH2 was also shared between village goats from Eastern Cape and KwaZulu-Natal while haplotype GH3 was between Eastern Cape village goats and the Tankwa feral population. Haplotype GH6 and GH7 were shared between only two SA village populations, GH6 (KZN and L) and GH7 (L and NW).

Haplotype GH8 was observed only in the SAB populations whilst GH10 was observed only in the TWA populations. Haplotype GH9 was shared between SAB and TWA populations. Haplotype GH8 and GH10 were observed in single populations of SAB and Tankwa goats, respectively.

4.3.4 Pairwise F_{ST} genetic distances

Genetic distances (F_{ST}) between pairs of populations ranged from 0.005 between Tankwa feral goats and village goats from Eastern Cape province to 0.356 between South African Boer and village goats from KwaZulu-Natal province (Table 4.6).

Table 4.6 Pairwise F_{ST} values showing genetic distances between South African goat populations.

*Population	EC	KZN	L	NW	SAB	TWA
EC						
KZN	0.084					
L	0.005	0.153				
NW	0.051	0.234	0.108			
SAB	0.229	0.356	0.148	0.323		
TWA	-0.005	0.049	0.037	0.034	0.296	

*EC = village goats from Eastern Cape province; KZN = village goats from KwaZulu-Natal; LIM = village goats from Limpopo province; NW = village goats from North West province; SAB = South African Boer goats; TWA = Tankwa feral goats.

The F_{ST} values between pairs of SA village goats ranged from 0.005 between Limpopo and Eastern Cape provinces to 0.234 between North West and KwaZulu-Natal populations (Table 4.6). The Tankwa goats differed greatly from the South African Boer (F_{ST} = 0.296) with very small differences observed between it and the SA village goat populations (F_{ST} ranged from 0.005-0.049). Overall highest pairwise F_{ST} values (0.148 - 0.356) were observed between the South African Boer and both the SA village and Tankwa feral goat populations. All values were significant at a $P \leq 0.05$.

Low within population variation (68.64%) was observed in the category consisting of the SA village goats and the Boer (Table 4.7).

Table 4.7 Analysis of Molecular Variance among South African goat populations.

Populations compared	**Percentage (%) of variation**		
	Among groups	**Among populations within groups**	**Within population**
SA Village goats only	0	7.44	92.56
SA Village and Boer	24.42	1.94	68.64
SA Village and Tankwa	5.27	7.90	97.37
All six groups	24.15	2.67	73.18

The level of population substructure was low in these four SA village goats as well as the category consisting of four SA village and Tankwa goats (% of among group variation <1.5). Moderately high level of population substructure was observed in the categories consisting of SA village and Boer goats as well as the group consisting of all six sub-populations (% of among group variation > 24%) as shown in Table 4.7.

4.3.5 Phylogenetic analysis

The phylogenetic analysis resulted in 9 clades distributed across the South African and GenBank sequences from Japan, Pakistan, India and Italy (Figure 4.2). Clades were constructed using the "*Nearest Neighbour Interchange*" described by Tamura Nei.

Clade V had the highest number of animals (n=23) whilst only one animal from EC (EC08) clustered in clade VIII. The maximum likelihood (ML) comparison indicated close relationships within the SA goat breeds that clustered together and represented in eight of the nine clades (Table 4.8). Almost all the clades were composed of the SA goat populations except for Clade IX consisting of only the Italian breeds. South African goat populations did not cluster per breed/population, clades consisted of a mixture of the South African goat populations irrespective of breed.

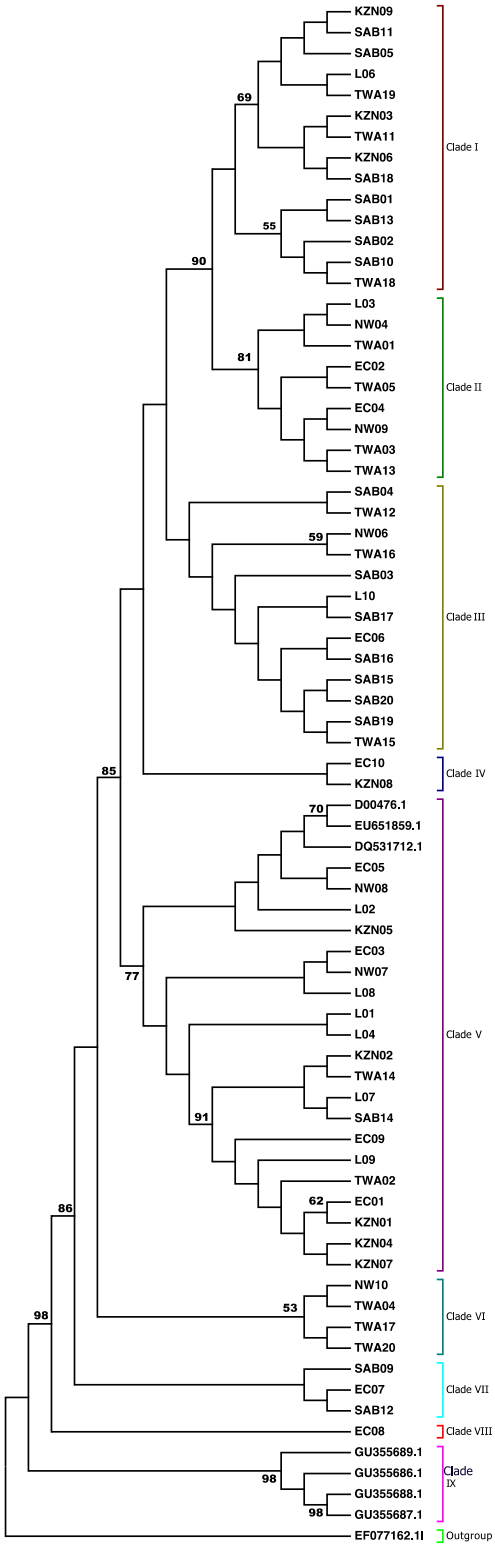


Figure 4.2 A maximum likelihood tree showing relationships between the South African and other GenBank sequences. Bootstrap "Cut off" was set to the value of 50 upon construction of the tree.

Table 4.8 South African goat breeds and their clade representations showing between breed relationships.

Clades	EC	KZN	LIM	NW	SAB	TWA	GenBank sequences	Total
I	0	3	1	0	7	3	0	14
II	2	0	1	2	0	4	0	9
III	1	0	1	1	7	3	0	13
IV	1	1	0	0	0	0	0	2
V	4	5	6	2	1	2	3	23
VI	0	0	0	1	0	3	0	4
VII	1	0	0	0	2	0	0	3
VIII	1	0	0	0	0	0	0	1
IX	0	0	0	0	0	0	4	4
Total	10	9	9	6	17	15	7	73

The South African populations only shared one clade (Clade V) with the populations from the Japan, Pakistan and India. The Sarda goat from Italy clustered separately in clade IX.

4.3 Discussion

The South African goat breeds were polymorphic for the growth hormone gene with the number of SNPs observed per sequence ranging from 14-142. The growth hormone gene is a highly polymorphic gene especially at exon 1 to 5 as reported in Serrana goat populations (Marques *et al*. 2003). Exon 4 has been reported to be the most polymorphic in previous studies and polymorphisms on this exon and on exon 5

have been associated with high milk yield in the Portuguese Algavia, Indian Jakhrana and Iranian Tali goats (Malveiro *et al*. 2001; Mousavizadeh *et al*. 2009; Gupta *et al*. 2009).

The sub-populations were polymorphic for growth hormone and the level of diversity (27.4-57.9 SNPs) was less than that of Osmanabadi and Sangamneri, Wickramaratne *et al*. (2010). There, 23 SNPs were observed for the Osmanabadi and 12 in the Sangamneri goat populations. The South African Boer population in this study was not as polymorphic as the Chinese Boer goat bucks that were observed to be more polymorphic particularly at exon 2 and exon 4 (Hua *et al*. 2009). Most of the studies performed on the growth hormone gene have previously been conducted on the exons and not on the complete growth hormone gene. These studies have proved that the growth hormone gene is polymorphic at exon/intron level (Gupta *et al*. 2007; Gupta *et al*. 2009), which is similar to observations in the South African goat populations. The polymorphisms in the South African goat population's growth hormone gene suggest that this gene can be used as a candidate gene in selection programs. Similar suggestions were observed by Marques *et al*. (2003) in Ribatejano and Jarmelista ecotypes and Mousavizadeh *et al*. (2009) in Iranian Talli dairy goat breeds.

Single nucleotide polymorphisms on exon 4, positions 1532 and 1585 caused AA changes from asparagine to glycine and arginine to tryptophan in the Savanna and Kalahari Red goats. Mutations resulting in amino acid changes may be associated with some of the economic traits in goats and other livestock species (Marini *et al*. 2012). A base mutation from *A>G* at position 781 was associated with litter size in the Boer and Matou goat breeds in China (Zhang *et al*. 2011). A similar mutation was observed in this study, which resulted in a *Gly>Ser* change. The Eastern Cape population had two amino acid changes of Tyr>Cys and Arg>Gly at exon 5 that were also observed in Black Bengal goats (Gupta *et al.* 2007). Although all the amino acid changes observed were around the exon region, it is important to note that not all the SNPs on exons resulted in AA changes. A similar observation was made by Marini *et al* (2012), of SNPs in exon 3 of the Savanna and Kalahari Red goats that did not result

in any AA changes.

The South African goat populations are relatively diverse and had numerous haplotypes that were either shared amongst or unique to specific populations. In a study on Black Bengal goat populations, 5 haplotypes were observed on exons 4 and 5 (Gupta *et al*. 2007), some of which as in the South African goats were shared amongst sub-populations. The current study as well as the study on the Black Bengal goats can be used as an indication of haplotype diversity of the goat growth hormone gene. Of the haplotypes observed, there were none that were shared by all the populations under study. This could be an indication of comparatively low similarity between the populations. Haplotype GH8 was found only in the Boer goat population further highlighting it's genetic distinction based on the growth hormone gene.

The growth hormone gene was more diverse within the South African Boer goats, as suggested by the highest gene diversity of 0.930 ± 0.005. The study by Hua *et al*. (2009) also reported high genetic diversity of the Chinese Boer goats. Kotze *et al*. (2014), in their first attempt to characterize genetic diversity of feral Tankwa goat concluded that the breed was as genetically diverse as other goat breeds of South Africa. Growth hormone gene diversity of the Tankwa goats was equally considerable with a gene diversity of 0.930 ± 0.006. The village goat populations although less diverse than the Tankwa, with fewer haplotypes (n=2 to n=6) and less gene diversity (0.827 ± 0.041 to 0.889 ± 0.014) as compared to the Boer and Tankwa goats yet this is still potentially sufficient for within breed selection for growth and other traits. Genetic potential can be considered in breeding programs.

The F_{ST} statistics have been used to measure genetic distances and diversity between populations. Genetic distances between pairs of SA village goats from KwaZulu-Natal, Limpopo and Eastern Cape ranged from 0.005 to 0.084, which shows that the populations are very close to each other and may be sharing genetic variants. Simela and Merkel (2008) also reported high genetic diversity and lack of population sub-structuring amongst village goats. Village goats from KZN were however relatively more different from those from the rest of the other provinces with F_{ST} values ranging

from 0.049 to 0.356. The SA Boer goat is a breed that has been specifically bred and developed for meat production (Visser *et* al. 2004), which explains the observed high population divergence from the village and Tankwa goat populations. The Tankwa population is a feral goat from the Northern Cape (Kotze *et al*. 2014) that, like the village goats, has not been artificially selected for growth or other economically important traits. Presumably, the unselected goat population will carry diverse genetic variants resulting in less between population divergences.

Analysis of molecular variance revealed more variation within population than across/among population diversity for all population categories. A high within breed diversity (97.37%) was observed in the category consisting of SA village and Tankwa goats. Pooling all the six populations together revealed that within population variation was still high (73.18%). However, some level of population sub-structuring was observed when all six populations were put together (F_{ST} = 24.15%) as well as in the population category consisting of the Boer and village goats (F_{ST} = 24.42%). Overall, the AMOVA results indicated high divergence of the Boer goats from the other South Africa goat populations. The high within-population variation implies a great potential for within-population selection programs versus crossbreeding in the SA village and Tankwa goat populations.

Whilst the ML phylogenetic tree supported the high genetic diversity and similarity of South African sub populations, it also shed insight on genetic ties South African goats have with other populations from Japan, India, Italy and Pakistan. Of these GenBank sequences, SA goats shared a clade with populations from Japan, India and Pakistan. The Italian Sarda breeds on the other hand clustered differently from the South African goat populations and other international breeds. These results demonstrate the similar genetic backgrounds of goats worldwide and suggests that natural and artificial selection pressure from similar rearing and production systems of the South African goats and those from Asia versus European populations explains some of this genetic homogeneity.

4.4 Conclusion

Overall, the growth hormone gene was polymorphic in the South African goat populations. Mutations that resulted in amino acid changes were observed that could result in alterations of phenotypes. High within, and low between, population diversity was observed in the South African village and feral goat populations. The Boer goat on the other hand had considerable levels of divergence from the village and feral goats. Results from this study suggest that the South African goat populations, especially the SA village, have some potential for good growth with room for within population selection programs.

Acknowledgements

The Agricultural Research Council Biotechnology Platform and the National Research Foundation funded this study. Ms. Ncube was funded through the NRF Innovative Scholarship and ARC-Professional Development Program. University of South Africa.

4.5 References

Alakilli S. Y. M., Mahrous K. F., Salem L. M. & Ahmed E. S. (2012) Genetic polymorphism of five genes associated with growth traits in goat. *African Journal of Biotechnology* 11, 14738-14748.

Campbell Q. P. (2003) The origin and description of Southern Africa's indigenous goats. *South African Animal Science* 4, 18-22.

Crosby L.D. & Criddle C.S. (2007) Gene capture and random amplification for quantitative recovery of homologous genes. *Molecular and cellular probes* 21, 140-147.

Gupta N., Ahlawat S.P.S., Kumar D., Gupta S.C., Pandey A. & Malik G. (2007) Single nucleotide polymorphism in growth hormone gene exon-4 and exon-5 using PCR-SSCP in Black Bengal goats – A prolific meat breed of India. *Meat Science* 76, 658-665.

Gupta N., Pandey A., Malik G. & Gupta S.C. (2009) Single nucleotide polymorphism (SNP) in growth hormone gene of Jakhrana, a prominent milk goat breed in India. *Small Ruminant Research* 81, 35-41.

Gwaze F.R (2009) Communal Production Systems of Goats Raised by Resource-Poor Farmers in the Eastern Cape Province of South Africa. Department of Livestock and Pasture Science, University of Fort Hare, South Africa.

Hua G.H., Chen S.L., Yu J.N., Cai K.L., Wu C.J., Li Q.L., Zhang C.Y., Liang A.X., Han L., Geng L.Y., Shen Z., Xu D.Q. & Yang L.G. (2009) Polymorphism of the growth hormone gene and its association with growth traits in Boer goat bucks. *Meat Science* 81, 391-395.

Kotzé A., Grobler J.P., van Marle-Köster E., Jonker T. & Dalton D.L. (2014) The Tankwa Karoo National Park feral goat population: A unique genetic resource. South African Journal of Animal Science 44, 43-48.

Lahiff E. & Cousins B. (2005) Smallholder Agriculture and Land Reform in South Africa. *IDS Bulletin* 36, 127-131.

Malveiro E., Pereira M., Marques P.X., Santos I.C., Belo C., Renaville R. & Cravador A. (2001) Polymorphisms at the five exons of the growth hormone gene in the algarvia goat: possible association with milk traits. *Small Ruminant Research* 41, 163-170.

Marini A.B., Aslinda K., Mohd. Hifzan R., Muhd. Faisal A.B. & Musaddin K. (2012) HaeIII-RFLP Polymorphism of growth hormone gene in Savanna and Kalahari goats. *Malaysian Society of Animal Production* 15, 13-19.

Marques, P.X., Pereira, M., Marques, M.R., Santos, I.C., Belo, C.C., Renaville, R. & Cravador, A. (2003) Association of milk traits with SSCP polymorphisms at the growth hormone gene in the Serrana goat. *Small Ruminant Research* 50, 177-185.

Masika PJ., Mafu JV., Goqwana MW., Mbutiand C. & Raats J. (2007) A comparison of goat growth performance in a communal and commercial farming system in the Central Eastern Cape Province, South Africa.

Moloko B. (2011) A profile of the South African goat market value chain. *Department of agriculture, fishery and forestry*.

Morrison J. W. (2007) A Guide to the identification of the Natural Indigenous Goats of Southern Africa. *South African Animal Science*, Vol 4.

Mousavizadeh A., Abadi M.M., Torabi A., Nassiry M.R., Ghiasi H. & Koshkoieh A. (2009) Genetic polymorphism at the growth hormone locus in Iranian Talli goats by polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP). *Iranian Journal Of Biotechnology* 7, 51-53.

National Agricultural Marketing Council (2005) Report on the investigation into the potential for the South African goat industry. Draft 1.

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

Pieters A. (2007) Genetic characterization of commercial goat populations in South Africa. Department of Animal Breeding and Genetics, University of Pretoria, Pretoria, South Africa.

Schneider S., Roessli D. & Excoffier L. (2000) ARLEQUIN Ver. 2.0: A Software for Population Genetics Data Analysis Department of Anthropology, University of Geneva, Geneva, Switzerland.

Simela L. & Merkel R. (2008) The contribution of chevon from Africa to global meat production. *Meat Science* 80, 101-109.

SUPAKON C (2006). The important candidate genes in goats – A review. Walailak Journal of Science and Technology 61, 17-36.

Visser C., Hefer C.A., van Marle-Köster E. & Kotze A. (2004) Genetic variation of three commercial and three indigenous goat populations in South Africa. *South African Journal of Animal Science* 34, 24-27.

Webb E. C. & Mamabolo M. J. (2004) Production and reproduction characteristics of South African indigenous goats in communal farming systems. *South African Journal of Animal Science* 34, 236-239.

Weir B. S. & Cockerham C. C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358-1370.

Wickramaratne S.H.G., Ulmek1 B.R., Dixit S.P., Kumar S. & Vyas M.K. (2010) Use of growth hormone gene polymorphism in selecting *Osmanabadi* and *Sangamneri* goats. *Tropical Agricultural Research* 21, 398 - 411.

Zhang C., Liu Y, Huang K., Zeng W., Xu D., Wen Q. & Yang Liguo. (2011) The association of two single nucleotide polymorphisms (SNPs) in growth hormone (GH) gene with litter size and superovulation response in goat-breeds. *Genetics and Molecular Biology* 34, 49-5.

CHAPTER 5

5.1 General Discussion

South Africa is one of the major goat producing countries with some of its goat populations belonging to well-characterized meat breeds, which are used by commercial farmers nationally and globally. In addition to these are the undeveloped and uncharacterized village populations and a Tankwa feral goat breed. The history of South African goats is still very vague, despite the size and diversity of the industry (Campbell, 2013). Although South Africa has high performing, fast growing goats such as the Boer, Kalahari and Savanna goats, the majority of these goats are non-descript village ecotype populations, which have various body sizes and are raised under marginalized production systems. There is very limited genetic information on South African goat populations in general. This lack of information on breed history, genetic background and genetic potential hinders efforts to develop, improve and conserve breeds and populations.

This study sought to investigate the genetic diversity and structure as well as maternal lineages of the South African goat populations to bring about knowledge and understanding on the genetic background of these goats as well as show how that diversity is distributed amongst commercial, feral and village goat populations. The study also analyzed the growth hormone gene within and between populations to investigate the potential of this gene as a candidate gene for selection and breeding as well as to get insight on the genetic potential of village goats for growth performance. The driving force behind this study was that there is limited information on the South African goat populations, with the origins and history of these goats greatly unknown. Knowledge of the maternal lineages is important in understanding the history of breeds, which influences its production performance. The mitochondrial DNA has a rapid evolutionary rate (Garrine 2007) and was considered the best gene for within species genetic studies (Garrine 2007). Village goats have been reported to have poor growth performance, which is speculated that it may be due to both poor management practices as well as genetic background (Masika *et al*. 2008).

Reports such as those of Simela and Merkel (2008) have however suggested the presence of high genetic variation in South African village goat populations to allow

for within population selection and breed improvement programs. The present study therefore sought to unravel that genetic variation using the mtDNA and growth hormone gene. The populations selected in this study were a representation of the production systems observed in South African goat farming and included (i) the SA Boer goat representing intensive production system practiced mostly by commercial farmers, (ii) the Tankwa goat representing captive feral system in the Northern Cape Karoo where it is kept under conservation and (iii) the SA village goats representing the extensive production system practiced mostly by communal farmers in rural villages across South Africa.

Goat domestication is a process that has been taking place over thousands of years ago and is believed to have involved three wild goat species namely (i) *Capra aegagrus* (bezoar), ii) *Capra falconeri* (markhors) and iii) *Capra ibex* (ibex) (Odahara *et al*. 2006). The results in this study suggested multiple and diverse maternal lineages within the South African goat populations. The results support mostly the presence of lineage haplogroup A and B in these populations. There was also evidence of novel and unique maternal lineages within South African goat populations observed as haplotypes that did not cluster with any of the GenBank lineages. There were within South Africa no breed specific clades implying that the multiple maternal lineages were shared between and within populations.

The Tankwa feral goat population though reported to be genetically unique (Mdladla *et al*. 2016), share maternal lineages with the farmed goats. The findings by Mdladla et al. (2016) raised expectations that the Tankwa would form mtDNA clade that will be unique and separated from the other South African goat populations. However, in this study, the Tankwa clustered with and around haplotypes from other South African farmed animals suggesting that these populations shared maternal lineages despite the genetic uniqueness presented by the autosomal SNPs.

The complete growth hormone gene was polymorphic in the South African goat populations. A high number of SNPs (142) were observed in this study, which is higher than that reported for populations in literature. High gene diversity was

observed in all the populations under study (Table 4.4) where the SAB and the TWA populations had the highest gene diversity. The Boer goat has been bred and developed for high growth performance, high meat yields and fast growth, which probably explain why the growth hormone gene was highly diverse in this population. The TWA population, although not an artificially developed breed, has a big body size compared to other farmed goat populations suggesting it is naturally selected for high growth. High within population diversity, and the absence of population substructure across SA village goat populations, implied that they are one population existing in different geographical locations. This has implications on the naming and conservation of the local indigenous goat populations. The results also showed that the SAB population is different from the other village and feral goat populations, which was expected since the Boer goat was specifically selected for growth and could have diverged from other population on growth related genes such as growth hormone.

Phylogenetic analysis of this gene produced nine clades (Fig 4.2), which were distributed across all populations also suggesting that the different goat populations shared some genetic background that presumably goes back to the development of these breeds. The phylogeny also clustered the South African goat populations with the Indian, Pakistan and Japanese goat populations implying that they either shared the same genetic background and or could have undergo the same selection pressures. The Italian population however formed its own clade separated from all the other populations further suggesting that this is a different breed and does not share any genetic information with the South African populations

5.2 Conclusion

The SA goats belong to a diverse genetic pool with multiple and diverse maternal lineages shared across the different breeds or populations. The presence of lineage A and B in South African goat populations suggest that the South African goats share maternal lineages with the Chinese goat populations. MtDNA results also indicated that the wild bezoar and the ibex have genetically influenced the South African goat populations, further suggesting that these two wild populations may be the ancestors of the South African goat populations. Lineage A from Tibet and elsewhere also

observed to cluster with the South African goat populations implying some genetic similarity between the populations.

Overall, the growth hormone gene was polymorphic in the South African goat populations. As with mtDNA, high within, and low between, population diversity was observed in the South African village and feral goat populations. Results from this study suggest potential for good growth and room for within population selection programs in the South African goat populations, particularly the uncharacterized and undeveloped village goats.

This study provided baseline information on the genetic structure of SA goat populations and will benefit from additional studies particularly on genetic association analysis.

APPENDICES

Table A1 A list of primers used for the rolling circle amplification of complete goat mtDNA gene of South African goat populations.

Primer name	**5' to 3' sequence**
Goat F1	CCACCGCGGTCATAC
Goat R1	AAGCACCGCCAAGTC
Goat F2	GGCCGCGGTATTCTG
Goat R2	GGTCTTCTCGTCTTG
Goat F3	AACCACAACCACCT
Goat R3	AGGAGGGTGATGAGG
Goat F4	ACCCAGCAGGAGGAG
Goat R4	ACTTCAGGGTGTCCA
Goat F5	AACCCATCCCTCACA
Goat R5	ACCTCTAGCAGTCGT
Goat F6	ACTCCAGCCGTTCAA
Goat R6	AGCAGCCGCCTAATT
Goat F7	TTCGGCGATAACAGC
Goat R7	AGGGAGTCGGAGAAGA
Goat F8	ACCGCACCCATCATA
Goat R8	TGGGTAGTTGGTGGA
Goat F9	CCGCCCTAGCAGTTA
Goat R9	AGCGGTTTCTAGTGC
Goat F10	TCTTCGCCTTCCACT
Goat R10	CGAGGGCTGTGATGA

Goat F11	CGTATCCCGTCCACT
Goat R11	GGGTTGCTGGTTTCA

Table A2 GenBank Sequences for D-loop and COX1 for comparison with South African goat populations.

Haplotype/ lineage	**Breed**	**Accession no.**	**Gene region**	**Location**	**Reference**
A	Langkazi	GU229278	Whole gene	Tibet	Wu *et al* 2012
B	Yunling Black	GU229279	Whole gene	Yunnan	Wu *et al* 2012
C	Wuzhumuqin White Cashmere	GU229280	Whole gene	Neimeng	Wu *et al* 2012
D	Tsaidam Cashmere	GU229281	Whole gene	Qinghai	Wu *et al* 2012
	Capra ibex, alpine ibex	FJ207526	Whole gene	European Mountains	Hassanin *et al*. 2006
	Capra nubiana, Nubian	FJ207527	Whole gene	Deserts of north-eastern Africa and	Hassanin *et al*. 2006

	ibex			Arabia	
	Hemitragus jemlahicus, Himalayan tahr*	FJ207531	Whole gene	Asia (only)	Hassanin *et al*. 2006
	Boer	KM233163.1	Whole gene	China	Niu *et al*. 2014
H1; LA	Huanghuai goat	DQ121560.1	D-loop	China	Liu *et al*. 2007
	Matou goat	DQ121580.1	D-loop	China	Liu *et al*. 2007
	Liaoning Cashmere goat	DQ188870.1	D-loop	China	Liu *et al*. 2007
	Neimonggol Cashmere	DQ188880.1	D-loop	China	Liu *et al*. 2007
H2;LB	Banjiao White Goat	DQ121491.1	D-loop	China	Liu *et al*. 2007
	Guizhou White Goat	DQ121540.1	D-loop	China	Liu *et al*. 2007
	Leizhou goat	DQ121570.1	D-loop	China	Liu *et al*. 2007
	Shaannan White goat	DQ121610.1	D-loop	China	Liu *et al*. 2007
	Bangladeshi Black	AB736098.1	COX1		Nomura *et al* 2013

Bengal			
Wild Bezoar	AB736122.1	COX1	Nomura *et al* 2013
Mongolian	AB736109.1	COX1	Nomura *et al* 2013

Table A3 Pairwise F_{ST} P-values and Standard Error. All values were significant at P ≤ 0.04.

***Population**	**EC**	**KZN**	**L**	**NW**	**SAB**	**TWA**
EC						
KZN	0.030±0.060					
L	0.045±0.024	0.012±0.042				
NW	0.054±0.010	0.051±0.003	0.046±0.025			
SAB	0.050±0.007	0.036±0.050	0.022±0.040	0.010±0.025		
TWA	0.021±0.031	0.045±0.016	0.037±0.035	0.031±0.029	0.021±0.038	

Table A4 SNPs, Insertions, Deletions, Replacements and Amino Acid changes.

SAMPLE ID	SNPs	INS	DEL	REPLACEMENTS	No. AAs	TYPE OF AA CHANGE
EC01	83	6	0	7	2	Try>Cys; Arg>Gly
EC02	53	2	1	0	1	Gly>Ser
EC03	23	1	1	0	0	None
EC04	73	4	0	1	1	Gly>Ser
EC05	23	1	0	0	0	None
EC06	33	3	1	0	0	None
EC07	19	1	0	0	0	None
EC08	53	1	2	1	0	None
EC09	35	1	0	0	0	None
EC10	129	3	1	1	1	Gly>Ser
KZN01	18	1	0	0	0	None
KZN02	26	1	1	0	0	None
KZN03	26	1	0	0	1	Gly>Ser
KZN04	17	1	0	0	0	None
KZN06	50	1	0	1	0	None
KZN07	52	1	0	0	0	None
KZN08	51	1	1	1	0	None
KZN09	70	3	0	2	0	None
L01	69	4	1	3	1	Gly>Ser
L02	24	2	0	0	0	None
L03	28	1	1	0	1	Gly>Ser

L04	75	4	2	5	0	None
L05	66	1	1	0	1	Gly>Ser
L06	65	3	0	1	0	None
L07	142	3	2	12	0	None
L08	114	3	3	3	0	None
L09	137	2	1	4	0	None
L10	75	2	1	0	0	None
NWI04	75	3	0	1	1	Gly>Ser
NW06	93	2	1	1	0	None
NW07	57	1	1	0	0	None
NW08	33	1	0	0	0	None
NW09	65	1	1	2	1	Gly>Ser
NW10	90	4	2	11	0	None
SAB01	83	2	1	0	0	None
SAB02	46	2	1	0	0	None
SAB03	115	4	3	1	0	None
SAB04	46	2	1	0	0	None
SAB05	14	1	0	0	0	None
SAB09	27	1	0	0	0	None
SAB10	45	2	1	0	0	None
SAB11	92	2	0	0	0	None
SAB12	87	4	3	3	0	None
SAB13	59	2	1	0	0	None
SAB14	27	1	0	0	0	None

SAB15	47	2	3	3	0	None
SAB16	69	4	2	1	0	None
SAB17	36	2	1	0	0	None
SAB18	57	2	1	0	0	None
SAB19	43	2	1	0	0	None
SAB20	47	2	1	0	0	None
TWA01	65	3	1	1	1	Gly>Ser
TWA02	30	1	0	1	0	None
TWA03	46	2	2	0	1	Gly>Ser
TWA04	29	2	1	2	0	None
TWA05	36	3	2	2	1	Gly>Ser
TWA11	44	1	0	0	1	Gly>Ser
TWA12	102	2	1	3	0	None
TWA13	41	1	0	0	1	Gly>Ser
TWA14	51	2	0	0	1	Gly>Ser
TWA15	61	3	1	0	0	None
TWA16	98	2	1	4	0	None
TWA17	39	10	0	0	0	None
TWA18	122	2	1	9	0	None
TWA19	71	2	0	0	0	None
TWA20	44	2	2	2	1	Gly>Ser

Table A5 A representation of the 31 haplotypes that were unique (not shared within and between populations).

Haplotype	EC	KZN	L	NW	SAB	TWA	Total
GH11	1	0	0	0	0	0	1
GH12	1	0	0	0	0	0	1
GH13	1	0	0	0	0	0	1
GH14	1	0	0	0	0	0	1
GH15	1	0	0	0	0	0	1
GH16	0	1	0	0	0	0	1
GH17	0	1	0	0	0	0	1
GH18	0	1	0	0	0	0	1
GH19	0	0	1	0	0	0	1
GH20	0	0	1	0	0	0	1
GH21	0	0	1	0	0	0	1
GH22	0	0	1	0	0	0	1
GH23	0	0	1	0	0	0	1
GH24	0	0	0	1	0	0	1
GH25	0	0	0	1	0	0	1
GH26	0	0	0	1	0	0	1
GH27	0	0	0	1	0	0	1
GH28	0	0	0	0	1	0	1
GH29	0	0	0	0	1	0	1
GH30	0	0	0	0	1	0	1

GH31	0	0	0	0	1	0	1
GH32	0	0	0	0	1	0	1
GH33	0	0	0	0	1	0	1
GH34	0	0	0	0	1	0	1
GH35	0	0	0	0	0	1	1
GH36	0	0	0	0	0	1	1
GH37	0	0	0	0	0	1	1
GH38	0	0	0	0	0	1	1
GH39	0	0	0	0	0	1	1
GH40	0	0	0	0	0	1	1
GH41	0	0	0	0	0	1	1
Total	**5**	**3**	**5**	**4**	**7**	**7**	**31**

Table A6 Pairwise F_{ST} P-values and Standard Error. All values were significant at P ≤ 0.04.

***Populatio n**	**EC**	**KZN**	**L**	**NW**	**SAB**	**TWA**
EC						
KZN	0.010±0.010					
L	0.0441±0.024	0.012±0.042				

NW	0.0541±0.020	0.019±0.003	0.036±0.015		
SAB	0.010±0.00	0.036±0.050	0.021±0.041	0.010±0.025	
TWA	0.042±0.061	0.045±0.015	0.0360±0.015	0.027±0.019	0.051±0.038