

GENETIC DIVERSITY AND RELATIONSHIPS OF AFRICAN SHEEP:

A Y CHROMOSOME PERSPECTIVE

EGERTON UNIVERSITY LIBRARY

OUNA BENARD ASWANI

**A thesis submitted to the Graduate School in partial fulfillment of the requirements for
the award of the Degree of Master of Science in Biochemistry of Egerton University**

EGERTON UNIVERSITY

January, 2007



Eger234660



2007/72630 X

DECLARATION AND RECOMMENDATION

Declaration

I hereby declare that this thesis is my original work and has not been presented for any degree award in any institution or University

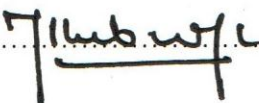
Date..02/02/2007..... Sign.. 

Name: Ouna Bernard Aswani

Reg No: SM14/0644/02

Recommendation

This thesis has been submitted with our approval as supervisors:

Date..5th February 2007..... Sign.. 

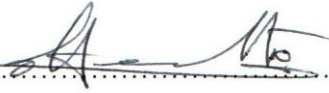
Dr. David Mburu

International Livestock Research Institute, Nairobi

Date..5th February 2007..... Sign.. 

Prof. Moses Limo

Department of Biochemistry and Molecular Biology, Egerton University

Date..2 February 2007..... Sign.. 

Dr. Olivier Hanotte

International Livestock Research Institute, Nairobi

DEDICATION

I dedicate this thesis to the Ouna family for their quest for education.

ACKNOWLEDGEMENTS

This study would not have been successful without financial, technical and moral support from several institutions, groups and individuals. My greatest and sincere appreciation goes to my supervisor Dr. David Mburu, for his technical guidance, supervision and encouragement that formed the basis for my understanding of Molecular Genetics. My special thanks also go to Dr. Olivier Hanotte for giving me a chance, facilitating and guiding me through this study. I also appreciate the supervisory effort made by Prof. Moses Limo, for his inspiration, recommendation and encouragement to pursue Molecular Biology. Dr. Han Jianlin's input in data analysis and brain storming on the way forward for this project is also highly appreciated. I am grateful to Dr. James Kijas and Dr. Jennifer Meadows of CSIRO industries for providing vital information on sequences and restriction enzyme.

ILRI provided financial, logistical support and enabling environment for the success of this project and I am very grateful for this. I am also grateful for the academic and administrative guidance given to me from Egerton University, particularly the Department of Biochemistry and Molecular Biology and the Graduate School.

I am indebted to the following individuals and groups of people; Harrison Kibogo, your close technical support and encouragement helped me to cope with the laboratory challenges. I am also grateful to any kind of support I received from my fellow biodiversity colleagues; Viranga, Jacky, Sheila, Victor, Emily, Allan, Lillian, Dorine and Grace. The assistance of all Lab 7 administrative and technical staff is also highly appreciated. To all my friends and colleagues; Mapesa, Nyabuga, Amwayi, Bonny, Grace, Muge, Gitau, Mala, Leah, Wasike, Immaculate, Kahindi and all my classmates at Egerton University, thank you.

I wish to thank the Ouna family for having encouraged, facilitated and tolerated this undertaking, the Omulama family as well as the Muyeras' for their hospitality during this period. Finally to the Almighty God, thank you for blessing me with life and education.

ABSTRACT

In livestock, evolutionary or population genetics studies at the male lineage are uncommon due to lack of informative markers on the Y chromosome. Recently, in sheep, two Y chromosome specific markers have been developed. A bi-allelic A/G single nucleotide polymorphism (SNP) and a multi-allelic microsatellite *SRYM18* marker. These two markers were used to analyze the Y chromosome diversity and population structure of 447 thin-tailed and fat-tailed African and 44 non-African male sheep from 35 distinct breeds. A total of five microsatellite alleles (131 bp, 139 bp, 141 bp, 143 bp and 145 bp) were observed. At *SRYM18*, the highest diversity was found in two thin-tailed sheep breeds from West Africa, Djallonke from Senegal and Maure of Mali, each with three alleles (139 bp, 141 bp and 143 bp). All the other breeds had at least two different alleles, with the exception of three fat-tailed sheep breeds, Gumuz, Sekota and Tukur, from Ethiopia that had only one allele of 143 bp. Continent-wide, the 143 bp allele was the most common, with a frequency of 72.9%, and it was distributed in all geographical regions. The 141 bp allele was observed in both the thin-tailed and fat-tailed sheep with continent-wide frequency of 19.7% while the 145 bp allele was only found in the fat-tailed sheep though at low frequency (8.8%). The 131 bp and 139 bp alleles are rare on the continent with their respective continent-wide frequencies being less than 1%. SNP screening detected only the A allele. A combination of results from the SNP and the microsatellite marker confirmed the presence of three haplotypes: 141/A; 143/A and 145/A with a respective frequency of 16.7%, 72.5% and 10.8% in the African breeds. Haplotype 143/A had the highest distribution in the East and southern Africa regions with a frequency of 90.5% and 60.0%, respectively, while in West Africa haplotype 141/A was the commonest with a frequency of 87.5%. These results are in agreement with archeological information suggesting a distinct origin for the West African thin-tailed and the fat-tailed sheep from East and southern Africa.

TABLE OF CONTENTS

DECLARATION AND RECOMMENDATION	ii
DEDICATION	iii
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
TABLE OF CONTENTS	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
ABBREVIATIONS	xi
 CHAPTER ONE	 1
INTRODUCTION	1
1.1 Background information	1
1.2 Statement of the problem	2
1.3 Objectives	2
1.3.1 Main objective	2
1.3.2 Specific objectives	3
1.4 Null hypothesis	3
1.5 Justification	3
 CHAPTER TWO	 4
LITERATURE REVIEW	4
2.1 Economic importance of the domestic sheep	4
2.2 Origin of the domestic sheep	4
2.3 Migration and distribution of the African sheep breeds	4
2.4 African sheep breed development	10
2.5 Threat to livestock biodiversity	10
2.6 Molecular markers commonly used for genetic characterization	11
2.6.1 Restriction fragment length polymorphism	11
2.6.2 Random amplified polymorphic DNA	11
2.6.3 Microsatellites	12
2.6.4 Single nucleotide polymorphism	12

2.7 Mitochondrial DNA 13

2.8 Y chromosome 13

2.8.1 Evolution of the Y chromosome 13

2.8.2 The structure of Y chromosome 14

2.8.3 Use of Y chromosome markers in livestock diversity studies 16

CHAPTER THREE 18

MATERIALS AND METHODS..... 18

3.1 DNA samples 18

3.2 DNA extraction..... 18

3.3 DNA quantification..... 20

3.4 Polymerase chain reaction 20

3.5 Agarose gel electrophoresis 21

3.6 DNA genotyping..... 22

3.7 Genotype analysis 22

3.8 Statistical analyses 22

3.9 *SRYM18* allele sequencing..... 23

3.10 Sequence Analysis 24

3.11 Single nucleotide polymorphism analysis 24

CHAPTER FOUR..... 26

RESULTS AND DISCUSSION 26

4.1 RESULTS 26

4.1.1 PCR optimization..... 26

4.1.2 Genotyping..... 27

4.1.3 Allele frequencies 30

4.1.4 *SRYM18* allele sequence analysis 40

4.1.5 Single nucleotide polymorphism analysis 43

4.1.6 Haplotype..... 44

4.2 DISCUSSION 45

4.2.1 SRY marker specificity..... 45

4.2.2 *SRYM18* microsatellite mutations..... 46

4.2.3 Y chromosome diversity within and between African sheep breeds..... 47

4.2.4 Origin, geographical distribution and migration of the African sheep 49

4.2.5 Sheep Y chromosome SNP 51

4.2.6 Sheep Y chromosome haplotypes 51

4.2.7 Usefulness of the Y markers in sheep diversity studies..... 52

CHAPTER FIVE 54

CONCLUSIONS AND RECOMMENDATIONS 54

5.1 Conclusions..... 54

5.2 Recommendations..... 55

REFERENCES 57

LIST OF TABLES

Table 1. Sheep samples used for microsatellite genotyping and SNP screening.....	19
Table 2. Microsatellite allele frequencies..	31
Table 3. Microsatellite allele frequencies of sheep breeds classified by type and region.....	37
Table 4. Logistic regression analysis of <i>SRYM18</i> microsatellite data.....	39
Table 5. Analysis of molecular variance within and between breeds.	40
Table 6. Microsatellites repeat units.	41
Table 7. Details of the breeds in which the A/G SNP was successfully screened.....	43
Table 8. Distribution of haplotypes in sheep breeds grouped by geographical regions.	45

LIST OF FIGURES

Figure 1. Photographs of the fat-tailed and thin-tailed sheep.	7
Figure 2. Migration routes of African sheep.....	9
Figure 3. Ideogram of the Y chromosome.	15
Figure 4. <i>SRYM18</i> PCR optimization gel photograph.	26
Figure 5. <i>SRY3</i> PCR optimization gel photograph.....	27
Figure 6. Electropherograms of the <i>SRYM18</i> microsatellite alleles.....	28
Figure 7. Allele distribution on the physical map of Africa	34
Figure 8. Origin and migration pattern of the <i>SRYM18</i> alleles in Africa.....	36
Figure 9. <i>SRYM18</i> allele multiple sequence alignment	42

ABBREVIATIONS

ABI	Applied biosystems incorporation
AGE	Agarose gel electrophoresis
bp	Base pair
<i>DAZ</i>	Deleted azoospermia
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FAO	Food and Agriculture Organization
<i>MSY</i>	Male specific Y
Myr	Million years
PAGE	Polyacrylamide gel electrophoresis
<i>SRY</i>	Sex determining region Y
TBE	Tris Borate EDTA buffer pH 8.3
U	Units
UV	Ultraviolet

CHAPTER ONE

INTRODUCTION

1.1 Background information

Since their domestication around 7000 to 9000 Before Christ (BC), sheep have significantly influenced the economic, social and cultural foundation of human society. In sub-Saharan Africa it is estimated that there are over 240 million sheep (FAO, 2004b), which are broadly classified as thin-tailed or fat-tailed (Mason & Maule, 1960). About 80% of the African sheep are local breeds that are maintained under traditional farming systems (Lebbie *et al.* 1996). Consequently, these breeds have become adapted to a wide range of environments, showing high levels of phenotypic variability and increased fitness under natural conditions. These naturally evolved genetic characteristics provides sustainable options for disease resistance, survival and efficient production that has been ignored in the drive for technological and management solutions to individual problems of livestock production in the tropics. For example, in most livestock breeding programs a small number of males with 'good characteristics' are normally used. If unchecked this could lead to inbreeding and therefore reduce diversity in subsequent generations (Bell, 2003). According to Rege *et al.* (1996), 30% of the indigenous animal genetic resource of Africa is at high risk of loss. An essential step in any strategy for the conservation of animal genetic resources involves documentation of the available livestock and measurement of their diversity and genetic relationships.

With the advent of molecular techniques, an increasing number of studies have focused on the genetic characterization of domestic breeds through molecular markers. These markers include: restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), microsatellites and single nucleotide polymorphisms (SNPs). These markers have provided useful information about population structure, gene flow, phylogenetic relationships, parentage testing and breed assignment. Studies into the genetic diversity of sheep have to date relied exclusively on autosomal and mitochondrial DNA markers. In a recent continent wide study, Muigai (2003) using a set of fourteen autosomal markers, concluded that there are two main groups of sheep on the African continent. The first group consists of the fat-tailed and fat-rumped sheep while the second

comprises of the thin-tailed sheep. This study further revealed that Africa is hosting a very large and diverse sheep genetic resource and it also supported the archeological theories that East and southern Africa fat-tailed sheep most likely entered the continent through the Horn of Africa, while the North and West African sheep most likely entered the continent through the Isthmus of Suez.

Unlike autosomal or X-specific loci, both mtDNA and the male-specific region of the Y chromosome do not recombine with autosomes or X chromosome. Therefore, polymorphisms on a mitochondrial molecule or on a Y chromosome share the history of a single female or a single male lineage, respectively. The use of Y chromosome haplotyping in natural populations of species, other than humans, is still hindered by the lack of Y chromosome markers and sequence information. In cattle, the analysis of Y chromosome has proven to be highly insightful for unraveling the processes of domestication and breed development (Bradley *et al.* 1996; Hanotte *et al.* 2000; Kikkawa *et al.* 2003). Recently in sheep, a novel bi-allelic SNP located in the *MSY* region (Meadows *et al.* 2004) and a polymorphic multi-allelic microsatellite *SRYM18* (Australian Sheep Gene Mapping Website, 2003) has been described in sheep. Together, these markers can be used to construct haplotypes to define individual male lineages and assess polymorphism at the ovine Y chromosome. In the present study, the two Y specific DNA markers were used to define Y chromosome diversity and relationships of African male sheep.

1.2 Statement of the problem

There is little information about genetic diversity and relationships of the African sheep. To be able to generate such information, there is need to analyze nucleotide diversity on the Y chromosome of the African sheep.

1.3 Objectives

1.3.1 Main objective

To determine the genetic diversity and relationships of the Y chromosome of the African fat-tailed and thin-tailed sheep breeds.

1.3.2 Specific objectives

- 1) To evaluate the Y chromosome diversity and relationships within and between the African sheep breeds.
- 2) To determine the usefulness of Y chromosome markers in tracking the history and differences between sheep breeds.

1.4 Null hypothesis

There is no significant nucleotide diversity on the Y chromosome that can be used to evaluate diversity status and track the history of various breeds of the African sheep.

1.5 Justification

Domestic animal biodiversity is essential for food security and for future challenges such as re-emergence of livestock diseases. The use of a few selected male breeds has been blamed for loss of genetic diversity in various livestock breeds. Y chromosome nucleotide diversity can provide important information about the history and diversity status of the African male sheep for rational utilization and conservation.

CHAPTER TWO

LITERATURE REVIEW

2.1 Economic importance of the domestic sheep

Domestic sheep are an important resource to many small holders farming systems in Africa. They provide meat, fat, blood, and milk for food; bone and horn for implements; skins and wool for clothing; gut for containers and manure for crop production. In addition, sheep have cultural and religious values among many African communities. Compared to other livestock, sheep appear to have a greater potential for survival in harsh environments unsuitable for other purposes ranging from sub-polar regions to tropical deserts (McCorkle, 1999).

2.2 Origin of the domestic sheep

Sheep domestication is thought to have started in the present day Iran and Turkey (Hiendleder *et al.* 2002; Pedrosa *et al.* 2005). The origin of the modern domestic sheep (*Ovis aries*) remains uncertain. Existing wild sheep includes urial (*Ovis vignei*, disputed scientific classification), mouflon (*Ovis musimon*), argali (*Ovis ammon*), bighorn (*Ovis canadensis*), thinhorn (*Ovis dalli*) and Barbary sheep (*Ammotragus lervia*). Of these the only presumptive wild ancestors of the domestic sheep are urial, mouflon and argali sheep (Ryder, 1984). Using mitochondrial DNA, Hiendleder *et al.* (1998a), suggested two ancestral maternal sources for *Ovis aries*. The first ancestor is shared by the modern domestic sheep and European mouflon, while the second ancestor is yet to be identified. This study also further suggested that European mouflon could have been derived from an early-domesticated sheep other than urial and argali sheep. This view has received more support (Hiendleder *et al.* 2002), but in addition a third maternal lineage that suggests multiple independent domestication events has also been reported in the Turkish sheep (Pedrosa *et al.* 2005).

2.3 Migration and distribution of the African sheep breeds

The sheep breeds of Africa are broadly classified into two main types, thin-tailed and fat-tailed (Figure 1). These sheep are not indigenous to the continent. They were likely domesticated in Asia (MacDonald, 2000) and entered into the African continent through two entry points. This

theory of sheep entry into the continent through two separate entry points initially based on archeological and breed observation information (Figure 2), has received further support from recent findings using molecular markers (Muigai, 2003). As illustrated by paintings in Egyptian tombs dating back to 4,000 BC, the thin-tailed hairy sheep were likely the first type of sheep to enter the African continent from the center of domestication, within the Fertile Crescent in western Asia (which encompasses the current Palestine, Lebanon, southern Turkey, Iran and Iraq). They entered the African continent possibly through the Isthmus of Suez into the present Egypt or the South extremity of the Sinai Peninsula. They then spread westward to West Africa, southward along River Nile into the present Sudan (Wilson, 1991; Muigai, 2003). A still disputed route is the movement of thin-tailed sheep southward from the western side of the continent following Bantu migration along the coastal area.

The sheep breeds of the hairy thin-tailed type include, Djallonke, Uda, Maure and Touareg. Djallonke also known as West African Dwarf sheep is a trypanotolerant sheep found in most West African countries including Senegal, Chad, Cameroon Gabon, Congo, Togo and Nigeria. Uda has a black and white pied coat colour. It is found in several countries including Niger, Nigeria, Chad, Cameroon, and Sudan where different colour variants of this breed are referred to Yankassa, Balami, Bali-bali, Bororo and Fellata. Maure probably originated from the Maghreb sheep of Morocco. Their coat colour is either black with white tip to the tail or white with black spots. They have a convex nose and exceptionally long legs for long distance walking. They are found in Mauritania, Mali, Niger and Senegal where they are also referred to as Moor, Mauritania and Arab. Touareg has white or red coat colour with some spotting. It is also long legged and has a prominent forehead. This sheep breed is found in Mali and Niger where it is also referred to as Ara-ara (DAGRIS, 2004). The thin-tailed coarse wool sheep is believed to have evolved from the thin-tailed hair sheep. They mainly inhabit North African countries including Algeria, Morocco, Tunisia and Libya where they are referred to Algeria Arab, Arrit and Ben Guil. They are characterized by a long, thin tail, heavy head, convex face profile and pendulous ears (Manson, 1967; DAGRIS, 2004). Some of them withstand extremes of temperature, draught and survive on poor nutrition.

The fat-tailed sheep may have originated from a second center of domestication in eastern Asia. They supposedly migrated into the Fertile Crescent after the first migration of the thin-tailed sheep had occurred (Muigai, 2003). The fat-tailed sheep likely entered the continent first,

through the Isthmus of Suez into the current Egypt at the beginning of the second millennium and later through the Horn of Africa into current Ethiopia. The fat-tailed sheep that entered Egypt spread westward into the current Libya, Tunisia and Algeria replacing some of the thin-tailed sheep of North Africa. These sheep breeds include the fat-tailed coarse wool Ossimi, Barbary, Barki of Egypt, Libya and Tunisia respectively (Manson, 1967; DAGRIS, 2004).

(a)



(b)

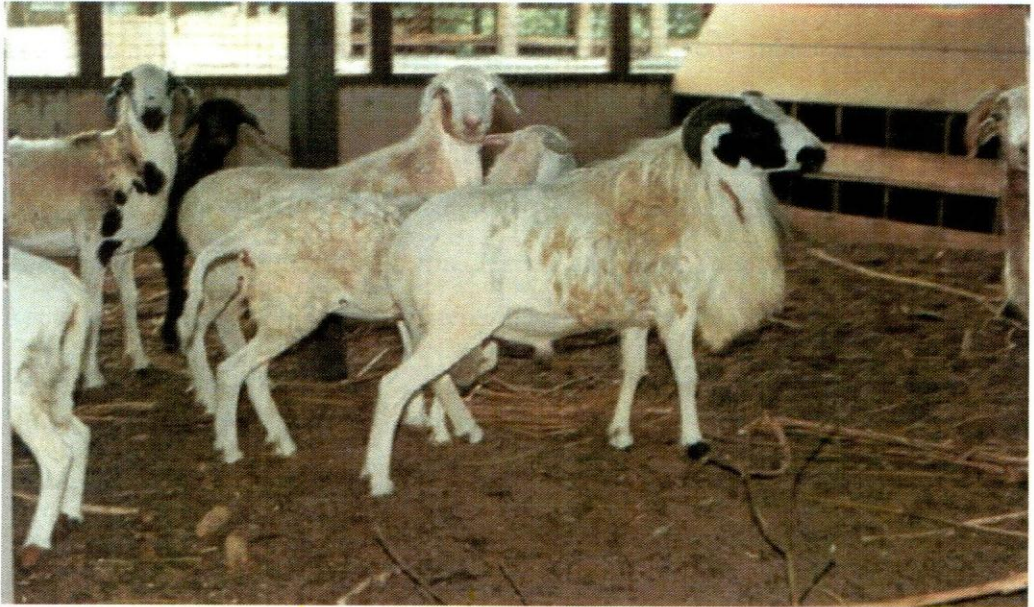


Figure 1. Photographs of the fat-tailed and thin-tailed sheep a) The fat-tailed Horro sheep which is an inhabitant of eastern and southern Africa b) A thin-tailed Djallonke sheep from West Africa (DAGRIS, 2004).

The fat-tailed sheep that entered through the Horn of Africa spread from Ethiopia to the whole of East Africa, southern Africa and to some extent the Sahelian belt. They include the fat-tailed hair sheep, fat-tailed coarse wool, fat-rumped hair and fat-tailed fur. The fat-tailed hair sheep vary in the tail form and extend of the woolly undercoat. They are found in the whole of East Africa and southern Africa where they have different breed names. They include Sukuma, Ugogo, Red Maasai, Afar, Horro, Damara, Pedi, Sabi and Tswana (Wilson, 1991; DAGRIS, 2004). The fat-tailed coarse wool probably descended from an ancient importation from Arabia. Their bodies are covered with a coarse open fleece of carpet wool. They have a heavy tail, pendulous ears and long slender legs. They are found in both the East Africa and southern Africa. These sheep breeds include Arsi-bale, Tukur, Ethiopian Highland, Menz, Wollo Somali-Arab of Ethiopia and Somalia respectively (Mason & Maule, 1960; Wilson, 1991; DAGRIS, 2004). Fat-rumped hair sheep have a short fat-rumped tail. They may have originated from the Blackhead of Somalia through crossbreeding of the fat-tailed and thin-tailed sheep. This breed was introduced into South Africa in 1869 by sailors. These breeds inhabits the dry areas of Kenya, Uganda, Ethiopia Somalia, Sudan and South Africa where they are referred to Blackhead Somalia, Turkana, Boran, Gabra, Moroto, Kabale and Black head Persian. Given their origin, the use of the name Black head Persian is a misnomer (Wilson, 1991; DAGRIS, 2004). The fat-tailed fur sheep probably originated from Syria, Jordan and Mesopotamia. They were imported into South Africa in 1907 from Germany where they had been introduced. Their bodies are covered with brown or gray fleece. The major populations of this breed commonly known as Karakul are found in South Africa, Namibia and Botswana with small numbers in Angola and Rwanda (Wilson, 1991; DAGRIS, 2004). The fine wool sheep were developed from selected local breeds and exotic breeds. Most of these breeds are mainly found in South Africa. They include Afrino, Dohne Merino and South African Merino (Rege *et al.* 1996; DAGRIS, 2004).

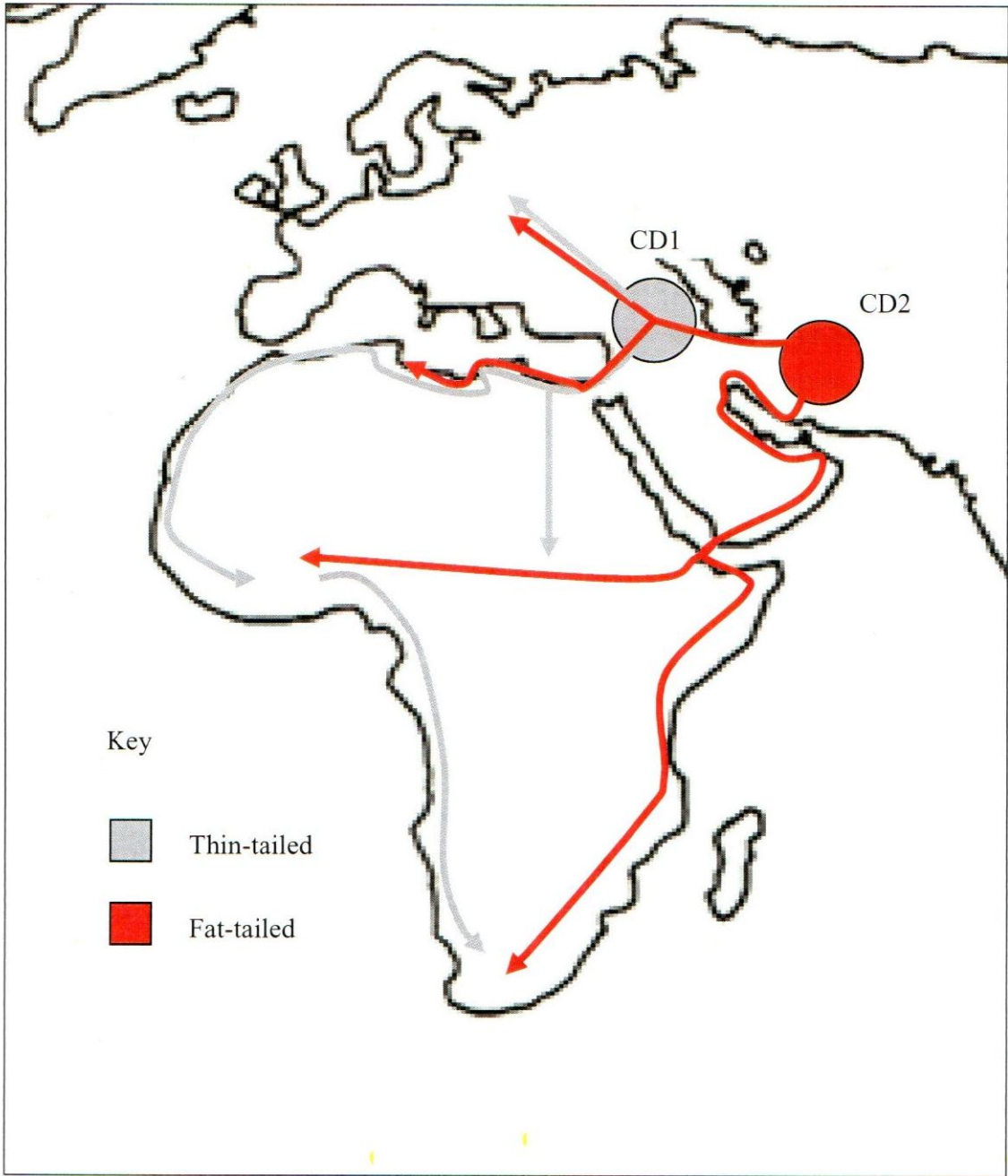


Figure 2. Migration routes of African sheep. Centers of domestication are labeled CD1 and CD2. The migration routes are coloured gray and red for the thin-tailed and fat-tailed sheep respectively (Muigai, 2003).

2.4 African sheep breed development

Due to different breeding goals, African stock raisers have developed over sixty different varieties of sheep through sharing, lending or exchanging breed stocks. Breeding was based on: resistance to major infectious diseases; ability to walk for long distances; perseverance to water and forage scarcity; high fecundity and good mothering abilities; and good production traits. For example, many dwarf sheep breeds from West Africa are preferred because they are resistant to blood parasites and other common diseases. Similarly, D'man sheep of Morocco is preferred by oasis dwellers for meat production. It frequently produces twins, triplets, quadruplets and even quintuplets. Some of these good traits and many others are now threatened by extinction (McCorkle, 1999).

2.5 Threat to livestock biodiversity

Extinction is an irreversible process in which identifiable breeds or genetically controlled characteristics disappear. It can occur at species, sub-species, breeds, and at gene level. For example extreme selection pressure acting against one possible allele in favour of another may result in the complete disappearance of the less favoured gene (Henson, 1992).

According to a review by McCorkle (1999), livestock biodiversity has declined over the years due to high-risk, cash-based economies of the modern world. This has forced stock raisers to either sell off their best animals in order to access other basic needs, or replace their multipurpose indigenous breeds with high producing exotic breeds, or abandon livestock keeping and concentrate on cash crop farming (Cunningham, 1995; FAO, 2005). Similarly, veterinary extension efforts are also to blame for weakening hardy gene pools by keeping sick and deformed animals until mating age. As a result, special genetic qualities are thought to have been lost or diluted. It is estimated that about two livestock breeds are lost every week and 18% of sheep breeds in the world are at a risk of extinction (Blench, 2001; Schearf, 2003). A good example is the Southdown and Oxford Down British sheep breeds which have been close to extinction due to competition they face from generalist meat breeds. Here in Africa the heartwater resistant Tswana sheep of Botswana were at the verge of extinction before a conservation program started (Henson, 1992).

Domestic animal biodiversity is essential for food security and for future challenges (Hall & Bradley, 1995). Of great importance are those breeds that have co-evolved with a particular

environment and farming system (FAO, 2005). African stock raisers keep detailed mental and oral livestock record that is based on morphological features (McCorkle, 1999). However, this is not sufficient for conservation purposes, because a lot of other good genetic traits are not known. Knowledge of genetic variation can provide useful information when decisions concerning management of genetic resources and conservation of unique populations for future breeding strategies are made (Hall & Bradley, 1995).

2.6 Molecular markers commonly used for genetic characterization

Molecular techniques provide a powerful tool that is used in the search for genetic markers. The most popular DNA markers include RFLP, RAPD, microsatellites and SNPs. These markers have provided useful information about the genetic diversity and population structure of different breeds.

2.6.1 Restriction fragment length polymorphism

RFLPs are codominant markers whose analysis is based on size difference of the restricted DNA. This occurs as a result of changes within the recognition sequence shifting the restriction site of a particular restriction enzyme. Individuals are differentiated basing on the size pattern and the number of obtained fragments. This procedure is simple and can be used to screen a large population. However, its requirement for sequence information and enzyme design limits its application (Botstein *et al.* 1980).

2.6.2 Random amplified polymorphic DNA

RAPD is a PCR based technique in which segments of DNA are randomly amplified with an identical pair of primers (Welsh & McClelland, 1990). Since low annealing temperatures are used, multiple products are formed with each product representing a different locus. Genetic variation and divergence within and between the taxa of interest are assessed by the presence or absence of each product, which is dictated by the changes in the DNA sequence at each locus (Liu & Cordes, 2004). RAPD markers are inherited as Mendelian markers in a dominant fashion making it difficult to differentiate homozygote dominant from heterozygote. These markers have

been used to study the origin and diversity of various sheep breeds. For example, Ali (2003) used nineteen random primers to study genetic relationships of four Egyptian sheep breeds.

2.6.3 Microsatellites

Microsatellites consist of tandemly arranged simple sequence repeats (SSRs) that range in size from 1-6 base pairs. They are highly polymorphic with discrete loci and co-dominant alleles. They are classified according to the number of repeat motifs. Di-, tri-, tetra-, penta- and hexanucleotide describe repeat motifs consisting of two, three, four, five and six base pairs respectively. A perfect repeat is one with a pure tract for one motif while a compound microsatellite has two or more different repeat motifs. An example of a compound microsatellite is the Y specific *SRYM18* marker isolated in sheep (Australian Sheep Gene Mapping Website, 2003). Microsatellite polymorphism is based on size differences due to varying numbers of repeat units contained by alleles at a given locus. Changes in the numbers of repeat unit are believed to occur in a stepwise manner due to polymerase slippage. Slippage seems to occur frequently to an immediate adjacent repeat unit leading to a gain or loss of one repeat unit. Different species have got different microsatellite mutation rates. In humans, autosomal mutation rates per generation (meiosis) ranges from $0.6-2.12 \times 10^{-3}$. In sheep, microsatellite mutation rates of up to 1.1×10^{-4} per generation (meiosis) have been reported (Brohede, 2003).

The uniform distribution of microsatellites within the genome (on all chromosomes and all regions of the chromosome) and high variability makes them useful in behavioural ecology, studying population structures and determination of paternity and kinship (Schlötterer, 1998). Microsatellites have been successfully used to characterize the genetic diversity of various livestock including cattle (MacHugh *et al.* 1997; Hanotte *et al.* 2000) sheep (Buchanan *et al.* 1994; Muigai, 2003), goat (Barker *et al.* 2001; Chenyambuga, 2002), camel (Mburu *et al.* 2002), yak (Xuebin *et al.* 2005) and horses (Bjørnstad & Røed, 2001).

2.6.4 Single nucleotide polymorphism

SNPs describes polymorphisms caused by point mutations that give rise to different alleles containing alternative bases at a given nucleotide position within a locus. For such a substitution to be considered as an SNP, the least frequent allele should have a frequency of at least 1%

(Vignal *et al.* 2002). SNPs are becoming important in molecular marker development since they are the most abundant polymorphism in any organism and reveal hidden polymorphism not detected with other markers. Theoretically, a SNP within a locus can produce as many as four alleles, each containing one of four bases (A, T, G and C) at the SNP site. Practically, however, most SNPs are usually restricted to one of the two alleles (most often, either the two pyrimidines C/T or the two purines A/G) and have been regarded as bi-allelic. Like microsatellites, SNPs are inherited as co-dominant markers (Vignal *et al.* 2002; Liu & Cordes, 2004).

2.7 Mitochondrial DNA

Mitochondrial DNA is only maternally inherited and therefore it provides a record of maternal lineage. Diversity studies utilize mtDNA sequence variation. mtDNA markers have been used in phylogenetic and diversity studies of sheep from Europe (Hiendleder *et al.* 1998a, 1998b, 2002; Meadows *et al.* 2005), Turkey (Pedrosa *et al.* 2005), China (Guo *et al.* 2005). Ongoing large scale studies looking at the origin and diversity of sheep breeds include: Nordic Gen Bank for farm animals, European Union Econogene project and the biodiversity project at the International Livestock Research Institute (ILRI).

2.8 Y chromosome

2.8.1 Evolution of the Y chromosome

All domestic sheep have $n = 54$ chromosomes. Of these, two are sex chromosomes (X and Y), while the rest are autosomes. The two sex chromosomes are thought to have evolved from autosomes. Recent reviews (Graves, 2002; Hellborg, 2004) show that the mammalian Y chromosome is about 170-310 Myr old. Its evolution is thought to have started by a mutation in the *SOX3* gene of an autosome producing the *SRY* gene on the proto- Y. This was followed by inversions that relocated the *SRY* from its original position to the top of the Y chromosome relative to the position of the *SOX3* near the bottom of the proto- X. These rearrangements made it impossible for proto-Y and proto-X to recombine. The lack of recombination made the Y chromosome susceptible to deletions that greatly decreased its size. This was followed by two expansions, following donations of genes from autosomes such as the *DAZ* gene from chromosome 3 by transposition. However to preserve its integrity additional rearrangements

occurred and the Y continued to lose genes and over time, shrank. It appears that the acquisition of the sex-determining allele by the Y chromosome triggered a continuous process of degradation that left the Y chromosome with at least 27 genes/ genes families and numerous pseudogenes with active homologues on the X. The Y chromosome seems to be evolving rapidly and due to this degradation process it has been predicted that, the entire human Y will disappear in about 10 million years (Aitken & Graves, 2002; Graves, 2002). Similar degeneration has also been reported in the fruit fly *Drosophila miranda* (Bachtrog & Charlesworth, 2002; Bachtrog, 2004). It is expected that after this demise, either the human Y chromosome will be recycled or a new sex chromosome order (speciation) will occur (Graves, 2005).

2.8.2 The structure of Y chromosome

The human Y chromosome (Figure 3) consists of pseudoautosomal regions at the tips of both the long and short arm and *MSY* region within which the *SRY* is located (Skaletsky *et al.* 2003). A similar map has been described for the horse Y chromosome (Raudsepp, *et al.* 2004). The Y chromosome shares a number of homology regions with the X chromosome and undergoes recombination with it only in the two pseudoautosomal regions. Though 95% of the Y chromosome does not undergo recombination with the X chromosome (Hurles & Jobling, 2001), a recent report shows that genes within the *MSY* recombine amongst themselves enabling them to eliminate any deleterious mutation (Skaletsky *et al.* 2003). The lack of recombination of the *MSY* with the rest of the genome enhances the resolution of the historic information relevant to evolutionary studies (Jobling & Tyler- Smith, 2000).

Y chromosome has been extensively studied in humans. These studies are used in tracking paternal lineages, forensic science and medical sciences (Jobling & Tyler- Smith, 2000). In humans, as in other mammals, the Y chromosome cause testis differentiation and so determines maleness in a dominant fashion through the action of a single gene *SRY* (Sinclair *et al.* 1990).

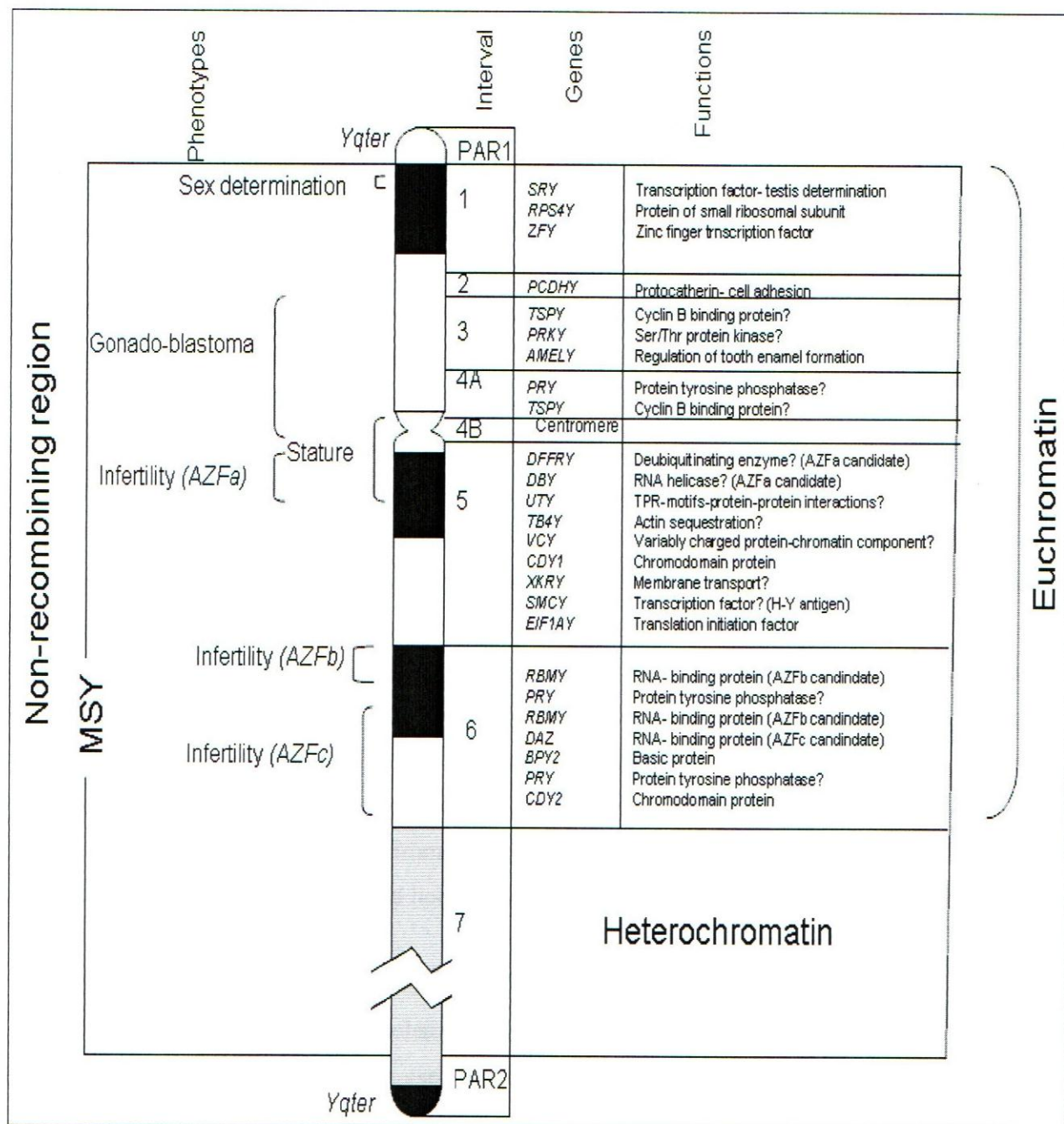


Figure 3. Ideogram of the Y chromosome. It shows the non recombining region, location of the pseudoautosomal regions (PAR), euchromatin with several genes including *SRY*, and the long arm heterochromatin (Jobling & Tyler-Smith, 2000).

2.8.3 Use of Y chromosome markers in livestock diversity studies

The use of Y chromosome markers in livestock diversity studies is still uncommon due to the lack of informative markers. In sheep, only two male specific markers have been described on the ovine Y chromosome (Meadows, *et al.* 2004; Australian Sheep Gene Mapping Website, 2003). Attempts for cross species amplification using bovine microsatellite markers *INRA189*, *INRA126* and *INRA124* showed that these loci reside outside the *MSY* region in sheep (Meadows *et al.* 2006). The lack of numerous markers on the Y chromosome can be attributed to three factors. Firstly is the earlier perception that the Y chromosome is a genetic wasteland (Skaletsky *et al.* 2003). This in effect slowed down the discovery of markers on the mammalian Y chromosome and therefore hindered progress on the utilization of the Y chromosome markers in diversity studies. Secondly, the low mutation rate on the mammalian Y chromosome (Hellborg, 2004; Hellborg & Ellegren, 2004). Unlike the X chromosome, the mammalian Y chromosome show low levels of nucleotide diversity (Brohede, 2003; Hellborg & Ellegren, 2004). One possible reason is that the Y chromosome is exposed to mutations arising from the male germ line only, while both male and female mutations hit the X chromosome (Hurles & Jobling, 2001; Butler, 2003). Lastly, the lack of enough sequence data on the Y chromosome has also slowed down the development of Y specific markers. Other than in humans and mice, the only complete physical map in livestock has been described in the horse Y chromosome (Raudsepp, *et al.* 2004), so mapping of the ovine Y chromosome is likely to open up avenues for the discovery of new male specific markers that could help in the construction of patriline as well as in diversity studies.

In the recent past, various attempts to identify and use Y-specific markers in livestock have been made. Wallner *et al.* (2004) isolated six Y-chromosome specific microsatellite markers in the horse. These markers were then used to study the relationship among three closely related equidae: horse, zebra and ass. Low sequence variability over a total length of 3033 nucleotides among fourteen individuals from different horse breeds was reported. The low variability observed is consistent with the findings of Hellborg & Ellegren (2004). Using *INRA124*, a Y specific microsatellite locus, Hanotte *et al.* (2000) reported that in cattle the indigenous taurine Y chromosome has been nearly eliminated from the Abyssinian region following male mediated zebu introgression.

Recently, Meadows *et al.* (2004) reported that approximately 10% of the expected nucleotide diversity in sheep is present on the ovine Y chromosome. However, comparative sequencing of nine fragments associated with the five *MSY* genes, revealed that only one fragment contained an A / G SNP located approximately 1685 base pairs upstream of the ovine *SRY* gene. A diversity analysis using this marker showed that four rams carried the A allele while six carried the G allele and none of the animals was heterozygous for the mutation as predicted from the *MSY* sequence.

CHAPTER THREE

MATERIALS AND METHODS

3.1 DNA samples

Fat-tailed and thin-tailed sheep breeds used in this study were sampled in 11 African countries, while breeds from outside Africa were obtained from Pakistan, USA, Portugal and UK. Sampling was done in accordance with the FAO guidelines (FAO, 2004b) following the history, relatedness and geographical distribution. A total of 556 animals were used. Details of each breed are provided in Table 1. DNA sample for each individual was obtained from ILRI DNA bank, but when unavailable genomic DNA was re-extracted from blood retrieved from the DNA bank.

3.2 DNA extraction

DNA was extracted from the peripheral blood mononuclear cells stored in 8 M urea using a modified phenol/chloroform protocol (Sambrook *et al.* 1989). A 200 µl aliquot of each sample was transferred into a 1.5 ml ependorff tube containing 100 µl of 2 M urea (BDH, Poole, England) and mixed by pipetting, before adding 150 µl of potassium acetate (with a molar concentration in the ratio 3:2 with respect to Potassium and Acetate respectively) (BDH, Poole, England). Each tube was vortexed gently for 5 seconds and put on ice for 10 minutes to allow cell lysis. Cell debris was precipitated by centrifugation at 14000 x g for 5 minutes. The supernatant was transferred into a fresh ependorff tube and 50 µl of 6 M NaCl (BDH, Poole, England) added. An equal volume (450 µl) of chloroform (BDH, Poole, England) was added and the contents mixed by inverting the tube several times for 5 minutes. Proteins were then precipitated at the interphase by centrifuging at 14000 x g for 6 minutes. The upper aqueous layer containing the nucleic acids was transferred into a fresh ependorff tube and 900 µl (two volumes) of absolute ethanol (Panreac, Barcelona, Spain) was added. The DNA was precipitated by inverting the contents several times and was pelleted by centrifuging at 14000 x g for 6 minutes. The DNA preparation was washed once with 70% ethanol to remove any solutes and air dried before reconstituting in 110 µl of triple distilled deionized water.

Table 1. Sheep samples used for microsatellite genotyping and SNP screening. A total of 556 samples from 35 breeds representing three regions of Africa, East, West and southern Africa together with outgroup breeds from Central America, Asia and Europe were used.

Breed	Breed type*	Source	Region	Sample size
Afar	Fat-tailed	Ethiopia	East Africa	32
Arsi-bale	Fat-tailed	Ethiopia	East Africa	22
Gumuz	Fat-tailed	Ethiopia	East Africa	19
Horro	Fat-tailed	Ethiopia	East Africa	19
Kabale	Fat-tailed	Uganda	East Africa	7
Moroto	Fat-tailed	Uganda	East Africa	20
North West Highland	Fat-tailed	Ethiopia	East Africa	25
Red Maasai	Fat-tailed	Kenya	East Africa	21
Red Maasai	Fat-tailed	Tanzania	East Africa	22
Sekota	Fat-tailed	Ethiopia	East Africa	15
Sukuma	Fat-tailed	Tanzania	East Africa	24
Tukur	Fat-tailed	Ethiopia	East Africa	17
Ugogo	Fat-tailed	Tanzania	East Africa	18
Wollo	Fat-tailed	Ethiopia	East Africa	19
Rutana	Thin-tailed	Ethiopia**	East Africa	21
Black head Persian	Fat-tailed	South Africa	southern Africa	10
Damara	Fat-tailed	Namibia	southern Africa	18
Karakul	Fat-tailed	Namibia***	southern Africa	25
Pedi	Fat-tailed	South Africa	southern Africa	15
Sabi	Fat-tailed	Zimbabwe	southern Africa	14
Tswana	Fat-tailed	Botswana	southern Africa	11
Bali-bali	Thin-tailed	Mali	West Africa	7
Balami	Thin-tailed	Nigeria	West Africa	12
Djallonke	Thin-tailed	Nigeria	West Africa	19
Djallonke	Thin-tailed	Senegal	West Africa	34
Maure	Thin-tailed	Mali	West Africa	12
Touareg	Thin-tailed	Mali	West Africa	19

Yankassa	Thin-tailed	Nigeria	West Africa	10
Barbados black belly	Thin-tailed	US Virgin Island	Central America	5
St Croix	Thin-tailed	US Virgin Island	Central America	12
Lohi	Fat-tailed	Pakistan	Asia	6
Salt range	Fat-tailed	Pakistan	Asia	4
Portuguese white Merino	Thin-tailed	Portugal	Europe	12
Swaledale	Thin-tailed	UK	Europe	7
Texel	Thin-tailed	UK	Europe	3
Total				556

* Information from DAGRIS whenever available

** Sampled in Ethiopia but originally from Sudan

*** Sampled in Namibia but originally from Russia

3.3 DNA quantification

DNA quality was analyzed on a 0.8% (w/v) agarose gel (See gel electrophoresis in the proceeding sections). Good quality DNA was then quantified using GeneQuant® (Biochrom, Cambridge, UK) and diluted to a working concentration of 40 ng/μl. Samples with poor quality or low yields were re-extracted using the same protocol.

3.4 Polymerase chain reaction

DNA amplification was done using GeneAmp® PCR system 9700 (Applied Biosystems, Norwalk, USA). Genomic DNA from 556 individuals was amplified using two Y chromosome specific primers using different conditions as follows.

A 10 μl reaction for the microsatellite *SRYM18*, consisting of 40 ng template DNA, 0.1 μM of each primer (Forward = 5' fluorescent Tag 6FAM – GGCATCACAAACAGGATCAGCAAT – 3' ; Reverse = 5' – GTGATGGCAGTTCTCACAATCTCCT - 3'), 1.5 mM MgCl₂, 0.125 mM dNTP, and 0.3 U *Taq* DNA polymerase (Promega, Madison, U.S.A) was set up. Initial

denaturation was done for 5 minutes at 95 °C followed by 30 cycles of 30 seconds at 95 °C, 30 seconds annealing at 60 °C, 1 minute of primer extension at 72 °C and a final extension of 10 minutes.

A 20 µl reaction for the single nucleotide polymorphism marker *SRY3* consisting of 40 ng template DNA, 0.1 µM of each primer (Forward = 5' fluorescent Tag PET – TCAGTAGCTTAGGTACATTCA – 3'; Reverse = 5' – GTGCTACATAAATATGATCTGC – 3', 1.5 mM MgCl₂, 0.125 mM dNTP, and 0.6 U *Taq* DNA polymerase (Promega, Madison, U.S.A). Initial denaturation was done for 15 minutes at 95 °C followed by 35 cycles of 45 seconds at 95 °C, 45 seconds annealing at 60 °C, 45 seconds of primer extension at 72 °C and a final extension of 7 minutes.

3.5 Agarose gel electrophoresis

Agarose gels were prepared by boiling 0.8 g, 1 g or 2 g (for 0.8%, 1% and 2% w/v gel respectively) of agarose (Invitrogen, U.S.A.) in 100 ml of 1 X TBE (0.09 M Tris, 0.09 M Boric acid and 0.002 M EDTA, pH 8.3) (Sambrook *et al.* 1989). The gel solution was allowed to cool to below 60 °C before adding ethidium bromide (final concentration 0.5 µg/ml) and the contents mixed by stirring (Sambrook *et al.* 1989). The gel tray was then sealed on both sides using paper tapes and combs put in place before pouring the gel. After casting, the gel was allowed to polymerise for about 30 minutes before electrophoresis.

After setting, the paper tapes were removed and the tray containing the gel put in the electrophoresis tank containing 1 X TBE. To check the efficiency of amplification, a 5 µl aliquot of the PCR product was mixed with 1 µl of 6 X loading buffer (15% Ficoll[®] 400, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 0.4% orange G, 10 mM Tris-HCl (pH 7.5) and 50 mM EDTA) was then loaded into the wells. A 5 µl aliquot of the PCR size marker (Promega, Madison, U.S.A.) mixed with the loading buffer was loaded in the first and the last well alongside the samples. A constant voltage of 100 V was used to fractionate the samples for about 45 minutes. The gel was visualized and image captured using the UVIpro Version 12.4 gel image software (UVitec, Cambridge, UK).

3.6 DNA genotyping

The *SRYM18* PCR product was diluted 2 X in a microtitre PCR plate using triple distilled deionized water and mixed by pipetting. The LIZ standard was prepared by mixing 10 µl of the standard with 1 ml of Hi-DiTM formamide. A 1 µl aliquot of the diluted PCR product was transferred into a 96 well PCR plate and 9 µl of LIZ standard/Hi-DiTM formamide (Applied Biosystems, Warrington, UK) was added. The PCR plate was then span for 30 seconds at 2000 x g. The mixture was denatured for 4 minutes at 95 °C and immediately put on ice. Fragment analysis was done using automated capillary DNA sequencers (ABI 3730 or 3100, Applied Biosystems, Warrington, UK)

3.7 Genotype analysis

Genotype data was analyzed using GeneMapperTM software V 3.7 (Applied Biosystems, Warrington, UK) using a bin (base pair range that define an allele) ranging from 70-200 bp that overlapped the expected size of the amplified fragment. Allele size calling was done using the third order least squares method. Data on GeneMapper software was exported to Excel and the genotype edited. Samples that failed to amplify or had a lot of background noise were re-amplified and re-genotyped using the same conditions and procedures. After re-amplification, samples that consistently failed to amplify were not scored and were excluded in subsequent analyses.

3.8 Statistical analyses

Scored alleles were used to calculate allele frequencies. The frequency was scored as the proportion of all alleles of the gene that are of the specified type. The allele frequency was calculated by dividing the sum of a given type of allele in a breed by the breed size (Hartl & Clark, 1997). In this case, the breed size refers to the sum of all the alleles scored in that breed.

Further, the allele data was used to calculate variance within breeds, breed types and groups using Arlequin software V 3.0 (Excoffier *et al.* 2005). Analysis was done on 35 samples (breeds) using default values for standard data. Breeds were classified into three groups for the first structure. All the African fat-tailed formed the first group while all the African thin-tailed and all non African breeds formed the second and third group respectively. In the second structure

analysis, only the African breeds were used, the first group comprised of all sheep breeds from East Africa while the second and third groups comprised of breeds from West and South Africa respectively. Similarly in the third structure only African breeds were used. The first group had all the fat-tailed while the second group had all the thin-tailed sheep breeds.

Logistic regression was done to test the significant differences in the proportion of each allele using Proc Logistic in SAS (PC-SAS V 9.1, SAS institute, Inc. Cary, NC, USA). Allele scores were compared to each other within the Africa regions, between the African regions, between the continents and between the two main breed types, thin-tailed and fat-tailed.

3.9 *SRYM18* allele sequencing

To confirm the scored genotypes and assess homoplasmy, the PCR products of thirteen individuals (Arsi-bale 10, 131 bp; Maure 16, Lohi 37, and Lohi 41, 139 bp allele; Pedi 31, Touareg 21 and Yankassa 27, 141 bp allele; Bali-bali 02, Balami 02 and Karakul 22, 143 bp allele; and Sabi 15, Sukuma 02 and Sukuma 16, 145 bp allele) were purified for sequencing. Details of the procedure are as described in QIAGEN[®] PCR purification kit (<http://molecool.wustl.edu/krolllab/PDFs/>) protocol. Briefly, the PCR products were separated on a 2% agarose gel. Each band was excised under UV, transferred into a 2 ml ependorff tube and the weight determined. Six volumes (1200 µl) of QG buffer was added and the gel slice solubilized by incubating at 50 °C for 15 minutes with regular vortexing. To increase the DNA yield, one gel volume (200 µl) of isopropanol was added and the solution mixed. The solution was transferred into QIAquick[®] column and centrifuged for 1 minute at 10,000 x g. Any remaining traces of agarose was removed by adding 500 µl of QG buffer to the QIAquick[®] column and centrifuged for 1 minute at 10,000 x g. The salts were washed for 5 minutes using ethanol containing PE buffer and the DNA eluted into a fresh ependorff tube using 30 µl of EB buffer (10mM TrisCl, pH 8.5).

The PCR fragment was then sequenced using BigDye[®] terminator cycle sequencing kit according to the manufacturer's recommendation (<http://docs.appliedbiosystems.com/>). Briefly, a 10 µl reaction was set by adding 0.5 µl of the terminator ready mix, 1 µl of 5 X sequencing buffer, 20 ng of the PCR product, 3.2 pmoles of the *SRYM18* primer and de-ionized water into each well of a clear optical reaction plate. The plate was sealed, the contents mixed and spun

briefly. Samples were then subjected to 25 cycles of rapid thermal ramp to 96 °C, followed by a 10s hold; rapid thermal ramp to 50 °C, followed by a 5s hold and a final rapid thermal ramp to 60 °C, followed by a 4 minutes hold.

After amplification, samples were purified to remove unincorporated dye terminators as follows: The plate was briefly spun before adding 1 µl of 125 mM EDTA, 1 µl of 3 M sodium acetate and 25 µl of absolute ethanol to each well. The contents were then incubated for 15 minutes at room temperature to precipitate the extension products before spinning for 30 minutes at 3000 x g. The plate was then inverted and spun to 185 x g to remove the supernatant. A 100 µl aliquot of 70% absolute ethanol was added to each well and spun for 15 minutes at 3000 x g. The plate was then inverted and spun to 185 x g for 1 minute to remove the supernatant. The samples were incubated at 90 °C for 1 minute, resuspended in 10 µl HiDi formamide and denatured by heating for 2 minutes at 95 °C and cooled on ice. Sequences were analyzed using automated capillary Applied Biosystems DNA sequencers (ABI 3730 or 3100).

3.10 Sequence Analysis

Raw data of the DNA sequence was edited using Chromas lite software V 2.01. Consensus sequences were generated from the forward and reverse sequences. The consensus sequences were aligned together with sequences from the study done by Meadows *et al.* (2006) using ClustalX (1.81) (Thompson *et al.* 1997).

3.11 Single nucleotide polymorphism analysis

Screening was based on RFLP using A/G SNP specific *Fnu4HI* restriction enzyme (recognition sequence is 5'- GC[↓]NGC- 3', 3'- CGN[↓]CG- 5'). A 27 µl digestion reaction consisted of 15 µl aliquot of the *SRY3* PCR product, 10 µl of triple distilled deionized water, 2 U *Fnu4HI* restriction enzyme and 1 X NEB buffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM DTT, pH 7.9) (New England BioLabs® Inc., Beverly, USA). The reaction was incubated for 2 hours at 37 °C and was stopped by adding 5 µl of the loading buffer. The digest was run on a 2% agarose gel. Existence of two bands (138 bp and 473

bp) indicates the A SNP while existence of 3 bands (49 bp, 89 bp and 473 bp) indicates the G SNP.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 RESULTS

4.1.1 PCR optimization

The male specific microsatellite marker *SRYM18* was optimized against both male and female individuals. The PCR products were then run on 2% agarose gel and the results are shown in Figure 4 below.

A single PCR product of approximately 145 bp was observed in the male samples but was lacking in the female samples. This result confirms that the *SRYM18* primer set is male specific. Subsequently all male samples were amplified using these primers. Four female samples were selected and included in each reaction plate together with water as negative controls.

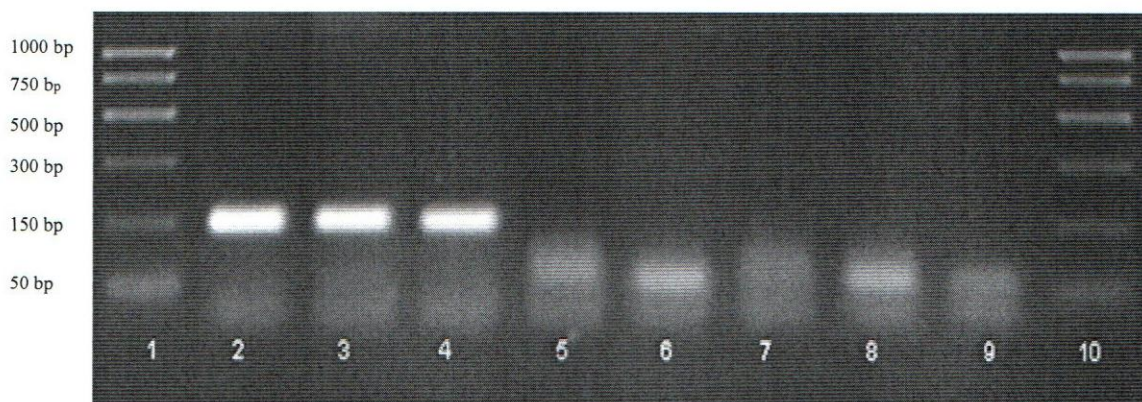


Figure 4. *SRYM18* PCR optimization gel photograph. Lane 1 and 10 was loaded with PCR marker. Lane 2-4 shows PCR products of about 145 bp from three male individuals. Lanes 5-8 contained products from female individuals, while lane 9 was loaded with water.

Similarly, the SNP marker *SRY3* was optimized against both male and female individuals and the PCR digested with *Fnu4HI* restriction enzyme. The resultant fragments were run on 2% agarose gel and the results are shown in Figure 5 below. Subsequently all the 556 individuals were screened for the SNP using this primer set.

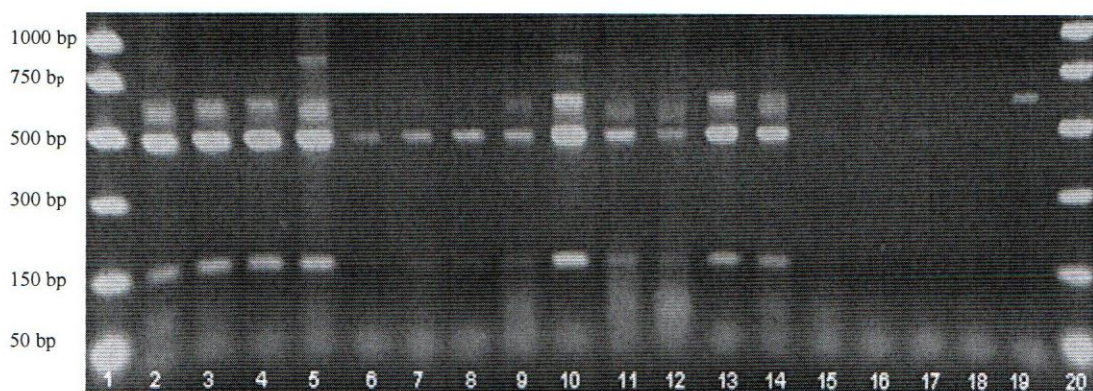


Figure 5. *SRY3* PCR optimization gel photograph. Lane 1 and 20 have PCR size marker. Lane 19 shows the undigested PCR products of about 611 bp from a male individual. Lanes 2 -14 is a partial restriction digest of the 611 bp fragment from 13 male individuals. The presence of two bands (of about 500 bp and 150 bp) is an indication of the A allele. Lanes 15-17 contained female individuals, while lane 18 had water.

4.1.2 Genotyping

Of the 556 male samples, only 491 samples amplified using the *SRYM18* marker and were subsequently genotyped. A total of 65 DNA samples did not amplify even after DNA re-extraction. Genotyping revealed the presence of five alleles of sizes 131 bp, 139 bp, 141 bp, 143 bp and 145 bp. Electropherograms of the peaks are presented in Figure 6 and the distribution of alleles in populations is presented in Table 2.

The 143 bp allele was the most wide spread allele in the African sheep breeds. It was observed in 90.2% of individuals from East Africa and in 53.0% and 44.6% of individuals from southern Africa and West Africa, respectively. Within the breeds outside Africa, it was observed in three thin-tailed breeds (St Croix, Portuguese white merino and Swaledale) and one fat-tailed (Salt range) breed.

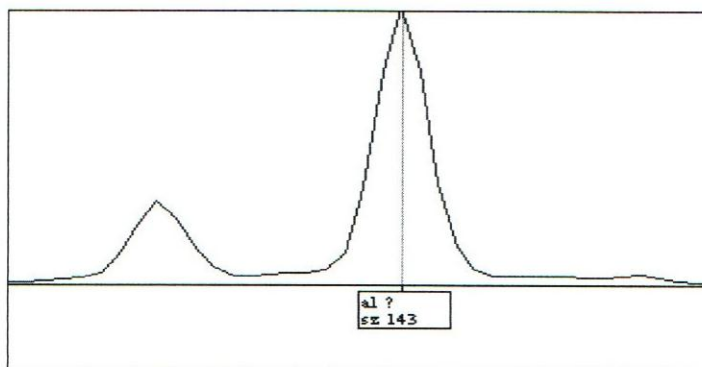


Figure 6a. Electropherogram of the 143 bp allele

The 141 bp allele was the second most widely distributed allele within the African breeds. In total, 19.7% of individuals from all the three regions of Africa had this allele. This allele was the most common allele (53.5%) in West Africa. It was also observed in 35.0% of individuals from southern Africa and in 1.2% of individuals from East Africa. However, none was observed in breeds from outside Africa.

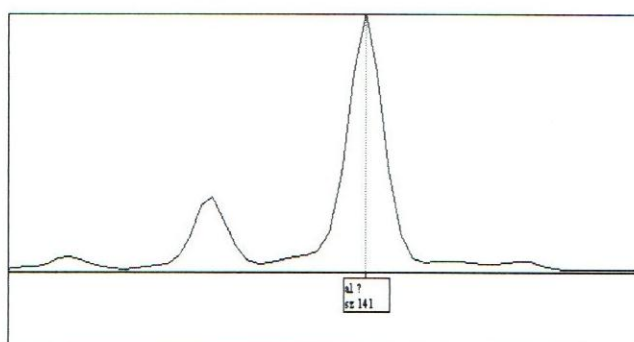


Figure 6b. Electropherogram of the 141 bp allele

The largest allele in size was the 145 bp and it appeared in fat-tailed sheep from East Africa (7.8%), southern Africa (12.0%) and in most thin-tailed individuals from Europe and Central America. However, it was not observed in any of the breeds from West Africa and Pakistan.

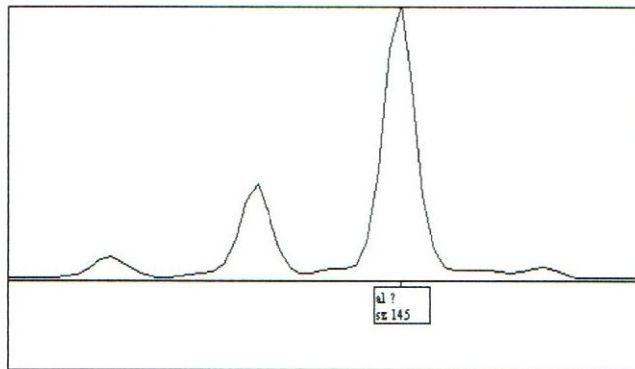


Figure 6c. Electropherogram of the 145 bp allele

Within the African breeds, the 139 bp allele was found in three individuals. Two of the individuals were from the West Africa's thin-tailed Djallonke and Maure breeds from Senegal and Mali respectively while the third one was the fat-tailed Wollo sheep from Ethiopia. This allele was also observed in all the six, fat-tailed, Lohi sheep from Pakistan.

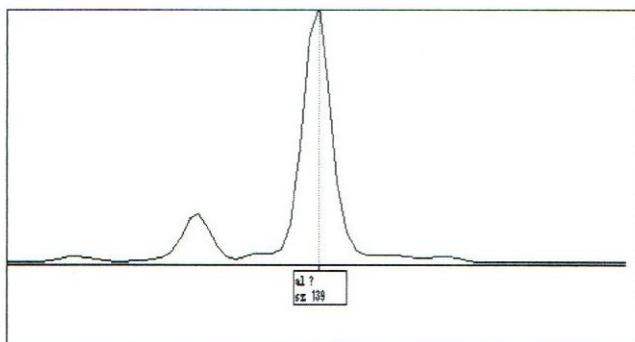


Figure 6d. Electropherogram of the 139 bp allele

The 131 bp allele was only observed once in an Arsi-bale sheep, which is an East African fat-tailed breed from Ethiopia. This was a unique allele and therefore this sample was re-amplified and re-genotyped and the same result was confirmed.

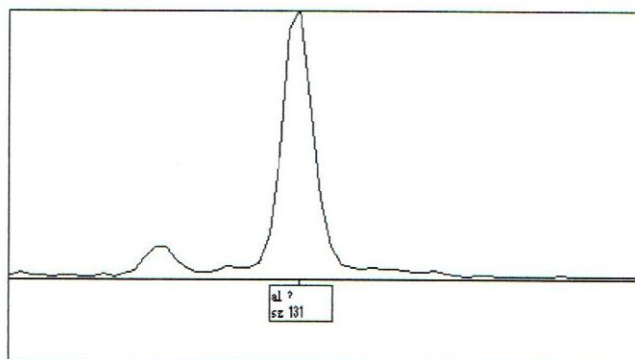


Figure 6e. Electropherogram of the 131 bp allele

4.1.3 Allele frequencies

The allele frequencies of each breed were calculated based on the total number of genotypes observed within each breed. The resultant frequencies are shown in Table 2. In most breeds, two different alleles were detected with varying frequencies indicating Y chromosome diversity. The 141 bp and the 143 bp alleles were observed in a total of 11 African breeds while the 143 bp and 145 bp alleles were observed in 10 African breeds as well as in two breeds from outside Africa, the St. Croix and the Swaledale. Two African breeds showed the highest Y chromosome diversity with three different alleles, the 139 bp, 141 bp and 143 bp, these were Djallonke from Senegal and Maure of Mali. This level of diversity was not observed in any breed from outside Africa. In another three African breeds, Gumuz, Sekota and Tukur which are fat-tailed sheep from Ethiopia, only one allele type (143 bp) was observed implying lack of Y chromosome diversity. A similar scenario was also observed in a majority of breeds from outside Africa with the exception of St Croix from the US Virgin Island and Swaledale from the UK that had two alleles each.

Table 2. Microsatellite allele frequencies. Twenty eight sheep breeds were from Africa while the other seven breeds from outside Africa were included as outgroups.

Breed	Source	Breed type	n	Allele	Total	% allele Frequency
1 Afar	Ethiopia	FT	32	143	25	96.0
				145	1	4.0
				n/a	6	
2 Arsi-bale	Ethiopia	FT	22	131	1	5.0
				143	21	95.0
3 Bali-bali	Mali	TT	7	141	2	29.0
				143	5	71.0
4 Balami	Nigeria	TT	12	141	5	45.0
				143	6	55.0
				n/a	1	
5 Black head persian	South Africa	FT	10	143	9	90.0
				145	1	10.0
6 Damara	Namibia	FT	18	141	6	33.0
				143	11	67.0
				n/a	1	
7 Djallonke	Nigeria	TT	19	141	6	35.0
				143	11	65.0
				n/a	2	
8 Djallonke	Senegal	TT	34	139	1	4.0
				141	23	85.0
				143	3	11.0
				n/a	7	
9 Gumuz	Ethiopia	FT	19	143	17	100
				n/a	2	
10 Horro	Ethiopia	FT	19	143	15	88.0
				145	2	12.0
				n/a	2	
11 Kabale	Uganda	FT	7	143	3	75.0
				145	1	25.0
				n/a	3	
12 Karakul	Namibia	FT	25	141	18	95.0
				143	1	5.0
				n/a	6	
13 Maure	Mali	TT	12	139	1	9.0
				141	6	55.0
				143	4	36.0
				n/a	1	
14 Moroto	Uganda	FT	20	143	12	86.0
				145	2	14.0
				n/a	6	
15 North West Highland	Ethiopia	FT	25	141	1	4.0

16 Pedi	South Africa	FT	15	143	24	96.0
				141	6	43.0
				143	8	57.0
17 Red Maasai	Kenya	FT	21	n/a	1	
				141	2	18.0
				143	9	82.0
18 Red Maasai	Tanzania	FT	22	n/a	10	
				143	20	95.0
				145	1	5.0
19 Rutana	Ethiopia	TT	21	n/a	1	
				141	1	6.0
				143	15	94.0
20 Sabi	Zimbabwe	FT	14	n/a	5	
				143	7	50.0
				145	7	50.0
21 Sekota	Ethiopia	FT	15	143	14	100
				n/a	1	
				143	13	57.0
22 Sukuma	Tanzania	FT	24	145	10	43.0
				n/a	1	
				143	16	100
23 Tukur	Ethiopia	FT	17	n/a	1	
				143	16	100
				n/a	1	
24 Touareg	Mali	TT	19	141	6	33.0
				143	12	67.0
				n/a	1	
25 Tswana	Botswana	FT	11	143	9	82.0
				145	2	18.0
				143	14	88.0
26 Ugogo	Tanzania	FT	18	145	2	12.0
				n/a	2	
				139	1	5.0
27 Wollo	Ethiopia	FT	19	143	18	95.0
				141	6	60.0
				143	4	40.0
Barbados black belly	US Virgin Island	TT	5	145	4	100
St Croix	US Virgin Island	TT	12	n/a	1	
				143	2	18.0
				145	9	82.0
Lohi	Pakistan	FT	6	n/a	1	
Portuguese white merino	Portugal	TT	12	139	6	100
				143	10	100
				n/a	2	
Salt range	Pakistan	FT	4	143	4	100
Swaledale	UK	TT	7	143	2	30.0
				145	5	70.0
Texel	UK	TT	3	145	2	100

N = Sample size
n/a = Not amplified
FT = Fat-tailed
TT = Thin-tailed

To allow easy comprehension of results, pie charts representing the allele frequencies within each breed studied were constructed for each country on the map of Africa (Figure 7). Each allele is represented by a unique colour. The presence of various colours within a breed is an indication of high Y chromosome diversity and *vice versa*. A comparison of colour pattern between breeds shows that the Y chromosome diversity of the West African breeds is different from that of the East African breeds. However, the southern African breeds are a mixed pool representing both the East African and West African breeds.

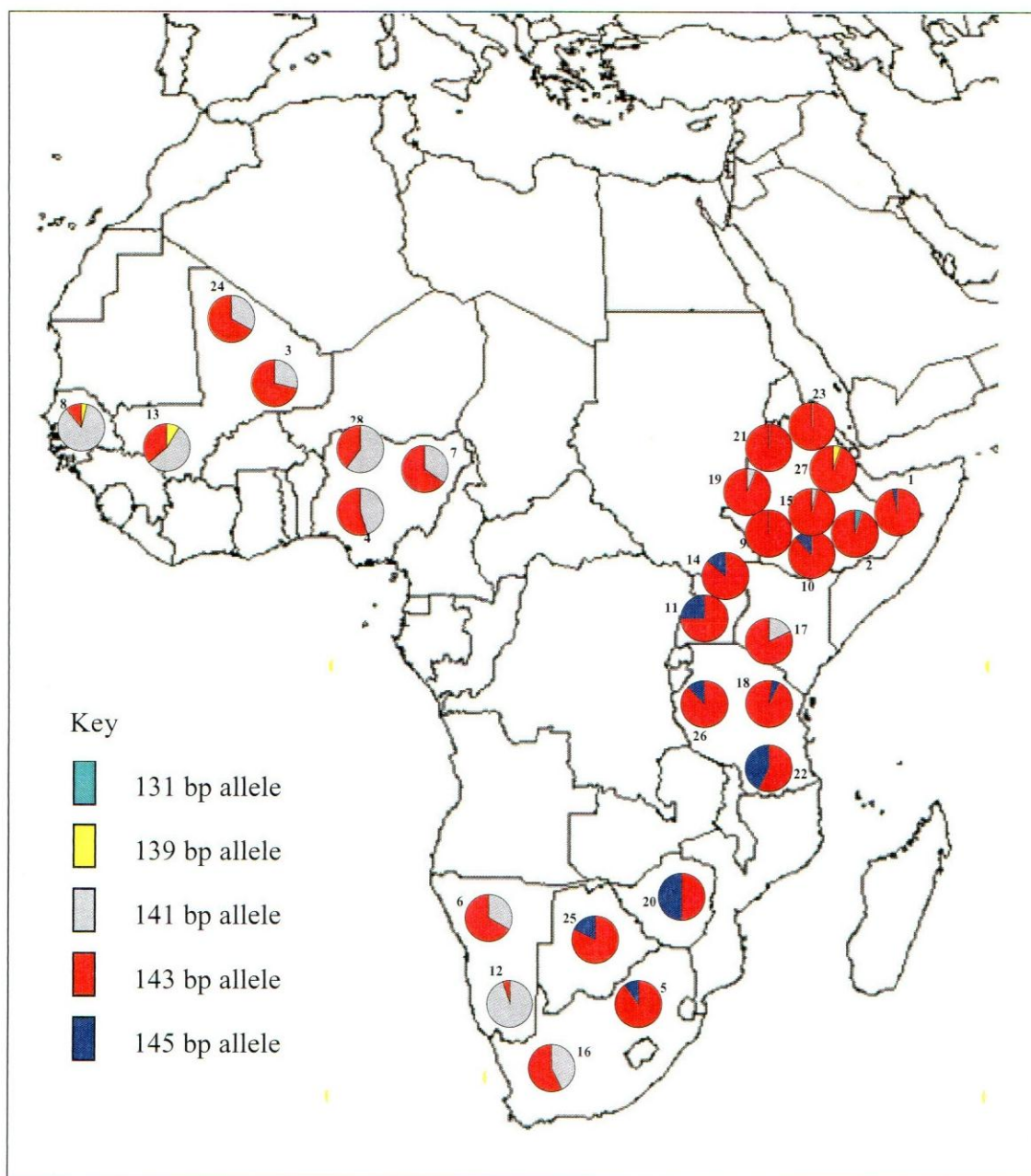


Figure 7. Allele distribution on the map of Africa. Allele frequencies in each breed studied are presented in a pie chart within their respective country of origin. Multiple colouration indicate Y chromosome diversity in breeds. Each pie chart is identified by a unique number given to each breed as shown in Table 2 above.

The distribution pattern of two alleles (141 bp and 143 bp) in Figure 7, gave an indication of the possible origin and migration routes for these alleles based on the history of the breeds in which these alleles were detected. The 141 bp allele is commonly found in the thin-tailed sheep. Since these sheep were the first one to enter the African continent; most likely the 141 bp allele entered the continent first. Similarly the 143 bp allele is the most common among the fat-tailed sheep which entered the African continent after the entry of the thin-tailed sheep. This allele most likely entered the African continent after the entry of the 141 bp allele. The presence of both the 141 bp and 143 bp alleles in some fat-tailed and thin-tailed sheep of West, East and southern Africa could have occurred due to migration of the African communities, trade and crossbreeding. For better visualization, the proposed origin and the likely route of entry and migration in Africa were constructed for the two alleles (Figure 8). Each allele is represented by the same colour assigned in Figure 7.

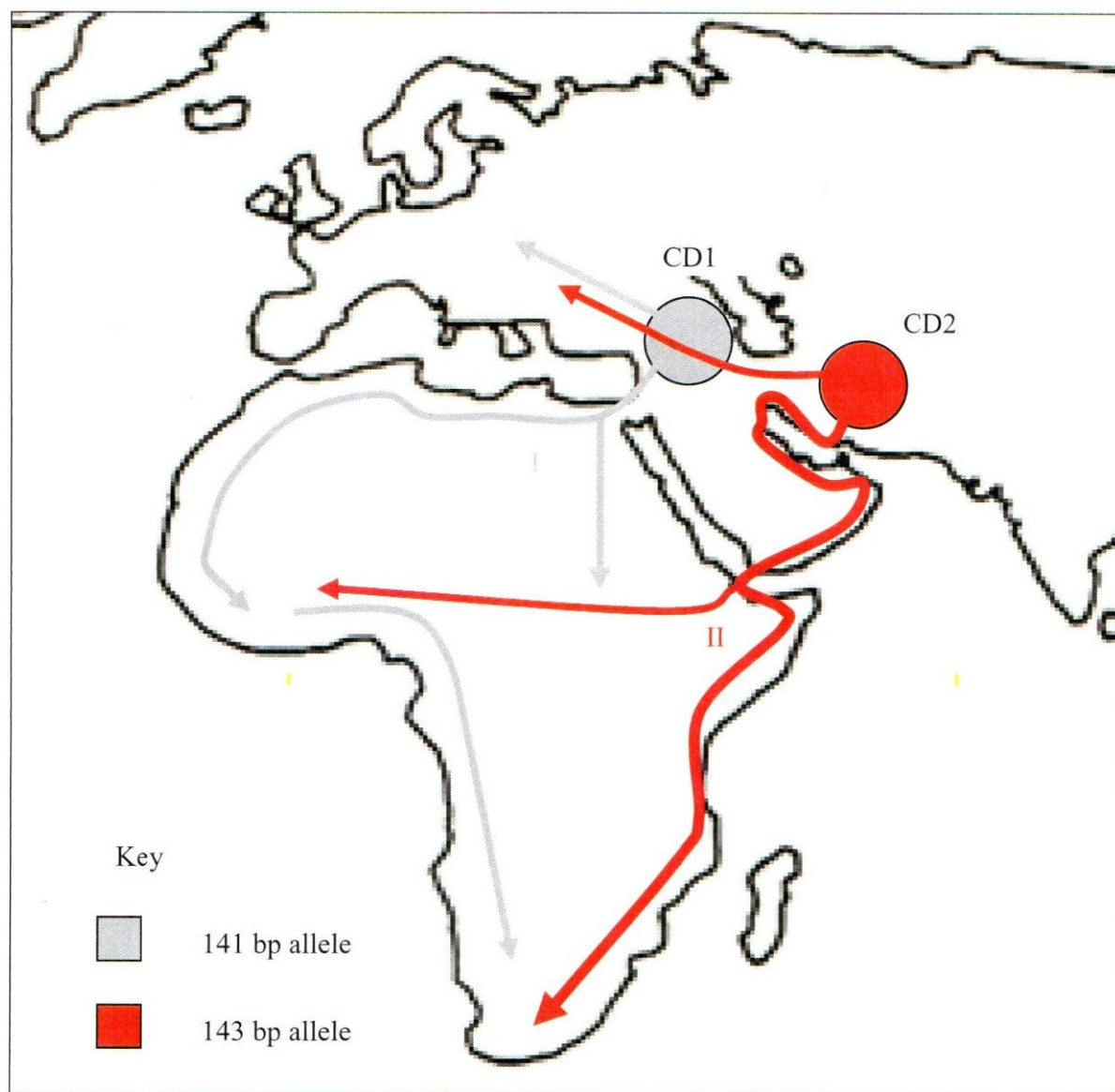


Figure 8. Origin and migration pattern of the *SRYM18* alleles in Africa. Each allele is represented by a unique colour. Centers of domestication are numbered CD1 and CD2 for the Fertile Crescent and the second center of domestication respectively. The migration routes taken by the 141 bp and 143 bp alleles are numbered I and II respectively.

The sheep breeds from each region of Africa were segregated into two main breed types, fat-tailed sheep and thin-tailed sheep. The allele frequencies for either of the breed type were then worked out and the results are presented in Table 3.

Table 3. Microsatellite allele frequencies of sheep breeds classified by type and region

Breed type	Number and % frequency					Total
Allele in bp	131	139	141	143	145	
East African fat-tailed	1	1	3	221	19	245
%Frequency	0.4	0.4	1.2	90.2	7.8	100
Southern African fat-tailed	0	0	30	45	10	85
%Frequency	0.0	0.0	35.0	53.0	12.0	100
African fat-tailed	1	1	33	266	29	330
%Frequency	0.3	0.3	10.0	80.6	8.8	100
East African thin-tailed	0	0	1	15	0	16
%Frequency	0.0	0.0	6.3	93.7	0.0	100
West African thin-tailed	0	2	54	45	0	101
%Frequency	0.0	2.0	53.5	44.6	0.0	100
African thin-tailed	0	2	55	60	0	117
%Frequency	0.0	1.7	47.0	51.3	0.0	100
African fat- tailed and thin-tailed	1	3	88	326	29	447
%Frequency	0.2	0.7	19.7	72.9	6.5	100
Asian fat-tailed	0	6	0	4	0	10
%Frequency	0.0	60.0	0.0	40.0	0.0	100
European thin-tailed	0	0	0	12	7	19
%Frequency	0.0	0.0	0.0	63.0	37.0	100
Central American thin-tailed	0	0	0	2	13	15
%Frequency	0.0	0.0	0.0	13.0	87.0	100
Total population size						491

The 143 bp allele is the most common in both the fat-tailed (80.6%) and thin-tailed (51.3%) sheep of Africa. The allele frequencies are significantly different between the two types of sheep ($P < 0.0001$) (Table 4). The 143 bp allele is also significantly more frequent in the fat-tailed sheep of East Africa (90.2%) compared to the fat-tailed sheep from southern Africa (53.0%) and

thin-tailed sheep of West Africa (44.6%) ($P < 0.0001$). In the Asian fat-tailed sheep, it has a frequency of 40.0% while in the thin-tailed sheep of Europe and Central America, it has a frequency of 63.0% and 13.0% respectively. The frequency of this allele in the Asian and European sheep is not significantly different from the frequency of this allele on the African continent ($P > 0.05$). But, the frequency in Central America is significantly lower than the frequency on the African continent ($P < 0.05$).

The 141 bp allele appears in both the thin-tailed (47.0%) and fat-tailed (10.0%) breed types of Africa. These allele frequencies are not significantly different between the two sheep types ($P > 0.05$). It has a significantly higher frequency (35.0%) in the southern Africa fat-tailed sheep compared to 1.2% in the East African fat-tailed sheep ($P < 0.0001$). West African thin-tailed sheep had significantly higher frequency (53.5%) than the southern Africa fat-tailed sheep (35%) ($P < 0.0001$). The 141 bp allele does not appear in any breed outside Africa.

In the fat-tailed sheep, the 145 bp allele had a frequency of 8.8% and it was not observed in the thin-tailed sheep breeds of Africa. Regionally, its frequency is significantly higher within the southern African fat-tailed sheep breeds (12%) compared to 7.8% in East African fat-tailed sheep breeds ($P < 0.05$). On the contrary, in non African breeds it had a significantly higher frequency of 87.0% in the thin-tailed breeds of Central America and 37.0% in the thin-tailed breeds of Europe compared to 6.5% of the African sheep ($P < 0.05$). However, it was absent in the fat-tailed breeds of Asia.

The 139 bp allele was detected in only one fat-tailed individual (0.4%) and two thin-tailed individuals (2.0%) from East and West Africa, respectively. In non African breeds, it has a frequency of 60.0% in the fat-tailed sheep breeds from Pakistan. None was observed in breeds from Europe and Central America.

The 131 bp allele was observed in a single fat-tailed sheep (Arsi-bale 10) from East Africa and the calculated frequency was 0.4%. To rule out on artifacts, the DNA from this individual was re-amplified and re-genotyped and the same result was confirmed.

Table 4. Logistic regression analysis of *SRYM18* microsatellite data.

Parameter	DF	Estimate	Standard error	Wald χ^2	Pr> χ^2
143 bp allele					
Intercept	1	0.3505	0.1040	11.3546	0.0008
Fat-tailed Vs Thin-tailed	1	0.5601	0.1040	28.9887	<.0001
Intercept	1	0.6280	0.1270	24.4501	<.0001
E. Africa vs S. Africa	1	0.6926	0.1270	29.7333	<.0001
Intercept	1	0.2707	0.1052	6.6205	0.0101
E. Africa vs W. Africa	1	1.0188	0.1327	58.9431	<.0001
S. Africa vs W. Africa	1	-0.3352	0.1594	4.4209	0.0355
Intercept	1	7.6908	347.0	0.0005	0.9823
Africa Vs Asia	1	-6.7802	347.0	0.0004	0.9844
Intercept	1	-0.4267	0.3085	1.9134	0.1666
Africa Vs Europe	1	0.2170	0.3244	0.4474	0.5036
America Vs Europe	1	-1.1826	0.5433	4.7390	0.0295
141 bp allele					
Intercept	1	-0.5641	0.1371	16.9221	<.0001
Fat-tailed Vs Thin-tailed	1	-0.2020	0.1371	2.1701	0.1407
Intercept	1	-0.9254	0.2028	20.8112	<.0001
E. Africa vs W. Africa	1	-1.8312	0.3602	25.8440	<.0001
S. Africa vs W. Africa	1	0.9944	0.2533	15.4102	<.0001
145 bp allele					
Intercept	1	-1.3920	0.2240	38.6071	<.0001
E. Africa vs S. Africa	1	-0.4757	0.2240	4.5092	0.0337
Intercept	1	0.1320	0.3063	0.1858	0.6665
Africa Vs Europe	1	-1.7619	0.3280	28.8610	<.0001
America Vs Europe	1	1.0466	0.4503	5.4019	0.0201
139 bp allele					
Intercept	1	-2.9907	0.6281	22.6687	<.0001
E. Africa vs W. Africa	1	0.1004	0.6281	0.0255	0.8731
Intercept	1	4.9673	133.3	0.0014	0.9703
Africa vs Asia	1	-7.9958	133.3	0.0036	0.9522

Analysis of molecular variance using the microsatellite data of all breeds studied (Table 5) showed a variation of 49.7% within breeds, 32.4% of the variance is among breeds while 17.8% variation is among groups of Africa, Asia, Central America and Europe. When the African breeds were grouped into the three regions West, East and southern Africa, there was an increase in variation. Within breeds, variation increased to 55.3% while variation among regions increased to 27.9%. When the African breeds were further analyzed by breed type, there was a

slight increase of the within breed (57.0%) and among breeds (28.6%) variation. However, distribution of the variation between the thin-tailed and fat-tailed was low (14.4%).

Table 5. Analysis of molecular variance within and between breeds.

Source of Variation	Degrees of freedom	Sums of squares	Variance component	Percentage variation
Worldwide				
Among groups	2	11.907	0.04485 Va	17.8
Among breeds	32	38.880	0.08153 Vb	32.4
Within breeds	432	54.014	0.12503 Vc	49.7
Regions of Africa				
Among groups	2	16.764	0.06494 Va	27.9
Among breeds	25	18.062	0.03932 Vb	16.9
Within breeds	396	51.044	0.12890 Vc	55.3
Africa's thin-tailed and fat-tailed				
Among groups	1	6.093	0.3259 Va	14.4
Among breeds	26	28.734	0.06467 Vb	28.6
Within breeds	396	51.044	0.12890 Vc	57.0

4.1.4 SRYM18 allele sequence analysis

The *SRYM18* amplified products for thirteen individuals were sequenced. The full sequences for each allele as well as representative sequences from Meadows *et al.* (2006) were aligned and are shown in Figure 9 below. Details of the repeat arrays are presented in Table 6. For all the thirteen individuals the pentanucleotide repeat is fixed at three units while the dinucleotide unit varies between individuals. It repeats nine, thirteen, fourteen, fifteen and sixteen times respectively for the 131 bp, 139 bp, 141 bp, 143 bp and 145 bp alleles.

Table 6. Microsatellites repeat units. The pentanucleotide unit is fixed at 3 for all the alleles while the dinucleotide repeat varies with allele size. Comparable repeat units from Meadows *et al.* (2006) are also provided.

Sample	Source	Breed type	Allele	Repeat units
Arsi-bale10	Ethiopia	Fat-tailed	131 bp	(TTTTG) ₃ G(TG) ₉
Maure16	Mali	Thin-tailed	139 bp	(TTTTG) ₃ G(TG) ₁₃
Lohi 37	Pakistan	Thin-tailed	139 bp	(TTTTG) ₃ G(TG) ₁₃
Lohi 41	Pakistan	Thin-tailed	139 bp	(TTTTG) ₃ G(TG) ₁₃
Pedi 31	South Africa	Fat-tailed	141 bp	(TTTTG) ₃ G(TG) ₁₄
Touareg 21	Mali	Thin-tailed	141 bp	(TTTTG) ₃ G(TG) ₁₄
Yankassa 27	Nigeria	Thin-tailed	141 bp	(TTTTG) ₃ G(TG) ₁₄
Karakul 22	South Africa	Fat-tailed	143 bp	(TTTTG) ₃ G(TG) ₁₅
Bali-bali 02	Mali	Thin-tailed	143 bp	(TTTTG) ₃ G(TG) ₁₅
Balami 02	Nigeria	Thin-tailed	143 bp	(TTTTG) ₃ G(TG) ₁₅
Sabi 15	Zimbabwe	Fat-tailed	145 bp	(TTTTG) ₃ G(TG) ₁₆
Sukuma 02	Tanzania	Fat-tailed	145 bp	(TTTTG) ₃ G(TG) ₁₆
Sukuma 16	Tanzania	Fat-tailed	145 bp	(TTTTG) ₃ G(TG) ₁₆
<i>O. aries</i> African Ars10	Meadows <i>et al.</i> (2006)	Fat-tailed	131 bp	(TTTTG) ₃ G(TG) ₉
<i>O. canadensis</i> 137	Meadows <i>et al.</i> (2006)	N/A	139 bp	(TTTTG) ₄ (TG) ₁₁
<i>O. aries</i> L112	Meadows <i>et al.</i> (2006)	N/A	141 bp	(TTTTG) ₃ G(TG) ₁₄
<i>O. aries</i> LAF2	Meadows <i>et al.</i> (2006)	N/A	143 bp	(TTTTG) ₃ G(TG) ₁₅
Mouflon M4	Meadows <i>et al.</i> (2006)	N/A	143 bp	(TTTTG) ₃ G(TG) ₁₅
<i>O. aries</i> L149	Meadows <i>et al.</i> (2006)	N/A	145 bp	(TTTTG) ₃ G(TG) ₁₆
<i>O. dalli</i> 1587	Meadows <i>et al.</i> (2006)	N/A	145 bp	(TTTTG) ₄ (TG) ₁₄

4.1.5 Single nucleotide polymorphism analysis

All the 556 male individuals were screened for the SNP using *SRY3* marker. Only 136 male individuals amplified. The rest of the samples did not amplify even after DNA re-extraction. All the 136 individuals that were successfully amplified and digested had the A allele. The A SNP was detected in both the thin-tailed and fat-tailed sheep breeds of Africa as well as breeds from outside Africa. Table 7 lists all the breeds in which the A SNP was detected.

Table 7. Details of the breeds in which the A/G SNP was successfully screened.

Breed	Breed type	Country	SNP type	Total
East Africa				
Afar	Fat-tailed	Ethiopia	A	18
Arsi-bale	Fat-tailed	Ethiopia	A	1
North-West Highland	Fat-tailed	Ethiopia	A	13
Tukur	Fat-tailed	Ethiopia	A	4
Wollo	Fat-tailed	Ethiopia	A	12
Sukuma	Fat-tailed	Tanzania	A	11
Ugogo	Fat-tailed	Tanzania	A	13
Moroto	Fat-tailed	Uganda	A	3
Southern Africa				
Black-head Persian	Fat-tailed	South Africa	A	4
Pedi	Fat-tailed	South Africa	A	6
Damara	Fat-tailed	Namibia	A	4
Karakul	Fat-tailed	Namibia	A	3
Sabi	Fat-tailed	Zimbabwe	A	10
Tswana	Fat-tailed	Botswana	A	4
West Africa				
Djallonke	Thin-tailed	Nigeria	A	1
Yankassa	Thin-tailed	Nigeria	A	1
Djallonke	Thin-tailed	Senegal	A	14
Europe				
Portuguese White Merino	Thin-tailed	Portugal	A	8
Swaledale	Thin-tailed	UK	A	2
Central America				
St. Croix	Thin-tailed	US Virgin Island	A	4
Total				136

4.1.6 Haplotype

A combination of alleles at different loci leads to the formation of haplotype. This provides more information compared to the use of one locus. Combining *SRYM18* and *SRY3* data resulted in identification of three different haplotypes. It is possible that more haplotypes would have been identified if there was more success in SNP discovery using more sensitive methods such as Applied Biosystems allelic discrimination Taqman[®] assay. Table 8, shows the number of haplotypes with their respective frequencies. Since only the A allele was observed for the SNP, the differences in haplotypes were confined to the fragment size. Haplotype H6 was the most common haplotype on the African continent with a frequency of 72.5%. It had highest occurrence in the East African and southern Africa regions, with a frequency of 90.5% in the East African fat-tailed breeds and 60.0% in the fat-tailed sheep of southern Africa. However, in the West African breeds, the frequency was low, 12.5%. In non African breeds, it had a frequency of 80.0% and 25.0% in the European and the Central American breeds, respectively.

Haplotype H8 had a low frequency of 16.7% continent wide, though regionally it is most common in the West African region, with a frequency of 87.5%. It scored least in the southern and East African breeds with a frequency of 16.7% and 1.4% respectively. This haplotype was not observed in any breed outside Africa.

Haplotype H4 is the most rare of the three haplotypes on the African continent (10.8%). However, regionally it was the second most observed haplotype in the southern Africa breeds with a frequency of 23.3% and 8.1% in the East African breeds. It was not observed in the West African breeds. In breeds from outside Africa, it is common in the Central American breeds where the frequency is 75.0% and it is second most frequent haplotype in the European breeds with a frequency of 20.0%.

Table 8. Distribution of haplotypes in sheep breeds grouped by geographical regions.

Haplotype Region	Haplotype frequency			Total
	H4 145 / A	H6 143 / A	H8 141 / A	
East Africa	6	67	1	74
% frequency	8.1	90.5	1.4	100
Southern Africa	7	18	5	30
% frequency	23.3	60.0	16.7	100
West Africa	0	2	14	16
% frequency	0.0	12.5	87.5	100
Total for Africa	13	87	20	120
% frequency for Africa	10.8	72.5	16.7	100
Europe	2	8	0	10
% frequency	20.0	80.0	0	100
Central America	3	1	0	4
% frequency	75.0	25.0	0	100

4.2 DISCUSSION

4.2.1 SRY marker specificity

In animals, the *SRY* gene is located within the *MSY* region and it is only found on the Y chromosome (Graves, 2002; Skaletsky *et al.* 2003). Normal females lack the Y chromosome and by extension do not show characteristics conferred by the *SRY* gene. Therefore, molecular markers located within the *MSY* region of the Y chromosome are useful in the study of patriline as well as male introgression in livestock (Hanotte *et al.* 1997; Jobling & Tyler-Smith, 2000).

The specificity of any molecular marker is of great importance. A good molecular marker has to be specific to the extent that only the target sequences are amplified when the correct PCR conditions are applied. Y chromosome specificity was tested by amplifying the female DNA samples which acted as the negative controls (Gusmão & Carracedo, 2003). The two markers used, *SRYM18* and *SRY3*, were tested against both the male and female DNA samples and it was found that they were specific to the latter. This confirmed that they are male specific. The

microsatellite marker *SRYM18* produced fragments within the expected size range (Australian Sheep Gene Mapping Website, 2003; Meadows *et al.* 2006). Similarly, the *SRY3* SNP also produced the expected fragment of about 611 bp. The Y chromosome shares homology with the X chromosome at the tips of both the short and the long arm (Hurles & Jobling, 2001; Graves, 2002; Hellborg, 2004; Raudsepp *et al.* 2004). If these regions were amplified, a similar product would be observed in the female samples. Lack of PCR products from the female samples, further, confines the specificity of the two markers to the *MSY* region.

4.2.2 *SRYM18* microsatellite mutations

A total of 491 male individuals were genotyped successfully using the *SRYM18* microsatellite marker. Five alleles, 131 bp, 139 bp, 141 bp, 143 bp and 145 bp were observed. The *SRM18* is a compound microsatellite comprising of a pentanucleotide (TTTTG)_m and a dinucleotide (TG)_n repeat separated by a G/- indel. The amplified fragments are therefore expected to differ in size depending on the value of m, n and the presence or absence of a G in between. It has been observed by Meadows *et al.* (2006) that a fragment of 145 bp could either be lacking a G with m and n having a value of 4 and 14 respectively or could have a G with m and n having a value of 3 and 16 respectively. Similarly the 143 bp fragment could either be lacking a G with m and n having a value of 4 and 13 respectively or could be having a G with m and n having a value of 3 and 15 respectively. The 141 bp allele has a G with the value of m and n being 3 and 14 respectively while the 139 bp allele lacks a G with its m and n values being 4 and 11 respectively. The 131 bp allele has a G with m and n having a value of 3 and 9 respectively (Meadows *et al.* 2006). The 131 bp, 141 bp, 143 bp and 145 bp alleles were also observed within the African breeds and in some of the individuals from outside Africa. In addition our study detected the 139 bp in the African and Asian sheep which was only reported in *O. canadensis* a wild sheep by Meadows *et al.* (2006). Our results from the few sequenced individuals are in agreement with this observation.

The observation that individuals sharing the same fragment size have different number of repeat units is known as homoplasy. A comparison of sequence data of the 139 bp allele from Maure 16, Lohi 37, Lohi 41 and *O. canadensis* 137 (bighorn sheep) showed that there is homoplasy between species. The pentanucleotide and dinucleotide units of this allele from Maure 16, Lohi 37 and Lohi 41 were three and thirteen respectively and were separated by a G.

However, in *O. canadensis* 137, the pentanucleotide and dinucleotide units were four and eleven respectively and lack a G between them. This result illustrates that great care is needed when interpreting genotyping results of microsatellites that show homoplasy (like *SRYM18*) especially when sequence data is unavailable.

4.2.3 Y chromosome diversity within and between African sheep breeds

Most of these alleles with the exception of the 131 bp have been reported in sheep breeds residing outside Africa (Meadows *et al.* 2006). The 131 bp allele was found in the Arsi-bale sheep from Ethiopia. This unique allele could represent a breed specific allele. The low frequency (0.22%) of this allele does not allow us to conclude that it is diagnostic for the African continent. A narrowed diversity study using a larger sample size of this breed and other African breeds could provide more insight on the distribution and therefore the uniqueness of this allele.

Microsatellite analysis of the male lineage indicates that there is Y chromosome variation within and between the African sheep breeds. Variation was highest within breeds (57.0%) as compared to between breeds (28.6%). In three fat-tailed sheep (Gumuz, Sekota and Tukur) from Ethiopia only one allele type (143 bp) was observed, indicating that diversity in these breeds is limited. In contrast, Djallonke from Senegal and Maure from Mali, both of which are thin-tailed sheep from West Africa, had the highest level of diversity with three different alleles (139 bp, 141 bp and 145 bp) types each. The higher diversity in the West African sheep compared to the eastern and southern Africa sheep could be an indication of different breeding structures in their centers of origin before entry into Africa. The West African sheep and the eastern and southern Africa sheep breeds entered the African continent separately. The West African sheep breeds entered first and could have originated from a center of domestication with a higher genetic pool compared to that of the eastern and southern Africa sheep breeds. Other than these five breeds, the rest of the African breeds had two different alleles. Most of breeds from outside Africa had a single allele type except for St Croix (USA) and Swaledale (UK) in which both the 143 bp and 145 bp alleles were observed. This result interpreted alone indicates that the diversity observed within the African breeds is slightly higher than that observed in breeds from outside Africa. However, results obtained from a parallel study (Meadows *et al.* 2006) using more European breeds show that they have almost the same number of different types of allele with a similar distribution as is the case with the African breeds in this study. Comparing these results, the Y

chromosome diversity for the African breeds appears similar to that observed in breeds from outside Africa (Meadows *et al.* 2006). The low diversity observed in samples from outside Africa in this study was most probably as a result of the small sample size used.

Selection and breeding based on a few individuals could lead to fixation of the selected allele within the breed (Henson, 1992) which in turn could reduce diversity. Given the culture of African pastoralist, selection is not strictly practiced. Animals are kept for multipurpose use rather than for a single commercial objective (McCorkle, 1999) and mating within and some times between herds is not strictly controlled. Therefore, by the time these males are sold off they may have served several females. It is therefore tempting to conclude that within the African context, crossbreeding can be used to explain the high diversity observed. However, even in the specialized breeds for production where selection based on a few traits is high, their sheep Y chromosome diversity compares well with that of the African sheep.

Distribution of genetic variation between thin-tailed and fat-tailed was low (14.4%). The observed allele type among the fat-tailed sheep differs from that in the thin-tailed sheep breeds of Africa. The allele frequency difference for the 141 bp allele (47.0%) in the African thin-tailed compared to 10% in the fat-tailed is not significant ($P > 0.05$). On the other hand, the allele frequency difference between the 143 bp allele (51.3%) in the thin-tailed and (80.6%) in the fat-tailed sheep is significant ($P < 0.0001$). In addition, the presence of the 145 bp allele within the fat-tailed sheep and its absence in the thin-tailed sheep of Africa is another difference between these breed types. The thin-tailed and fat-tailed sheep of Africa are thought to have originated from two different centers of domestication, the Fertile Crescent and eastern Asia respectively (Ryder, 1984). Therefore, the slight variation may be attributed to differences in the development pattern of the two breed types. However, there is no clear cut distinctiveness between the fat-tailed and thin-tailed sheep breeds of Africa due to the sharing of alleles (141 bp and 143 bp). This could have resulted from the crossbreeding of fat-tailed with thin-tailed sheep breeds before their entry into the African continent or their subsequent crossbreeding within the African continent (Muigai, 2003).

Similarly the East African fat-tailed breeds show some variation from the fat-tailed sheep of southern Africa as demonstrated by differences in allele frequencies. In the East African fat-tailed sheep breeds, the most frequent allele is the 143 bp allele (90.2%), with the 141 bp allele and 145 bp allele observed at low frequencies (1.2% and 7.8% respectively). However, the

situation among the southern African fat-tailed sheep breeds is different. The 143 bp allele had the highest frequency (53.0%) and this was closely followed by the 141 bp allele with a frequency of 35.0%, while the 145 bp allele had a frequency of 12.0%. The frequencies for the three alleles were significantly different from each other ($P < 0.005$). Further the 131 bp and the 139 bp alleles found in the East African fat-tailed breeds were absent in the southern African fat-tailed breeds. In a recent study using autosomal microsatellite markers, Muigai (2003) concluded that the East African and the southern African fat-tailed sheep breeds share a common origin. Therefore, the absence of the 131 bp and 139 bp alleles is probably a result of region specific evolutionary processes such as genetic drift following migration of these sheep from East Africa to the southern part of the continent.

Other studies have also shown variations between sheep breeds of different regions (Muigai, 2003; Meadows *et al.* 2006). The AMOVA results show that the distribution of variation resulting from geographical groupings (East Africa, West Africa and southern Africa) is 27.9%. This variation decreased to 17.8% when the African sheep were analyzed with sheep from outside Africa (four groups; Africa, Asia, Europe and Central America). This is low compared to variation within breeds (55.3% and 49.7% for the three regions of Africa and for the four regions of the world respectively) in the same geographical grouping. It has been reported that goats are genetically more diverse than sheep but sheep show higher within-breed variation (Lenstra & Econogen Consortium, 2005). The results of this study are therefore in agreement with that observation. It could be explained by possibly a legacy of crossbreeding of sheep from different regions.

4.2.4 Origin, geographical distribution and migration of the African sheep

All domestic sheep are believed to have originated from Asia. In Meadows *et al.* (2006), the 141 bp allele, was detected in both African as well as sheep from Middle East/Central Asia and Europe. The thin-tailed sheep of Africa were most likely domesticated in the Fertile Crescent in West Asia. These sheep were the first to enter the African continent via the Isthmus of Suez into the current Egypt and spread to West Africa and along river Nile into the current Sudan (Manson, 1967; DAGRIS, 2004). The allele frequency for the 141 bp allele was significantly different ($P < 0.005$) in all the three regions of Africa. The high frequency (53.5%) of the 141 bp allele in the West African thin-tailed sheep compared to that in East and southern Africa sheep

(1.5% and 35% respectively), agrees with this migration theory. On the other hand, the fat-tailed sheep could have possibly been domesticated in eastern Asia where they could have crossbred with the thin-tailed sheep before migrating to Europe and Africa. They entered the African continent via the Horn of Africa into present day Ethiopia and spread to eastern Africa and later to southern Africa (Manson, 1967; DAGRIS, 2004). This is supported by the high frequency of the 143 bp allele in the fat-tailed sheep of East Africa (90.2%), southern Africa (53%), Asia (40%); the thin-tailed sheep of Europe (63%), and Central America (13%). Similarly in Meadows *et al.* (2006), the 143 bp allele was the most common in African, Asian and European sheep. Although the African thin-tailed and fat-tailed sheep entered the continent separately (MacDonald, 2000), following human migration and intercommunity interaction, sheep genetic admixture occurred between the two sheep types (McCorkle, 1999). The presence of the 143 bp allele (44.6%), in the West African thin-tailed sheep and the presence of the 141 bp allele (1.2%) in the East African fat-tailed sheep are likely the result of these introgression following trade between the West African and the East African communities and subsequent crossbreeding.

Similarly the southern African sheep is a mixed pool of the 141 bp allele (35%) and the 143 bp allele (53%). A similar observation, using autosomal microsatellite markers, was made by Muigai (2003) leading to the proposition of a southward migration of the West African thin-tailed sheep. Therefore, our results provide more evidence for this movement. However, introgression from recently imported breeds such as Karakul can not be ruled out. Further studies on such breeds would provide more information about their contribution.

The 145 bp allele seems to be the most recent on the African continent. It has been reported in the sheep of Europe, Australia and Middle East (Meadows *et al.* 2006). In our study, it is only found in the thin-tailed sheep of Central America and Europe (with a frequency of 87% and 37% respectively) and the fat-tailed sheep of eastern and southern Africa (with a frequency of 7.8% and 12% respectively). This could have been introduced into the continent by European settlers during colonization. The 139 bp allele was only detected in the African sheep (3%) and in the Asian (Pakistan) sheep (60%). In a study by Meadows *et al.* (2006) it was not reported in any domestic sheep. The 131 bp allele does not tell us much about the origin and migration of the African sheep since it was detected in only one Arsi-bale sheep.

Evidence from archeological, behavioural, morphological and mtDNA studies have shown that mouflon is the most likely ancestor of the domestic sheep. In the study by Meadows *et al.*

(2006) the pentanucleotide unit of *SRYM18* microsatellite in the wild sheep (thinhorn sheep and bighorn sheep) for the 145 bp, 143 bp and 139 bp allele is repeated four times and there is no nucleotide in between the pentanucleotide unit and the dinucleotide unit. In the same study the 143 bp allele in mouflon was reported as having three pentanucleotide repeat units with a G in between the pentanucleotide and dinucleotide repeat units. Our data shows that in all these alleles the pentanucleotide repeat is fixed at three and there is a G in between the pentanucleotide unit and the dinucleotide unit. This confirms that the domestic sheep share patriline with the mouflon.

4.2.5 Sheep Y chromosome SNP

All the 136 samples screened for the SNP yielded the A allele. Most of the samples did not amplify for unknown reasons, possibly due to mutations at the primer site. Comparative sequencing (Meadows *et al.* 2004; Meadows *et al.* 2006) has shown that an A/G allele exists on the ovine Y chromosome, though the G allele is very rare and so far it has not been reported in any African sheep. For a mutation to be considered a SNP the frequency of the rare allele should be at least 1%. In this study, the G allele was not detected. A similar scenario was also observed in a parallel study using a few African breeds (Meadows *et al.* 2006). The absence of the G SNP in the African samples suggest that the A allele is fixed for the African sheep. It is also possible that there is a mutation at the *Fnu4HI* restriction site. The presence of such a mutation within the larger African breeds could in itself be useful for future analysis.

4.2.6 Sheep Y chromosome haplotypes

A total of eleven haplotypes have been reported in sheep (Meadows *et al.* 2006) basing on the two loci. The five haplotypes reported in the African and European samples were, H4 (145 bp / A), H6 (143 bp / A), H8 (141 bp / A), H9 (131 bp / A) and H10 (110 bp / A). On the contrary, haplotype H5 (145 bp / G) and haplotype H7 (143 / G) reported by Meadows *et al.* (2006), were not observed in the African breeds studied (Arsi-bale, Balami, Pedi, Sabi, Sukuma, Tswana and West African dwarf sheep). In our case, the haplotypes detected were, H4, H6 and H8. Though sequence data and therefore the pattern and number of both pentanucleotide and dinucleotide repeat units is still lacking in some samples, our results compare well with the

findings of Meadows *et al.* (2006). Haplotype H9 and H10 were not detected in our study, though they were each reported in one individual in the African reference breeds (Arsi-bale and Balami respectively) studied by Meadows *et al.* (2006).

The three haplotypes observed in the African sheep seem to vary within the breeds, breed types as well as between regions. Haplotype H6 is the most widely distributed haplotype among the fat-tailed sheep of East Africa (90.5%) and southern Africa 60.0%. While haplotype H4 only appeared in the fat-tailed sheep with a high frequency (23.3%) in the southern African sheep. H8 (87.5%) was most common in the West Africa thin-tailed sheep. These results are in agreement with archeological information suggesting a distinct origin for the West Africa thin-tailed and the East and southern Africa fat-tailed sheep.

Basing on the data that we have, we confirm that haplotype H6 is the most common in the sheep of Africa and therefore, most likely the African sheep share a patriline with mouflon, a wild feral. Haplotype H5 and H7 were reported to define a different paternal lineage and were missing in the African breeds (Meadows *et al.* 2006). We also failed to detect the two haplotypes which have been associated with the European sheep due to lack of the G allele.

4.2.7 Usefulness of the Y markers in sheep diversity studies

The analysis of Y microsatellite marker was insightful for establishing the diversity within breeds as well as breed types and breeding structure. Similarly, this marker provided some information necessary for establishing detailed relationship and possible origin of the African sheep breeds. Similar observations have been made in sheep studies using autosomal microsatellite markers (Muigai, 2003), mtDNA (Meadows *et al.* 2005) and Y chromosome markers (Meadows *et al.* 2004; 2006). However, the SNP marker and haplotypes did not provide any new information. Haplotypes provide more information compared to the use of one marker (Bradman & Thomas, 1998; Meadows *et al.* 2006). More information could have been obtained if the SNP markers had higher success of amplification.

The use of Y chromosome markers in livestock diversity studies is still rare as very few markers have been described. Mapping of the ovine Y chromosome is likely to open up avenues for the discovery of new male specific markers that could help in the construction of patrilines as well as in diversity studies as is the case in humans. Lately, studies on the Y chromosome have become interesting and advances in the discovery of new Y specific markers are being made

(Australian Sheep Gene Mapping Website, 2003). In the near future, this is likely to generate more information that will allow better understanding of the origin, domestication and diversity of the domestic sheep.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

This study aimed at evaluating genetic diversity and relationships within and between African sheep breeds based on the polymorphisms on the Y chromosome.

Microsatellite data shows that genetic variation on the Y chromosome exists within African sheep breeds. The distribution of this diversity was highest between breeds in different geographical regions. It was low between breeds in the same geographical location, most likely due to trade and crossbreeding (McCorkle, 1999), as well as low mutation rates on the mammalian Y chromosome (Hellborg & Ellegren, 2004; Meadows, *et al.* 2004). These results agree with an earlier study using autosomal microsatellite markers (Muigai, 2003). Diversity of the Y chromosome microsatellite was found to be higher in the African sheep compared to breeds from outside the African continent. However, when these results were compared to that by Meadows *et al.* (2006), the African sheep Y chromosomes are as diverse as that of sheep from outside Africa. It is likely that the low number of non African sheep used in this study may have contributed to the apparent low diversity observed in non African samples.

This study also aimed at determining the usefulness of these two Y chromosome specific markers for the study of the history and differences between sheep breeds.

The sheep breeds in different geographical locations of Africa are phenotypically different. This study has shown that the 141 bp allele is most common within the thin-tailed sheep. Similarly, the 143 bp and 145 bp alleles are common within the fat-tailed sheep. This provides strong evidence supporting the hypothesis that the African fat-tailed and thin-tailed sheep entered into the continent separately. The 141 bp allele is likely to be the most ancient on the African continent followed by the 143 bp and 145 bp alleles respectively. These findings are consistent with archeological studies that suggest a different origin for the West Africa thin-tailed sheep and the fat-tailed sheep of East and southern Africa. The presence of both the 141 bp and 143 bp alleles in all the three regions shows that African stockraisers may have used crossbreeding strategies to improve their sheep stock as proposed by McCorkle (1999).

Characterization of the A/G SNP described in sheep produced only the A allele. No polymorphisms were observed at the SNP and therefore no new information was obtained

through the study of this marker compared to the microsatellite marker. In other words, haplotypes combining the microsatellite locus and the SNP, had frequencies identical to the microsatellite allele frequencies. However, it should be noted that we failed to amplify the *SRY3* loci in most individuals. From these results, despite the fact that the A SNP allele was the only one observed, it is premature to conclude that it is the only one present on the African continent. Using data obtained in our study, we can confirm the presence of haplotype, H4, H6 and H8 in African sheep. The H6 haplotype has been associated with the European mouflon. Since this haplotype is also common in the sheep of Africa, it is therefore most likely that the African sheep share a patriline with the mouflon (*Ovis musimon*). The African and Asian domestic sheep share the 139 bp allele which is lacking in sheep from Europe and Central America; suggestive that the African and Asian sheep share a patriline.

This study has contributed to the understanding of the origin, migration, diversity and breeding structure of the African sheep. Studies of the Y chromosome polymorphism in livestock are still rare. There is lack of informative markers at the Y chromosome. However, in the recent past interest in Y chromosome studies has been rekindled by the development of new markers (Australian Sheep Gene Mapping Website, 2003; Meadows, *et al.* 2004). The results presented in this study are an initial attempt for the use of Y chromosome markers in the genetic characterization of the African sheep. New information generated from future genetic and phenotypic studies, together with these and previous genetic and phenotypic findings if harmonized will be invaluable for breeding, conservation and utilization of the African sheep.

5.2 Recommendations

In order to have a better understanding of the origin, history and development of the breeds of sheep present in Africa today, further work is needed to address new questions raised in this study as follows.

- 1) Although this work was continent wide, it could not cover all the African breeds due to political, technical, economical and intellectual property issues. In order to generate more information regarding history and diversity of the African sheep, an in-depth country analysis using large sample sizes is required. Samples from North Africa were unavailable and it would be interesting to see how they relate with the rest of the

continent. In a more focused approach, it will be interesting to unravel the distribution of the 131 bp allele, this will provide important information about its origin.

- 2) This study has confirmed that the *SRYM18* is a compound microsatellite that exhibits homoplasy between species. To be able to assess if homoplasy within the domestic African sheep is also present, sequencing of alleles from individuals of different breeds covering different geographical regions is required. The 139 bp allele was only observed in three samples from Africa, the fat-tailed Wollo sheep from Ethiopia and the thin-tailed Maure and Djallonke sheep from Mali and Senegal, respectively. More sampling and analysis is required from these three breeds so as to understand the distribution of this allele and further confirm the nature of the repeat arrays.
- 3) The screening of the A/G SNP allele was only partly successful. Further work is required using a primer set that spans the *SRY3* flanking region to determine whether amplification failure was due to a mutation within the *SRY3* flanking region. Alternatively, a more sensitive but expensive SNP screening method using Applied Biosystems allelic discrimination Taqman[®] assay could be considered.
- 4) In this study, only domestic sheep breeds of Africa, and a few breeds from outside Africa were considered. It was not possible to obtain DNA from, the presumptive wild ancestors of sheep (urial, mouflon and argali), though alleles present in the Mouflon have been observed elsewhere. The results presented here show that more research needs to be done to unravel the wild origin of the different alleles observed in *SRYM18* microsatellite marker.
- 5) As demonstrated in this study and in Meadows *et al.* (2006), the use of these two markers did not reveal clear Y chromosome differentiation between breeds. Therefore, more work is needed to identify more Y specific markers to allow a better understanding of the evolutionary relationships of the Y chromosome of domestic sheep breeds.

REFERENCES

- Aitken, J. R. and Graves, J. A. M. (2002).** The future of sex. *Nature*, **415**: 963.
- Ali, B.A. (2003).** Genetic similarity among four breeds of sheep in Egypt detected by random amplified polymorphic DNA markers. *Afric. J. Biotech.* **7**: 94-197.
- Australian Sheep Gene Mapping Website. (2003).** Sheep markers: *SRYM18*.
<http://rubens.its.unimelb.edu.au/~jillm/jill.htm>. Last modified, 30th January, 2003.
 Accessed, 1st April, 2005.
- Bachtrog, D. and Charlesworth, B. (2002).** Reduced adaptation of a non-recombining neo-Y chromosome. *Nature*, **416**: 323-325.
- Bachtrog, D. (2004).** Evidence that positive selection drives Y-chromosome degeneration in *Drosophila miranda*. *Nature Genet.* **36**: 518-522.
- Barker, J.S.F., Tan, S.G., Moore, S.S., Mukherjee, T.K., Matherson, J.L. and Selvaraj, O.S. (2001).** Genetic variation within and relationships among populations of Asian goats (*Capra hircus*). *J. Anim. Breed. Genet.* **118**: 213-233.
- Bell, J. S. (2003).** The popular sire syndrome and concerns of genetic diversity. A report on the proceedings of the 2003 American Kennel Club (AKC), Canine Health Foundation (CHC) National Parent Club (NPC) Canine Health Conference (CHC). St. Louis Missouri, USA.
- Bjørnstad, G. and Røed, K.H. (2001).** Breed demarcation and potential for breed allocation of horses assessed by microsatellite markers. *Anim. Genet.* **32**: 59-65.
- Blench, R. (2001).** "til the cow comes home" why conserve livestock biodiversity? In: Living off Biodiversity: Exploring Livelihoods and Biodiversity Issues in Natural Resources Management. (Koziell, I. & Saunders, J. eds.) International Institute for Environment and Development (IIED), London.
- Botstein, D., White, R.L., Skolnick, M., Davis, R.W. (1980).** Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**: 314-331.
- Bradley, D.G., MacHugh, D.E., Cunningham, P. and Loftus, T.R. (1996).** Mitochondrial diversity and the origins of African and European cattle. *Proc. Natl. Acad. Sci. USA.* **93**: 5131-5135.
- Bradman, N. and Thomas, M. (1998).** Why Y? The chromosome in the study of human evolution, migration and prehistory. *Sci. Spectra*, **No. 14**.
- Brohede, J. (2003).** Rates and patterns of mutation in microsatellite DNA. PhD thesis, Uppsala University, Uppsala, Sweden.

- Buchanan, F.C., Adams, L.J., Littlejohn, J.F., Maddox, J.F. and Crawford, A.M. (1994).** Determination of evolutionary relationships among sheep breeds using microsatellites. *Genome*, **22**: 397-403.
- Butler, J.M. (2003).** Recent development in Y-short tandem repeat and Y-single nucleotide polymorphism analysis. *Foren. Sci. Rev.* **15**: 91-111.
- Chenyambuga, S. W. (2002).** Genetic characterization of indigenous goat populations of sub-Saharan Africa using microsatellite DNA markers. PhD thesis, Sokoine University of Agriculture, Morogoro, Tanzania.
- Cunningham, E.P. (1995).** Global Impact Domain - Animal Genetic Resources. Working document livestock and the environment: Finding a balance. FAO/World Bank/USAID. Rome.
- DAGRIS (2004).** Domestic Animal Genetic Resources Information System (DAGRIS). (Rege, J.E.O., Ayalew, W. and Getahun, E. eds.). International Livestock Research Institute, Addis Ababa, Ethiopia. <http://dagris.ilri.cgiar.org>. Last modified, 2004. Accessed, 1st April, 2005.
- Excoffier, L., Laval, G. and Schneider, S. (2005).** Arlequin V 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics online*, **1**: 47-50.
- FAO (2004a).** Food and Agriculture Organisation (FAO). The statistics division, economics and social department. Global livestock production and health atlas. Animal population and health division.
- FAO (2004b).** Food and Agriculture Organisation (FAO). Secondary guidelines for development of national animal genetic resources management plans. Measurement of domestic animal diversity (MoDAD): Original working group report.
- FAO (2005).** Food and Agriculture Organisation (FAO). Pastoralism in the new millennium: maintaining livestock biodiversity. FAO animal production and health; Paper **150**.
- Graves, J. A. M. (2002).** The rise and fall of *SRY*. *Trend Genet.* **18(5)**: 259-264.
- Graves, J. A. M. (2005).** Recycling the Y chromosome. *Science*, **307**: 50-51.
- Guo, J., Du, L.-X., Ma, Y. -H., Guan, W.-J., Li, H.-B., Zhao, Q.-J., Li, X. and Rao S.-Q. (2005).** A novel maternal lineage revealed in sheep (*Ovis aries*). *Animal Genet.* **36**: 331-336.
- Gusmão, L. and Carracedo, A. (2003).** Y chromosome – specific STRs. *Profiles DNA*. **6(1)**: 3-6.

- Hall, S.J.G. and Bradley, D.G. (1995).** Conserving livestock breed biodiversity. *Trends Ecol. Evol.* **10**: 267-270.
- Hanotte, O., Okomo, M., Verjee, Y., Rege, E. and Teal, A. (1997).** A polymorphic Y chromosome microsatellite locus in cattle. *Anim. Genet.* **28**: 318-319.
- Hanotte, O., Tawah, C.L., Bradley, D.G., Okomo, M., Verjee, Y., Ochieng, J. and Rege, J.E.O. (2000).** Geographical distribution and frequency of a taurine *Bos taurus* and an indicine *Bos indicus* Y specific allele among sub-Saharan African cattle breeds. *Mol. Ecol.* **9**: 387-396.
- Hartl, D. L. and Clark, A. G. (1997).** Principles of population genetics. 3rd ed. Sinauer Associates, Inc. Sunderland, pp 20-21.
- Hellborg, L. (2004).** Evolutionary studies of the mammalian Y chromosome. PhD thesis, Uppsala University, Uppsala, Sweden.
- Hellborg, L. and Ellegren, H. (2004).** Low level of nucleotide diversity in mammalian Y-chromosomes. *Mol. Evol.* **21**: 158-163.
- Henson, E. L. (1992).** *In situ* conservation of livestock and poultry. FAO animal production and health: paper **99**. FAO and UNEP. Rome.
- Hiendleder, S., Mainz, K., Plante, Y. and Lewalski, H. (1998a).** Analysis of mitochondrial DNA indicate that domestic sheep are derived from two different ancestral maternal sources: No evidence for contributions from urial and argali sheep. *J. Hered.* **89**: 113-120.
- Hiendleder, S., Lewalski, H., Wassmuth, R. and Janke, A. (1998b).** The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major ovine haplotype. *J. Mol. Evol.* **47**: 441-448.
- Hiendleder, S., Kaupe, B., Wassmuth, R. and Janke, A. (2002).** Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proc. R. Soc. Lond. B* **269**: 893-904.
- Hurles, M.E. and Jobling, M.A. (2001).** Haploid chromosomes in molecular ecology: lessons from the human Y. *Mol. Ecol.* **10**: 1599-1613.
- Jobling, M. A. and Tyler-Smith, C. (2000).** New uses for new haplotypes; the human Y chromosome, disease and selection. *Trends Genet.* **16 (8)**: 356-362.
- Kikkawa, Y., Takada, T., Sutopo, K., Nomura, K., Namikawa, T., Tonekawa, H. and Amon, T. (2003).** Phylogenies using mtDNA and *SRY* provide evidence for male-mediated introgression in Asian domestic cattle. *Anim. Genet.* **34**: 96-101.

- Lebbie, S.H.B., Yapi-Gnaore, C.V., Rege, J.E.O. and Baker, R.L. (1996).** Development in the management of small ruminant genetic resources. In: *Genetic Resource: Sub-Saharan Africa* (Devendra, C. ed.), FAO, Bangkok, Thailand.
- Lenstra, A. J. and Econogene Consortium. (2005).** Evolutionary and demographic history of sheep and goats suggested by nuclear, mtDNA and Y-chromosome markers. In: The proceedings of the International workshop on the role of biotechnology for the characterization and conservation of crop, forestry, animal and fishery genetic resources. Turin, Italy. pp 97-100.
- Liu, Z.J. and Cordes, J.F. (2004).** DNA marker techniques and their application in aquaculture genetics. *Aquaculture*, **238**: 1-37.
- MacDonald K. C. (2000).** The Origins of African Livestock: Indigenous or imported? In: The Origin and Development of African Livestock: Archeology, Genetics, Linguistics and Ethnography (Blench R.M. and MacDonald K.C. eds.), University College London Press, London. pp 2-17.
- MacHugh, D.E., Shriver, M.D., Loftus, R.T., Cunningham, P. and Bradley, G.D. (1997).** Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics*, **146**: 1071-1086.
- Manson, I. L. (1967).** Sheep breeds in Mediterranean. Commonwealth Agricultural Bureau, Farnham Royal, UK.
- Mason, I. L. and Maule, J. P. (1960).** The indigenous livestock of eastern and southern Africa, pp 84-85.
- Mburu, D.N., Ochieng, J.W., Kuria, S.G., Jianlin, H., Kaufmann, B., Rege, J.E.O. and Hanotte, O. (2002).** Genetic diversity and relationships of indigenous Kenyan camel (*Camelus dromedarius*) populations: Implications for their classification. *Anim. Genet.* **34**: 26-32.
- McCorkle, C.M. (1999).** Africans manage livestock diversity. *COMPAS Newsletter*. October. No. **2**: 14-15.
- Meadows, J.R.S., Hawken, R. J. and Kijas, J.W. (2004).** Nucleotide diversity on the ovine Y chromosome. *Anim. Genet.* **35**: 379-385.
- Meadows, J.R.S., Li, K., Kantanen, J., Tapio, M., Sipos, W., Pardeshi, V., Gupta, V., Calvo, J.H., Whan, V., Norris, B. and Kijas, W.J. (2005).** Mitochondrial sequence reveals high levels of gene flow between breeds of domestic sheep from Asia and Europe. *J. Hered.* **96** (5): 494-501.
- Meadows, J.R.S., Hanotte, O., Drögemüller, C., Calvo, J., Godfrey, R., Coltman, D., Maddox, J. F., Marzanov, N., Kantanen, J. and Kijas, W.J. (2006).** Globally dispersed

- Y chromosomal haplotypes in wild and domestic sheep. *Published online in Anim. Genet.* **37 (5)**: 444-453.
- Muigai, A.T.W. (2003).** Characterisation and conservation of indigenous animal genetic resources: Genetic diversity and relationships of fat-tailed and thin-tailed sheep of Africa, PhD thesis, Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya.
- Pedrosa, S., Uzun, M., Arranz, J.J., Gutiérrez-Gil, B., Primitivo, F. S. and Bayón, Y. (2005).** Evidence of three maternal lineages in near eastern sheep supporting multiple domestication events. *Proc. R. Soc. Lon. B*, 1-7.
- Raudsepp, T., Santani, A., Wallner, B., Kata, R. S., Chengwei, R., Zhang, H.-B., Womack, J. E., Skow, C. L. and Chowdhary, B. P. (2004).** A detailed physical map of the horse Y chromosome. *Proc. Natl. Acad. Sci. USA.* **101(25)**: 9321-9326.
- Rege, J.E.O., Yapi-Gnaore, C.V. and Tawah, C.L. (1996).** The indigenous domestic ruminant genetic resources of Africa. In: "2nd All Africa Conference on Animal Agriculture, Pretoria, South Africa.
- Ryder, M.L. (1984).** *Sheep*. In: Evolution of domesticated animals. (Mason, S.L. ed.) Longman. pp 63-85.
- Sambrook, J., Fritsch E. F. and Maniatis, T. (1989).** Molecular cloning. 2nd ed. **Vol. 1.** Cold Spring Harbor Laboratory Press. New York.
- Schearf, B. (2003).** World Watch List for Domestic Animal Diversity. FAO.
- Schlötterer, C. (1998).** Microsatellites. In: molecular genetic analysis of populations. 2nd ed. (Hoelzel, A. R. ed.). Oxford university press. Oxford. pp 237-261.
- Sinclair, A. H., Berta, P. and Palmer, M. S. (1990).** A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motive. *Nature*, **346**: 240-244.
- Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P.J., Cordum, H. S., Hillier, L., Brown, L. G., Repping, S., Pyntikova, T., Ali, J., Bieri, T., Chinwalla, A., Delehaunty, A., Delehaunty, K., Du, H., Fewell, G., Fulton, L., Fulton, R., Graves, T., Hou, S-F., Latrielle, P., Leonard, S., Mardis, E., Maupin, R., McPherson, J., Miner, T., Nash, W., Nguyen, C., Ozersky, P., Pepin, K., Rock, S., Rohlffing, T., Scott, K., Schultz, B., Strong, C., Tin-Wollam, A., Yang, S.-P., Waterston, R. H., Wilson, R. K., Rozen, S. and Page, D. C. (2003).** The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature*, **423**: 825-837.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997).** The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **24**: 4876-4882.

- Vignal, A., Milan, D., SanCristobal, M. and Eggen, A. (2002).** A review on SNP and other types of molecular markers and their use in animal genetics. *Genet. Sel. Evol.* 34: 275-305.
- Wallner, B., Piumi, F., Brem, G., Müller, M. and Achmann., R. (2004).** Isolation of Y Chromosome- specific microsatellites in the horse and cross-species amplification in the genus *Equus*. *J. Hered.* **95**:158-164.
- Welsh, J. and McClelland, M. (1990).** Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* **18**: 7213-7218.
- Wilson, R.T. (1991).** Small Ruminant Production and the Small Ruminant Genetic Resource in Tropical Africa, FAO, UN, Rome.
- Xuebin, Q., Jianlin, H., Lkhagva, B., Chekarova, I., Badamdorj, D., Rege, J.E.O. and Hanotte, O. (2005).** Genetic diversity and differentiation of Mongolian and Russian yak populations. *J. Anim. Breed. Genet.* **122 (2)**: 117-126.

GERTON UNIVERSITY LIBRARY