GENETIC CHARACTERISATION OF INDIGENOUS GOAT POPULATIONS OF SUB-SAHARAN AFRICA USING MICROSATELLITE DNA MARKERS

BY

SEBASTIAN WILSON CHENYAMBUGA

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY OF SOKOINE UNIVERSITY OF AGRICULTURE, MOROGORO, TANZANIA.

2002

ABSTRACT

Two studies were carried out to genetically characterize the sub-Saharan African goats using 19 microsatellite markers. The aims were to understand the genetic uniqueness and relationships among breeds and to quantify the level of genetic diversity within each breed. The first study investigated the genetic relationships among the major goat types in sub-Saharan Africa. Ten African breeds were sampled from eastern Africa (Maasai, Kigezi, Mubende, North West Highland, Arsi-Bale), southern Africa (Ndebele, Pafuri) and West Africa (West African Dwarf, Maure, Djallonke). Two European breeds (Grisons Striped, Toggenburg), two Asian breeds (Mongolian Cashmere, Bandipur) and one Middle East breed (Arab) were included as reference breeds. Twenty to forty eight animals per breed were genotyped at the 19 microsatellite loci. Within breed genetic diversity was determined as the mean number of alleles per locus and average gene diversity. Two measures of population differentiation, G_{ST} and θ were computed to measure the genetic differentiation among the breeds. Breed assignment test was performed to identify the source populations of individual animals. Three measures of genetic distances, D_S, D_A, and $(\delta\mu)^2$ were used to estimate the genetic distances between pairs of breeds. These distances were used to construct neighbour-joining (NJ) trees to assess the relationships among breeds. The genetic relationships among breeds were further assessed using principal component analysis (PCA) and multidimensional scaling (MS).

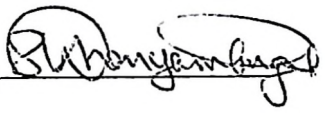
In total, 263 alleles were detected across all the breeds. Among the sub-Saharan African breeds, the mean number of alleles per locus ranged from 5.26 ± 0.464 (Djallonke) to 7.05 ± 0.516 (Mubende). The lowest and highest average gene diversities were observed in Pafuri (0.542 ± 0.036) and Ndebele (0.672 ± 0.031), respectively. For all the breeds, the within breed genetic diversities of European breeds were slightly less than those of African and Asian breeds. Between 14.6% (θ) and 15.7% (G_{ST}) of the total genetic variation was due to differences between breeds. In the assignment test, 92.1% of all animals were correctly assigned to their original population. The three measures of genetic distances indicated that the largest genetic distances were observed between West African and southern African breeds and the lowest genetic distances were found between the pairs of breeds within the same country. The $(\delta\mu)^2$ distances were found less effective in revealing the true genetic relationships among the breeds compared to D_S, and D_A distances. The NJ trees constructed from D_S, and D_A showed some differences, however, both grouped the breeds according to their geographic origins. The PCA and MS supported the grouping of the breeds according to their geographic origins. Unlike the NJ trees, the PCA and MS clearly separated the European breeds from the Asian breeds. Furthermore, the MS separated the Arab breed from the other breeds more clearly.

The second study examined the genetic diversity and structure of Small short-eared eastern African goats as a case study. The same 19 microsatellite markers were used to genotype 13 populations. The populations sampled were Afar, North East Highland, Boran, Galla, Kenyan Small East African, Maasai, Ugogo, Sukuma, Ujiji, Tanzanian Coastal, Newala, Mbeya and Landim. The reference breeds were Tswana, Venda, West African Dwarf, Red Sokoto and Toggenburg. The Galla had the lowest (5.53 ± 0.599) mean number of alleles per locus while Afar had the highest (6.53 ± 0.646). The level of gene diversity ranged from 0.667 ± 0.035 (Afar) to 0.553 ± 0.036 (Newala). Between 11% (θ) and 12% (G_{ST}) of the total genetic variability could be attributed to differences among the populations. In the assignment test, 79.1% of the individuals were correctly assigned to their source populations. NJ tree using individual animals as the taxonomic units was constructed to study the structure of eastern African goats. Among the Tanzanian populations, only the Newala appeared to be tightly clustered together, the rest of the populations were mixed together. Among the Kenyan and Ethiopian populations, only the Afar and Kenyan Small East African featured in their distinct clusters, the rest of the populations showed unclear pattern of clustering. The NJ trees based on D_S, and D_A distances, though differed in their topology, clearly marked the separation of the Tanzanian populations from the Ethiopian and Kenyan populations. The Landim was grouped together with the Tswana and Venda goats from southern Africa. The MS supported the separation of the Tanzanian populations from the Ethiopian-Kenyan populations. However, the PCA indicated that the eastern African goat populations are genetically close to each other.

It was concluded that microsatellite analysis is very useful in examining the genetic relationships of closely related populations of sub-Saharan African goats. The relationships of the breeds were according to their geographic origins implying that the goats of eastern Africa, West Africa and southern Africa are substantially differentiated from each other. This shows the necessity of having separate conservation programmes for each region.

DECLARATION

I, Sebastian Wilson Chenyambuga, do hereby declare to the Senate of Sokoine University of Agriculture that this thesis is my own original work and that it has not been submitted for a degree award in any other University.

Signature 

Date 25/3/2002

COPYRIGHT

ACKNOWLEDGEMENTS

I offer my sincere thanks to the SUA-MU ENRECA project staff for giving me the opportunity to undertake a PhD study in the section of Animal Breeding and Genetics. Many thanks to Dr. G.C. Kifaro and Prof. P.S. Gwakisa who gave me useful guidance and advice at the time of preparing the proposal. The financial supports granted to me by the SUA-MU ENRECA project (DANIDA sponsored project at Sokoine University of Agriculture), ILRI, through the Small Ruminant Biodiversity project (DFID sponsored project at ILRI) and the Association of Commonwealth Universities Scholarship Commission are highly appreciated.

A number of individuals provided assistance during the conduct of the experiment and preparation of the manuscript and I need to acknowledge them. The list below may not be exhaustive, but the help of all individuals who contributed to the success of this work, in one way or another, is highly appreciated.

I am indeed grateful to Dr. O. Hanotte of ILRI, Kenya and Dr. J.E.O. Rege of ILRI, Ethiopia, not only for collecting samples for the breeds from outside Tanzania and supervising this work while I was working in the laboratory, but also for correcting and offering invaluable suggestions regarding the text. I am highly indebted to Prof. S.J. Kemp and Dr. P.C. Watts who guided my first step at the University of Liverpool and offered me a good environment to accomplish the experiment for the second part of my

study. Their untiring encouragement and fertile suggestions are appreciated. The assistance rendered by Dr. P.C.Watts for genotyping six markers for the study of eastern African goats is highly acknowledged. I owe many thanks to my supervisors at Sokoine University of Agriculture, Dr. G.C. Kifaro and Prof. P.S. Gwakisa, not only for correcting and guiding the write up of the thesis, but also for making suggestions to improve it.

Special thanks to all members of the Animal genetics laboratory at the University of Liverpool, namely, Derek, Anthea, Harry, Jian and Angel, to name a few. I enjoyed their company. I would like to express my deep gratitude to the technical staff of lab. 7 at ILRI. Thanks to J. Hirbo who assisted in genotyping some of the breeds. Thanks are recorded to J. Mwakaya, J.Ochieng, M. Ogugo, J. Wambugu and T. Mbulo, to mention a few. Without their kind cooperation, the conduct of this work would have been difficult. Thanks are due to my fellow PhD students at ILRI, namely, Anne Muigai, Simon Kanga, David Menge and Joseph Nganga. Their company is highly appreciated. I would like to sincerely thank all member of staff and my fellow PhD students in the Department of Animal Science & Production, Sokoine University of Agriculture, for their cooperation and moral support.

This list would not be complete without a mention of my family, my wife and children in particular, who gave me the time to engage in this demanding work. Hopefully, the final product justifies their sacrifices.

DEDICATION

To my wife, my parents and my children who gave me the gifts of time and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS

A –	Adenine
AD –	Anno Domini (years after the birth of Christ)
B.C. –	Before Christ
bp –	Base pair
B.P. –	Before present
C -	Cytosine
^{O}C –	degree centigrade
D_A –	Nei's *et al.* (1983) Angular genetic distance
DANIDA –	Danish International Development Agency
DFID –	Department for International Development in UK
DNA –	Deoxyribonucleic acid
dNTP –	Deoxynucleoside triphosphate
dATP –	Deoxyadenosine triphosphate
dCTP –	Deoxycytidine triphosphate
dGTP –	Deoxyguanosine triphosphate
D_S –	Nei's standard genetic distance
dTTP –	Deoxythymidine triphosphate
$(\delta\mu)^2$ -	Goldstein *et al.*'s (1995b) genetic distance
EDTA –	Ethylene diamine tetra-acetate
ENRECA –	Enhancement of research capacities in developing countries

f –	Inbreeding coefficient
FAO –	Food and Agriculture Organization of the United Nations
F_{IS} –	Inbreeding coefficient which measures the reduction in heterozygosity of an individual due to non-random mating
F_{ST} –	Standardised variance in allele frequencies among populations
g –	gram
G -	Guanine
G_{ST} –	The coefficient of gene differentiation
ILRI –	International Livestock Research Institute
h –	Hour(s)
kb –	Kilobase pair(s)
km –	Kilometre (s)
M –	Molar
ml –	Millilitre(s)
mM –	Millimolar
min –	Minute(s)
ng –	Nanogram
PC –	Principal component
θ -	The correlation of genes between individuals in the same population
rpm –	Revolutions per minute
SUA-MU –	Sokoine University of Agriculture – Makerere University

T –	Thymine
Taq –	*Thermus aquaticus*
Tris –	Trizma® base
µg –	microgram
µl -	microlitre

CHAPTER 1

1.0 INTRODUCTION

The population of goats in sub-Saharan Africa is estimated at 163 millions (FAO, 1996). Indigenous goats constitute over 95% of the total goat population. This reflects the importance of the local genetic resources to the people of this region. Goats play a big role in providing meat and cash income and to a small extent milk to smallholder farmers. In addition, goats supply manure and are important for a number of social functions such as secure form of investment, payment of bride price, religious sacrifices and ceremonial functions. In sub-Saharan Africa, goats are functionally integrated into farming systems and they are the current account or working capital as opposed to cattle which are often considered as equity investment for rural farmers. Furthermore, goats have special advantages over cattle in that their adaptive features such as feeding behaviour, efficient feed utilisation and, in part, their overall hardiness enable them to effectively cope with the stressful nature, even of marginal lands. Compared to sheep, goats are found in larger number, have a higher reproductive rate, have got a wider ecological distribution and have higher off-take rate (Winrock International, 1983). Thus, indigenous goats play an important role in the livelihood of many people of sub-Saharan Africa, especially those living in areas with low agricultural potential.

Indigenous goats are at risk of loss due to indiscriminate slaughter, neglect, inappropriate husbandry practices and continuous crossbreeding with exotic breeds (Peters, 1987). This general trend has led to a global concern for the conservation of indigenous breeds, a concern that is motivated by the fear that the genetic resources of indigenous breeds, and particularly the complex of traits adapted to climatic and environmental stress, may be lost (Hammond, 1994). Indigenous breeds of sub-Saharan Africa have a high degree of heat tolerance, are resistant or tolerant to many diseases prevalent in the tropics and have the ability to survive long periods of feed and water shortages. These attributes, mostly likely with strong genetic bases, are all essential for successful animal production in Africa (Rege, 1992).

In many African countries south of Sahara, indigenous breeds have not been characterised, hence, there is a risk of losing the local types through uncontrolled and indiscriminate crossbreeding, before their special and perhaps valuable characteristics are identified and documented. Characterisation of African goat breeds based on morphological characters has been reported by Mason and Maule (1960), Epstein (1971), Mason (1981b), and recently by Rege *et al.* (1996). These authors divide the goats of sub-Saharan Africa into five major groups: (i) The Long lop-eared goats of North-east Africa which includes the Sudanese Nubian and Sudanese Desert found mainly in northern Sudan, the Shukria of Eritrea and Ethiopia and the Benadir of Somalia; (ii) the Small short-eared goats which extend from Ethiopia and Somalia through Kenya up to Zimbabwe and some parts of

South Africa; (iii) the Dwarf short-eared goats which are found along the coast of West and Central Africa; (iv) the Sahelian and Intermediate goats seen in the Savannah belt and (v) the Southern lop-eared goats in southern Africa.

Although the indigenous goats of sub-Saharan Africa have been broadly classified into the above groups, the distinction of many geographically isolated populations into breeds/strains has been of controversy. Often populations are named after locations or communities keeping them, for instances, Boran and Maasai goats are kept by Borana and Maasai people, respectively. Consequently, phenotypically different populations have the same names in some cases and phenotypically similar ones have different names in other cases (Rege *et al.*, 1996). The considerable variation observed among the goat populations, specially, with regard to size and coat colours and pattern has led to some inconsistencies in the classification of the various local populations into breeds. Furthermore, the grouping of the goats of sub-Saharan Africa into long lop-eared and small short-eared is controversial as the lop-eared and short-eared are often mixed together (Epstein, 1971). Nevertheless, it is true that, given the wide geographic area and the diversified climate and topography of the region, local populations in different places might have been isolated over a long period while also being subjected to varying selection pressures and genetic drift, thus have become genetically divergent. However, the extent to which various populations are different from each other is not clear. It is not known whether the goat populations found in different parts of the region represent distinct entities (breeds or

strains) with the same names or they are genetically similar but have been given different names in different places. The lack of knowledge on the distinctiveness of the breeds/populations has resulted in inefficient utilisation of this important genetic resource. This is because it is difficult to choose the appropriate breeds/populations, among the many available, on which genetic improvement and conservation programmes can be based.

There is a need to genetically characterise these animals in order to determine the genetic diversity and describe the relationship among the different populations. This is important for formulation of conservation programmes and development of improvement activities for meat and milk so as to improve the livelihoods of the people. The primary measure should be to determine the magnitude of genetic differentiation and relationships among the populations. One way of doing this is to determine genetic distances between pairs of breeds/populations. The computation of genetic distance requires estimation of allele frequency data using either protein markers or DNA markers. However, techniques for the determination of polymorphism of protein products lack the power to resolve the differences between closely related breeds since a great deal of genetic variations remains undetected by these methods (Meghen, *et al.*, 1994; Dowling *et al.* 1996).

In recent years, a range of innovations in molecular genetics have been developed for the study of genetic variation and evolution of populations using DNA marker genotype

information. One of the recent DNA markers based on the polymerase chain reaction (PCR) are microsatellites. The application of microsatellite markers is currently believed to be useful in the analysis of genetic diversity as they are numerous, randomly distributed in the genome, highly polymorphic and they show codominant inheritance (Ellegren, 1993). This makes it possible to study polymorphism of closely related populations. Microsatellites, as genetic markers, have been applied successfully in the study of genetic variation between and among European and African cattle breeds (MacHugh *et al.*, 1997; Okomo *et al.*, 1998). Microsatellite DNA variation has also been employed to assess genetic distances within and between sheep breeds (Buchanan *et al.*, 1994; Arranz *et al.*, 1998). Unfortunately, for goats, throughout the world, very limited work has been done to determine genetic variation among breeds/populations and, in fact, information on microsatellite polymorphism of African breeds is completely lacking.

The general objective of this study was to quantify genetic diversity within and among sub-Saharan African goat populations and to clarify the genetic relationships among the breeds/populations found in the region.

Specific objectives were:

1) To test whether the distribution of alleles (allele frequencies) in sub-Saharan African goat populations deviate from Hardy-Weinberg equilibrium.

2) To estimate within-population genetic diversity of selected breeds/populations from sub-Saharan Africa and to compare it with the genetic diversity of breeds from outside Africa.

3) To evaluate the extent of genetic differentiation between breeds/populations of sub-Saharan Africa.

4) To estimate genetic distances among the goat breeds/populations of sub-Saharan Africa and between sub-Saharan African breeds and breeds from outside Africa.

5) To establish phylogenetic relationships among different breeds/populations of goats in sub-Saharan Africa and between sub-Saharan African breeds and breeds from outside Africa.

CHAPTER 2

2.0 LITERATURE REVIEW

2.1 Classification and domestication of goats

2.1.1 Classification

Taxonomically, goats are placed in the order *Artiodactyla*, suborder *Ruminatia*, family *Bovidae* and tribe *Caprini*. The other members of the Caprini are the sheep (*Ovis*) and the two goat-like sheep or sheep-like goats, the bharal or blue sheep of Himalayas (*Pseudois nayaur*) and the Aoudad or Barbary sheep of the Sahara (*Ammotragus lervia*). Goats are divided into two genera - *Capra* and *Hemitragus*. The latter is the tahr which has three species; *Hemitragus jayakari*, the Arabian tahr, *H. Jemlahicus*, the Himalayan tahr and *H. Hylocrius* in the Nilgiri hills of southern India. The genus *Capra* is divided into six species (Mason, 1984):

(i) *Capra aegagrus* - the bezoar or wild goat and ancestor of the domestic goat. Its distribution extends from Greece to Pakistan.

(ii) *Capra ibex* - the ibex, with subspcies in the Alps, the western Caucasus, Central Asia, the Near East and Ethiopia.

(iii) *Capra caucasica* - the west caucasian tur. This species has also been called the Kuban or west Caucasian ibex, *C.ibex severtzovi*.

(iv) *Capra cylindricornis* - the tur of eastern Caucasus.

(v) *Capra pyrenaica* - the Spanish ibex or Spanish wild goat.

(vi) *Capra falconeri* - the markhor of Afghanistan, Pakistan and Tajikistan.

All these different species or subgenera have been alleged by various authors to have contributed to the domesticated stock (Epstein, 1971). According to Mason (1984), the bezoar is the direct ancestor of the domestic goat and the markhor is said to have contributed to some breeds of Central Asia (Mason, 1981a). However, Mason (1984) argues that it is less likely that the markhor has influenced some breeds. The ibexes and turs have not been domesticated.

2.1.2 Domestication of goats

It is generally accepted that domestic goats mainly descended from the bezoar. This is because its horns have the anterior keel of domestic goats and their scimitar shape is mirrored in the horns of many modern breeds, those of the ibexes, the turs and markhor are quite different. Recently, Takada *et al.* (1997), studying the diversity of the cytochrome b gene of mitochondrial DNA (mtDNA) in Asian goats, concluded that the strongest candidate for a matriarch ancestor of domestic goats is the bezoar or pasang

(*Capra aegagrus*). The goat is the earliest domestic animal, next oldest to the dog, to be domesticated. It was probably the first ruminant to be domesticated. There are two principal reasons for this: Firstly, the wild goat was reported to be present in the regions of Southwest Asia during the time when agriculture was developing. Secondly, the goat is an extremely hardy animal, hence, could have withstood the rigours of being reduced to the state of domestication better than other ruminants. Goats were domesticated at first for meat. As a dairy animal, the goat is regarded as the oldest, even older than cattle (Devendra and Burns, 1983). This is due to its convenience for milking.

Domestication of goats is considered to have occurred in the mountainous area of western Asia between the 7th and 9 th millennium B.C. (Epstein, 1971; Devendra and Burns, 1983). According to Zeder and Hesse (2000) goats were domesticated 10,000 years before present (B.P.) in the slopes of the Zagros Mountains on the border of present-day Iran and Iraq. The evidence for time and place is based on the identification of bones from archaeological sites. However, it has been argued that the goats could have been domesticated not only in the Zagros range but also in Palestine or elsewhere in the Fertile Crescent (Mason, 1984). This is because domestic goats from succeeding centuries have been recorded in other sites in the Near East. Furthermore, phylogenetic analysis of mtDNA has revealed three divergent lineages suggesting three separate maternal origins of domestic goats (Luikart *et al.*, 2001). It is believed that goats were probably present at Cayönü in Southeast Anatolia about 7000 B.C. They were present at

Jarmo in Iraqi Kurdistan where the site is dated from about 6500 B.C. to the end of the 6 th millennium B.C. They were also present in Palestine (Beidha, Elkhitam, Jericho) in the 7 th millennium B.C. and at Sialk on the Iranian plateau as early as the 6 th millennium B.C. Domestic goats with scimitar horns like the bezoar spread over a large part of the Old World during the Neolithic period. They were represented in vase paintings from ancient Mesopotamia in the 4 th millennium B.C. Goats with twisted horns are also depicted in ancient Mesopotamia from the 4 th millennium B.C. and the small short-eared sabre-horned generalised type was occasionally illustrated in the tomb paintings of the 4 th millennium B.C. It is believed that by the 5 th millennium B.C. goats had reached Egypt and by about 3500 B.C. goats with spiral or corkscrew horns entered Egypt from the Middle East. From Egypt the goats moved to the south and west. A dwarf goat was the first true goat received by the ancient Egyptians who passed it to the rest of Africa up the Nile valley and across to Lake Chad. The dwarf goat is recorded at Shaheinab near Khartoum in the Sudan from 3300 B.C. The screw-horned lop-eared type later spread in the same directions and is illustrated from Nubia and Libya in 2650 B.C. The goats of Africa south of the Sahara are predominantly small, short-eared, short-haired and short horned. They presumably derive from the goats that spread from Egypt southwards through Sudan and Ethiopia into East African regions. Goats are noted to appear in East Africa by 4000 – 3500 years B.P. (Marshall, 2000) and in southern Africa by 4th to 7th centuries AD (Smith, 2000).

2.2 Goat genetic resources of sub-Saharan Africa

2.2.1 Classification criteria for African goats

Over the millennia, goats have been migrating with their human keepers throughout the world. Consequently, through natural and artificial selection, breeds and types, which vary greatly in appearance and performance, have evolved in different parts of the world. The indigenous breeds of Africa have developed mainly through genetic isolation and natural selection. Also African livestock keepers have played an important role in shaping the present-day populations through traditional breeding systems dating back to the original African livestock keepers. However, African goats are not classified in terms of breeds in the western sense in which a breed has animals of a more or less closed population of interbreeding animals similar in some conspicuous traits such as colour, conformation etc, with a breed society, breed standard and herd book. Breeds in Africa are largely populations/ecotypes found in different countries and the classification is based on groupings of those populations/ecotypes which seem to be similar. The names of breeds normally refer to tribe or location of the people keeping such animals. The genetic composition of these populations reflect their values to their traditional owners.

Many methods of classifying domestic goats have been used by various authors using different criteria such as ear length (Mason and Maule, 1960), geographical origin

(Devendra and Burns, 1983), climate (Epstein, 1971), function (Payne, 1990), a combination of horns and ear characters, with geographical and functional consideration as supplementary criteria (Mason, 1981b, Rege *et al.*, 1996). Although the various authors differ in details of breeds/strains, the main groups are the same in each case. They divide the goats of sub-Saharan Africa into three main groups; the long lop-eared types of the north-east and south Africa, the short-eared dwarf goats of west Africa and the small short-eared types of eastern Africa. In between, however, there are various transitional groups. For example, the lop-eared Sahelian goats in northern Africa gradually mix with the dwarf goats toward the southwest in the border of the Sahara desert and the lop-eared Nubian goats in Northern Sudan and Ethiopia give way to small short-eared goats towards the south. The lop-eared goats reappear again in the southern part of Africa.

2.2.2 Breed types

African breeds in general have been described by Epstein (1971), those in eastern and southern Africa by Mason and Maule (1960), in South Africa by Joubert (1973), and in West Africa by Mason (1951). The distribution of the major breeds of sub-Saharan Africa is shown in Figure 1. In the review under this section, African goats will be grouped according to the classification of Rege *et al.* (1996), with some modification whenever necessary. Nevertheless, the above-mentioned references (Mason and Maule,

1960; Epstein, 1971; Mason, 1981b) are the main sources of information used in the description of the following breeds.

2.2.2.1 The long lop-eared goats of Northeast Africa

This group includes the Sudanese Nubian, Sudanese Desert, Shukria, and Benadir.

i) Sudanese Nubian

The Sudanese Nubian goat is also known as Beladi. It is the typical goat in urban and riverain Sudan north of latitude 12°N. It is basically of medium size with straight or convex profile. It is usually horned in both sexes and the horns are small and curved backwards. It has long lop-ears and it is black in colour and hairs are long. It is a good milker.

ii) Sudanese Desert

This breed is a typical goat of arid areas (desert and semi-desert) of northern Sudan and south of Egypt. It is variable in size (medium to large) and both sexes are horned, the horns being long, twisted and projecting laterally. The male usually has a thick mane and beard. Ears are of medium size and pendulous. Hair is short and the coat colour is

variable ranging from white to black, but most of the goats are grey, marked with black or brown. Milk yield is much lower than the Sudanese Nubian but the females are prolific, usually producing two kids at each parturition.

iii) Shukria

This goat is similar to the Sudanese Nubian. It is found in western Eritria's lowlands and northwest Ethiopia. It is of medium size with a convex profile and curved or straight horns. The ears are long and pendulous and the males have beards and ruffs but no tassels. It is brown in colour with long hairs. The goat produces single kids but it has good milk yield.

Figure 1: Geographical distribution of the major goat breeds of sub-Saharan Africa

iv) Benadir

This is a goat breed of Webi Shibeli valley in southern Somalia. It is large in size with a good conformation and good muscling. Horns are more or less robust, pointing backwards and running parallel. Ears are large and pendulous with turned up tips. The coat is usually short and smooth but sometimes long and coarse. The coat colour is often red-spotted or black-spotted; plain white is rare.

2.2.2.2 The Intermediate Types or Short Eared Twisted Horns

This group represents goats intermediate between the West African dwarf and the Sahelian goats. Among the several types of intermediate goats, the famous and well developed ones are the Red Sokoto or Maradi and the West African long legged or Sahelian.

i) Red Sokoto or Maradi

This breed is found in northwest Nigeria and in Niger where it is known as the Chévre rousse or Maradi. The goat is of variable size (small to medium size) but generally intermediate between the Sahelian goat and the dwarf goat. It is horned in both sexes. The ears are short and oriented horizontally. Hairs are short and the coat is soft with

bright mahogany-red colour. It produces milk, but the goat is most famous for its skins, which are the raw material for Moroccan leather. It is prolific, sometimes producing twins and triplets, even quadruplets. The major drawback of the goat is its lack of hardiness. It is not suited as a range goat and is not capable of long treks.

ii) Sahelian or West African long legged

Sahel goats are found in the Savannah belt lying between the tropical forest of West Africa and the desert of northwest Africa. This group of goats is of larger size and has long legs. Facial profile is straight or slightly convex. Horns are thin, flat and twisted in males and sickle-shaped in females. Sometimes males have long twisted horns projecting upwards. Beard and tassels are common. Ears are usually short and horizontal, sometimes moderately long and pendulous. Hair coat is short and fine. In males, the forehead is occasionally covered with a long forelock, while tufts of long hair extend over the shoulder and outer surface of the thighs. The commonest colours are white, black and brown, usually combined into a pied or tri-coloured coat. The goat is adapted to the arid conditions of sub-Saharan savannah regions. The females produce one kid per parturition but they are valued for milk, meat and skin.

2.2.2.3 The West African Dwarf (WAD) goats

This rather large and diverse population is also known as Fouta Djallon. It is widespread in the coastal countries of west and central Africa south of latitude 14°N. Similar types are also found in southern Sudan, west of Zaire, Angola and northern Namibia. Its main feature is that it is generally an achondroplastic dwarf, i.e., it has disproportionately short legs (about 40 - 50 cm long) which are often bent, a short, broad head and a pot belly on its plump body. The horns are very short such that their twist is not conspicuous, a few are polled. The ears are short to medium sized and erect or carried horizontally. The coat is short but the males often have a ridge of longer hair along the spine and beards. The most common coat colour are 'wild agout' (fawn to brown with black back line, belly and tail) or black, brown, grey or black and white or tri-coloured (white, red and black), but either pattern of colours are encountered. It is basically a meat goat but, most importantly, it is trypanotolerant. The dwarf goats have been widely exported to Europe and North America for many decades, originally for their novelty value in zoological gardens, and then as laboratory animals. More recently, they have become popular as pets in Europe and North America where different breeds have evolved such as the African pigmy and Nigerian dwarf in U.S.A., Pigmy in Britain and Dutch dwarf in The Netherlands.

2.2.2.4 The Small Short-eared goats

The major breeds grouped under Small short-eared goat type are Abyssinian, Somali (Galla) and Small East African goats.

i) Abyssinian

The Abyssinian goats are widely distributed in Ethiopia and Eritrea. They have straight to concave facial profiles. Both males and females are bearded and have ruffs and tassels are sometimes present. Horns tend to be straight with some animals being polled. Ears are short (average 14 cm). Coat length varies according to type from smooth short hair to coarse long hair. The colour is fawn, often patched with either white or brown. However, some animals are black, grey or reddish-brown with or without patches. The breed is prolific, often having twins and sometimes triplets or even quadruplets. Some strains like the Kefa is valued for milk, meat and skins.

ii) The Somali group

The Somali goats include most goat varieties in Somalia and in the Somali region of Ethiopia (Ogaden). A similar type is found in the northeast Kenya where it is called Galla. This group is intermediate in size between the desert goats of North Africa and

the small goats of East Africa. The goat has relatively small head with a slightly dished profile. The male has short and slender horns, which may be spiral or crescentic in shape. The female's horns are small but many are polled. The male has a short beard. Ears are short and either erect or directed forward. Hair is short and the colour is brilliant white, occasionally with a reddish tinge or with coloured patches on the neck and shoulders and often with brown or black spots around the eyes, ears and nose. The skin has black pigmentation. The breed is adapted to dry environment and is used for meat and skin production.

iii) Small East African goats

This is the most widespread and numerous goat type in Kenya, Uganda and Tanzania. It is also the dominant goat type of Rwanda, Burundi, Zambia, Malawi, Mozambique and here and there in Zimbabwe. The breed is a very variable type in body size and general characteristics. The majority have fine shapely heads, characterised by a straight or slightly concave facial profile. Horns grow backwards and upwards or laterally, varying considerably in length from very short (2.5 cm) to moderately long (25 cm) and often scimitar-shaped, especially in females. In males beard and tassels may or may not be present. The ears are short and in most cases erect, but sometimes horizontal. The hair is usually short and soft, though males often have longer hair along the spine and sometimes on the shoulders and flanks. The colours and pattern vary considerably, but

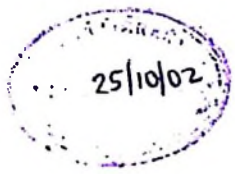
common colours are black, brown, white and grey in various combinations of bicolour or multi-colour. The breed is praised for its hardiness and is used mainly for meat and skin. In some areas this general type has become distinct, for example, in Uganda, varieties include the Mubende of Buganda which has short, fine and usually pure black hair (some animals are pied). The variety is valued for skin. The Kigezi of southwest Uganda has long hair and black colour or grey. It is prized for its long hair, which is used to make clothing.

2.2.2.5 The Lop-eared Southern goats

The most important indigenous breeds which are lop-eared in the south of Africa are Nguni, Damara, Tswana, Ndebele, Pafuri, Boer and Angora.

i) Pafuri

The Pafuri of Mozambique is found in the Limpopo, northwest of Gaza in Suldo province. It is also called Boer goat because it descended from male Boer goats which were introduced into the Pafuri's present breeding area and crossed with Landim in 1928. The goat has a convex profile and well developed horns. Ears are flat and drooping with rounded tips. Coat colour is very variable, it may be black, red, white,



grey, brown or pied. It is a meat and milk goat. The breed has been documented as endangered.

ii) Ndebele

The Ndebele of Gwanda-Tuli, southwest Zimbabwe are characterised by long lop-ears that turn up at the tips, but sometimes the ears are carried horizontally. The horns are short and erect, some animals are polled. Hairs are usually short, but some have longer hairs especially on the hind quarters. The variable coat colours include white, black and brown, and may be spotted or pied. It is a meat goat.

iii) Nguni

The Nguni goats of Swaziland and Zululand are also known simply as Swazi or Zulu goats. The goat is of fairly large size (30 - 40 kg adult live weight). Both sexes are horned and the horns are twisted. The ears are of medium length and flabby. The coat is usually short, but some males have long hairs on the upper part of the extremities. The coat colour is very variable, black, white, grey, whole or mixed. It is principally a meat goat.

iv) Damara

The Damara or Herero of Namibia is characterised by either straight or convex facial profile with medium length horns. The ears are long, wide and drooping. The coat is short and can be of many colours but is usually white, red and white or brown and white and sometimes red or grey. It is mainly a meat goat.

v) Tswana

The Botswana type is a lop-eared goat with erect, medium size horns but some animals may be polled. Ears are usually long and pendulous with turned-up tips, but sometimes may be horizontal. Hairs are usually short but may be long on the hindquarters. Colours include white, black, brown and spotted or pied. It is a meat goat.

vi) Boer

The Boer goat of South Africa was derived principally from Hottentot goats. It became known as Boer goat during the 19th century in order to distinguish it from the Angora goat. Ideally, the modern Boer is totally white with a chestnut-red head and white blaze. The horns are prominently rounded and set well apart growing with a gradual backward curve. The head is strong with roman nose and large eyes. Ears are broad, drooping and

of medium length. The coat is soft, smooth and grossy and the hairs are short to medium in length with no woolly undercoat. It has been bred deliberately as a meat goat. However, it has good milk yield (1.3 – 1.8 kg/day). Also it produces a good quality skin and has a high reproductive rate (7% triplets, 50% twins).

vii) Angora

Angora goats were introduced into South Africa in 1938 from Anatolia in Turkey. They are also found in Lesotho. The goat has a straight or slightly concave facial profile. Horns are short to medium in length with a homonymous twist. Ears are of medium length and pendulous or semi-pendulous. It is almost pure white in colour. The Angora goat is unique in producing true fleece (mohair).

2.3 Conservation of domestic animal diversity

2.3.1 Meaning and importance of Domestic Animal Diversity

Domestic animal diversity has been defined by Hammond (1993) as the spectrum of genetic variation existing among species, breeds and individuals, of all animal species which have been domesticated, and their immediate wild relatives. This genetic variation has been developed during the millions of years of evolution to form and stabilise each

species. Globally, there are about 40 species of domestic livestock (FAO, 1995) which have been domesticated by humans over the last 12,000 years. In the process of domestication, however, separate and genetically unique breeds and strains have developed within each species as a result of human development and occupation of new areas over the planet. According to FAO (1995), there are about 4,000 to 5,000 breeds and strains of domestic animals in the world, of which 300 are breeds of goats (Devendra and Burns, 1983). These breeds and strains are referred to as the global animal genetic resources and the genetic variation both between and within the breeds is described as the diversity within the species of domestic animals. The selection processes, both environmental- and human-directed, have resulted in much of the diversity existing between breeds. Differences among the breeds have been created by reproductive isolation, often imposed by humans through pursuit of different breeding objectives and physical separation for various lengths of time. As a consequence of physical separation, each breed/strain has been adapted to particular ecological conditions to suit the local climate and the requirements of the community. Thus, in the conservation of domesticated species, it is the diversity between breeds, rather than the variation between species, which is of crucial importance (Hammond, 1993; Barker, 1994).

Intraspecific diversity has got both within-population and between-population genetic components. Genetic variation at the population level consists of the differences in the

types of alleles present and their frequencies across all members of a population considered together. Within population genetic variation is caused by change of allele frequencies over time due to selection, random genetic drift and gene flow (immigration from or emigration to other populations). Within population genetic variation is important as it is related to heterozygosity which is known to enhance fitness-related characteristics (Allendorf and Leary, 1986). Between-population variability arises both from random processes (founder effects, demographic bottlenecks, genetic drift and mutations) and from local selection imposed by environment and humans (Hartl and Clark, 1997). Between-population variability is the result of adaptations of populations to their local conditions and it is important since locally adapted populations may have particular genes or gene combinations critical for viability in their local environment. For conservation of animal diversity, it is therefore, important that both within- and between-population variation be maintained. This will enable better utilisation of animals for production of food and agriculture to meet the current needs of humankind while maintaining the potential to meet the range of future possible changes and needs (Hammond, 1993; FAO, 1995). This is because genetic diversity is the basis for a species evolutionary flexibility and responsiveness to environmental changes.

One reason put forward to support conservation of genetic diversity is that the improvement of domestic animals to meet human needs is dependent on genetic variation both within- and between-breeds/strains. Loss of genetic diversity will likely

decrease the ability of animals to respond to environmental change and will also discard genetic information potentially useful to humans (Primack, 1993; Hunter, 1996). The earth's climate varies widely and is continuously changing. It is the diversity of species and breeds that will allow the environment to be utilised efficiently and throughout the years both in temperate areas and in the tropics. The arguments for conservation, particularly in developing countries, are strengthened by the predictions of the scale of human population growth. It has been predicted that, human needs for food will double over the next two generations with the demand for animal products increasing more rapidly than that of plants (Hammond, 1993; Hammond, 1994). With increasing population pressures, the quantity of food and other products must increase. In developed countries consumer emphasis on product quality is increasing. Changing the production level and product quality require different types and varieties. In addition, changes in food and agriculture production will require new combinations of plant and animal species and lines (e.g. breeds and varieties) and changes of management and production strategies. Sustainability in these new production systems will require different genetic types (FAO, 1995).

The second reason put forward to support conservation of domestic animal diversity is that, for a long time to come, the large differences between regions in human needs for food and agriculture and in production capability will persist. According to Hammond (1994), almost three quarters of the world agriculture will remain at the low- to medium-

input levels where animal production environments incorporate combinations of stress such as feed shortage, disease, drought and heat. Breeds in these environments are reservoirs of genetic information responsible for the adaptations necessary for production under adverse environment. Although they have been overlooked, the breeds which are able to produce in harsh conditions, in future may make important contribution to feeding humans most in need (Simm, 1998).

A third reason for conservation of domestic animal diversity is that sustainable crossbreeding schemes require at least two viable purebred populations and often more. Trends towards fewer breeds will reduce the opportunities for improvement of productivity through crossbreeding.

A fourth reason is that the wide variety of livestock breeds and strains we have today is part of our cultural heritage and deserves protection. At the level of breed, unique genetic differences exist between the breeds of each domestic species, for example in goats, the Nubians have high milk yield, the Small East African are hardy animals and the West African dwarf goats are trypanotolerant. Thus, genetic diversity at the level of species can be exploited simply to satisfy human needs for varieties.

2.3.2 Loss of genetic diversity in domestic animals

Concern about reduction in genetic diversity has been expressed primarily in terms of loss of breeds and strains. It is postulated that the current rate of extinction of species, breeds and strains is greater now than at any time in the past (Hammond, 1993). It is estimated that at least 30 to 40% of all animal genetic resources are currently at high risk of extinction (Hammond, 1994). However, adequate records do not exist to enable reliable estimates of either loss rates of the breeds or of domestic animal diversity itself. Thus, the existing data on the number of endangered breeds are likely to be underestimates of the magnitude of the problem. The loss of genetic diversity is occurring both within populations and among populations.

2.3.2.1 Loss of diversity within populations

The factors that diminish genetic diversity within populations are genetic bottlenecks, random genetic drift, inbreeding and human activities. All the four factors are functions of population size (Lacy, 1987; Lande and Barrowclough, 1987).

a) Genetic bottlenecks

A demographic bottleneck occurs when a large population experiences a severe, temporary reduction in size due to environmental or demographic events. In domestic

livestock, natural catastrophes which occur at unpredictable intervals include such things as drought, disease outbreak and war. These catastrophes kill a certain percentage of a population. One result is that the genetic variability of all subsequent generations is contained in the few individuals that survive the bottleneck and reproduce. Hence, some genetic diversity is lost in the process, the magnitude of the loss depends on the size of the bottleneck and the growth rate of the population afterward (Hunter, 1996). Another phenomenon related to bottleneck effects is a founder event. A founder event occurs when a few individuals of a population establish a new population. The genetic constitution of the new population depends upon the genetics of the founders. The genetic diversity of the original larger populations is reduced because the sample of genes in the few founders is not likely to be representative of the original gene pool. Generally, passing through a genetic bottleneck can create two problems (Carson, 1983): a loss of certain alleles, especially rare alleles, if no individuals possessing those alleles survive, and a reduction in the amount of variation in genetically determined characteristics due to the presence of fewer alleles and decline in heterozygosity. Baker and Moeed (1987) have shown that the founder populations of several species of introduced birds have fewer alleles and lower heterozygosity than the source populations. Bottlenecks have been shown to reduce allozyme heterozygosity (Leberg, 1992). Leberg (1992) suggested that extremely low levels of allozyme heterozygosity in broad geographical surveys imply the occurrence of one or more recent severe

bottlenecks. The overall effect of bottlenecks is the decline in fitness of the individuals in the populations.

b) Random genetic drift

Genetic drift is a random change in gene frequencies in small populations, attributable to sampling error. In small populations each generation retains just a proportion of the gene pool of the previous generation. After many generations of random genetic drift, small populations will usually result in loss of variability as a consequence of fixation of some alleles on loci and loss of other alleles.

c) Inbreeding depression

Mating among close relatives results in inbreeding depression. Its probability of occurrence increases in small populations if matings occur at random. Inbreeding allows the rare, harmful recessive alleles to become expressed in the homozygous form, with resulting harmful effects on the offsprings (Selander, 1983; Ralls *et al*, 1988; Charlesworth and Charlesworth, 1987; Lomker and Simon, 1994) such as reduction in fertility, fecundity, offspring size, growth and survival, and physical deformities. Since inbreeding depresses reproductive fitness, it is assumed to increase the risk of extinction. This presumption is supported by correlation between extinction and inbreeding in laboratory and domestic animals (Soulé, 1980).

d) Human activities

Extinction of species, breeds and strains/varieties through human activities represents the greatest threat to genetic diversity. The major threats to genetic diversity that result from human activity are habitat destruction and degradation, pollution, introduction of exotic species, and over-exploitation (Frankham, 1994). These threats are all caused by an ever-increasing use of the world's natural resources due to expanding human population and development of market economy. The growth of towns, factories and mines in developing countries creates a cash market for livestock products. Consequently, the traditional farmers who formerly kept animals for their own needs begin to supply the cash market. Genetic variation is being lost as farmers in developing countries abandon their local breeds in favour of high-yielding breeds for commercial production. Even if human activities do not directly eliminate a breed or strain, loss of genetic variation is taking place as the numbers of individuals in populations are reduced. In the long run the population size of a breed or strain may become so small that the breed/strain is no longer viable and may eventually go extinct.

2.3.2.2. Loss of diversity among populations

Loss of among-population genetic diversity occurs when historically divergent and isolated populations experience an artificially high rate of gene flow from other populations. In domestic animals this is typically happening as traditional farmers mix

and interbreed populations by moving from one place to another when they set up new settlements. Also it occurs through crossbreeding by increasingly promoting universal use of very few superior breeds per species to upgrade local breeds. The result of this is that the uniqueness of formerly distinct breeds/strains is diminished or lost. Moreover, there are concerns that, if there exist co-adapted groups of genes within each subpopulation, merging may result in the break-up of these and the loss of the particular adaptedness of each subpopulation to its environment (Templeton, 1986). Lacy (1987) has shown that the effect of gene exchange between subpopulations is to increase the variance within groups, decrease the variance between groups and decrease the total variance. According to Hammond (1994) about 50% of the total variation for several quantitative traits in cattle, sheep and pig breeds is at the between-breed level, the remainder being common to all breeds. Hence, a move to one breed would eliminate half of the total variation. Loss of genotypic variance is irreversible as is the loss of alleles from a population. This is because it is impossible to sort out the alleles into groups of genotypes resembling the original subpopulations (Nelson and Soulé, 1987).

2.3.3 Approach to conservation of animal genetic diversity

Conservation of genetic resources has become a major issue of public interest in the last decade or so. This is due to the fact that genetic diversity is the basis for a species evolutionary flexibility and responsiveness to environmental changes and loss of genetic

diversity will restrict the opportunities available to mould our domestic animals to meet unpredictable future changes and requirements. Hence, genetic diversity needs to be preserved. Although loss of variation within breeds is continually balanced by the introduction of new variation through new mutations, it is unwise to rely on new mutations as they generally have deleterious effects on reproductive fitness (Frankham, 1994). Moreover, the genetic variation present as differences among breeds cannot be readily regenerated. Each breed or strain is a product of mutation, genetic drift as well as separate adaptation and evolution, often over many centuries and different selection pressures imposed by climate, endemic parasites and diseases, nutrition and selection imposed by humans. Each breed, thus, comprises a unique set of genes which should be conserved.

Conservation of domestic animal diversity has been defined as the sum total of all operations involved in the management of animal genetic resources such that the pool of genetic diversity is maintained over time (Hammond, 1993). It encompasses management of human activities in such a way that these animal genetic resources are best utilised and developed to meet immediate and short term human needs while maintaining the diversity they harbour to meet possible longer term needs for future generations. Most attention in conservation of animal diversity has been directed towards rare breeds. However, in management of animal genetic resources, the fundamental problem is not the distinction between the breeds that are endangered and

those that are not, but between those that are perceived to have little or no current utility and those which do have current utility or seem likely to have an immediate future use (Barker, 1994). According to Notter *et al.*(1994), the critical issues in conservation of domestic animal diversity includes: First, defining utilisation strategies for available breeds to optimise production efficiency and sustainable utilisation throughout the global range of environments and production systems, and livestock improvement strategies that make appropriate use of global domestic animal diversity. Second, development of optimum strategies for assessment and preservation of the genetic diversity which exists in breeds that are little or not currently favoured in commercial production. FAO (1995) has recommended the following strategies for effective management of domestic animal diversity at global level and for each species:

- Identifying and listing all breeds,
- Describing and characterising breeds in order to understand their unique qualities and potential contributions and to understand which breeds have the potential to make the greatest variety of future contributions,
- Monitoring the population statistics for each breed and regularly reporting to the world those populations currently at risk of extinction,
- Storing adequate samples of as many breeds as possible, generally in the form of frozen semen, ova, and embryos, to enable the future regeneration of lost populations of animals.

In general, two methods are used for conservation of genetic resources to support future livestock improvement: *in-situ* conservation and *ex-situ* conservation.

2.3.3 1. *In-situ* conservation

This involves maintaining breeds in their normal farm environment. According to Hammond (1994) *in-situ* conservation is primarily the active breeding of animal populations for food and agriculture, such that diversity is both best utilised in the short term and maintained for the longer term. Operations pertaining to *in-situ* conservation include performance recording and development (breeding) programmes. It also includes ecosystem management for sustainable production of food and agriculture. It has been suggested that the number of animals required to maintain a breed in western countries ranges from 150 to 1500 breeding females, depending on the species and reproductive rate. However, in developing countries the number should not decline below 5000 breeding females (Hodges, 1990). The advantages of this approach are that the animals are still being utilised, the performance characteristics can be properly recorded and evaluated, and the breeds have the opportunity to evolve. The disadvantages are that selection and genetic drift may result in unfavourable genetic changes if the population is small. There is a risk of increasing inbreeding and hence homozygosity, which is associated with reduced fitness. The animals are at risk from diseases and other natural disasters. Also, they are likely to be less productive and so more costly to maintain.

2.3.3.2 *Ex-situ* conservation

In the context of conservation of domestic animal diversity, *ex-situ* conservation means storage. Generally three approaches are used for *ex-situ* conservation.

a) Maintaining breeds in farm parks or other collections

This involves the preservation as live animals of a sample of a breed in a situation removed from its normal production environment or habitat. Many of the pros and cons are similar to the *in-situ* conservation method. However, there is potentially more control over management of the population. It has been suggested that 50 individuals should be the minimum number necessary to maintain genetic variability (Franklin, 1980). With this population only 1% of the genetic variability will be lost per generation. This figure is on the safe side as it has been indicated that animal stocks can be maintained with a loss of 2 – 3% of the variability per generation (Primack, 1993). According to Notter *et al.* (1994) populations can be maintained at an inbreeding rate of 1% per generation corresponding to effective population size (Ne) of 50 which can be achieved with 14 males and 100 females under random mating.

b) Creating a conservation herd (gene pool)

This involves crossing several rare breeds together, then breeding them to maintain genetic variability. It is an effective way of conserving genetic variation for two or three

breeds. Maintenance of genetic diversity is almost better served by pooling five breeds in a conservation herd with Ne of 250 (Notter *et al.*, 1994). However, there is a greater risk of losing useful genes when more populations are combined. The disadvantage of this approach is that, although useful genes may be conserved, the identity of individual breeds is lost.

c) Cryopreservation

This involves frozen storage of rare breeds in the form of living semen, ova, embryos or tissues, which can be used to regenerate animals. Cryopreservation of semen and embryos is a powerful tool for preservation of genetic diversity (Moore, 1985). Where there is a critical threat of a breed becoming extinct, preservation of genetic material of individual animals in the form of germ cells and embryos is necessary to ensure that an adequate genetic pool is retained for future improvement programmes. Technologies for the use of artificial insemination (AI) and embryo transfer (ET) can provide support for this approach. The use of frozen semen in conservation programme is particularly feasible where a tradition of use of AI is already strong. Collection of semen of endangered local breeds should take place as part of the AI programme. According to Smith (1984), semen from 25 unrelated sires would provide an exceptionally good sample of the genetic diversity of a breed. However, Notter *et al.* (1994) recommended cryopreservation of semen from five breeds with 10 sires from each of the breeds and argued that this will provide greater preserved diversity than frozen semen of 25 sires of

each of two breeds. For multiple alleles, the probability of retaining all alleles from the rare breed is higher for the conservation herd supported with frozen semen than for a conservation herd maintained by random mating alone (Lomker and Simon, 1994). However, the ability to rapidly recreate living purebreds is a key issue in germplasm utilisation. If frozen semen is used as the only method of conservation, then several generations of backcrossing are needed to reinstate the breed concerned. In contrast, breeds can be reinstated rapidly from frozen embryos. Embryo preservation is particularly important when the reproductive rate is low or the generation interval is long. Furthermore, maternal cytoplasmic genes cannot be preserved from semen, as recombination of cytoplasmic alleles does not occur across maternal lineages. These alleles are effectively conserved by maintenance of even a modest number of female lineages. However, for all species, the likely cost per haploid genotype maintained or per animal sampled is much less for semen than for embryos (Notter *et al.*, 1994). In general, embryo preservation should be attempted only when ET is already in use for livestock improvement or when a breed has both clear potential value and a high likelihood of extinction.

Frozen storage of semen and embryos have the advantage that after the initial investment, they are relatively inexpensive. Semen and embryos provide much greater flexibility for finding ways of enhancing the reproduction of individuals which are refractory to natural breeding conditions (Moore, 1985). In addition, the breed being

conserved is free from unintended genetic change (Simm, 1998). During storage frozen genetic material is at less risk from disease and natural disasters than live animals, but obviously at risk from technological failures. These risks can be minimised by splitting the material collected from a particular breed or population and storing it in different locations. The disadvantages are that reproductive technologies are not uniformly successful or presently available for all individuals or all species and the expertise is not always available in the places where it is needed most. Also, the breeds conserved in this way are not able to adapt to changes in the production environment or new disease challenges. Semen or embryos frozen now may as well be unable to meet future requirements for improvement.

2.4 Molecular Techniques for the Assessment of Animal Genetic Diversity

2.4.1 Rationale for molecular genetic approach

Phenotypic characterisation has traditionally been used in the description of breeds. The characters vary from external features (horns, ears, coat colour), physical body measurements (heart girth, height, body length), production traits (body weight, milk yield, wool production, carcass traits), reproductive traits (fertility, age at first kidding, kidding interval) to adaptive traits (mortality, resistance to diseases, tolerance to heat and poor management and feeding) (Aboagye, 1992; Aboagye *et al.*,1994). The description of

breeds with respect to coat colour, horns, physical body measurements and productive traits is based upon observation of the phenotype. This is due to the fact that an organism's phenotype is principally a manifestation of its genotype. Phenotypic characters have the advantages of being easily seen and measured and usually much lower costs are incurred during phenotypic characterisation (Minelli, 1993). For these reasons, phenotypic characters have been used extensively for characterisation of breeds. One of the problems associated with phenotypic characterisation is the difficulty in combining different measures in order to provide a useful tool for the description of a given breed. Most phenotypic characters are polygenetically inherited and many of them are influenced by environmental effects, sometimes with a genotype x environment interaction. Furthermore, phenotypic characters are affected by natural selection. Therefore, due to the influence of different environmental conditions and different selection pressure on different kinds of characters, interbreed phenotypic comparison is unlikely to give meaningful results.

Earlier work to study genetic variability of individual animals and populations employed screening of protein variants by gel electrophoresis. Polymorphism in gene products such as enzymes, blood group systems and leucocyte antigens have been used for investigating genetic phenomena at the molecular level. Using the above protein markers, numerous studies have estimated genetic variability, gene flow and phylogenetic relationships among populations (Barbanchao *et al.*, 1984; Nguyen 1990; Pepin and Nguyen 1994). However, these techniques lack the power to resolve the differences between closely related breeds

(Meghen *et al.*, 1994; Dowling *et al.*, 1996) since a great deal of genetic variation remains undetectable by using protein markers. Moreover, the genotype frequencies estimated from protein markers may be influenced by natural selection among alleles (Allendorf *et al.*, 1983; Pemberton *et al.*, 1988; Mopper *et al.*, 1991) making it difficult to interpret interpopulation comparisons. Since the goal of breed characterisation is to determine the within- and between-breed genetic diversity, the near ultimate description of an animal should be a description of the sequence of nucleotides that comprise its genome. Describing differences and similarities in the DNA of two or more populations can provide the measure of relative genetic distances of such populations from each other (Kemp, 1992). The analysis of DNA has several significant advantages over protein markers for the study of molecular population genetics and systematics (Dowling *et al.*, 1996). First, the genotype rather than the phenotype is assayed. Second, the expression of DNA markers is not influenced by development or by environmental factors. Third, one or more sequences appropriate to a problem can be selected on the basis of evolutionary rate or mode of inheritance. Fourth, the methods are, for most part, general to any type of DNA and, fifth, DNA can be prepared from small amounts of tissues and is relatively stable, even in non-cryogenetically stored tissues. The last attribute means that genetic information on rare or endangered species can be obtained without destructive sampling (Taberlet *et al.*, 1997) and it is possible to analyse DNA from extinct populations or species (Taylor *et al.*, 1994). More recently, molecular data from DNA markers have received particular attention in the study of population variability because of their possible

use in determining the chronology of evolutionary events. This is because DNA markers are much less subject to natural selection than phenotypic traits.

2.4.2 Molecular techniques in the study of animal genetic diversity

The past decade has witnessed the development of DNA-based techniques for the study of molecular evolution and population genetics. A variety of different molecular techniques are now being used in various laboratories for the study of inter- and intra-specific genetic variation at the DNA level. The most widely used techniques are restriction fragment length polymorphisms (Southern, 1975) of nuclear DNA (Jeffreys and Morton, 1987) and mitochondrial DNA (Loftus *et al.*, 1994), minisatellites (Jeffreys *et al.*, 1985; Jeffreys and Morton, 1987), microsatellites (Weber and May, 1989), random amplified polymorphic DNA (Williams *et al.*,1990), amplified fragment length polymorphism (Vos *et al.*, 1995) and sequencing of mitochondrial DNA (Cunningham *et al.*, 1994). These techniques differ in the type of data they generate, in the way that they resolve genetic variations and in the taxonomic levels at which they may be most appropriate.

2.4.2.1 Restriction Fragment Length Polymorphism (RFLP)

In RFLP, DNA is digested with restriction enzymes that cut DNA at particular sites within a specific recognition sequence, typically of 4 - 6 base pairs (bp) long (Dowling *et al.*,

1996). Since each enzyme cleaves DNA at a characteristic recognition sequence, the complete digestion of a particular DNA produces an array of fragments. These fragments are separated by gel electrophoresis and blotted onto a filter and then probes are hybridised to the DNA. Variations in fragment lengths between individuals, populations or species is a result of mutations, which either alter restriction sites or lead to insertions/deletions of cleavage sites. Such polymorphism in a specific gene locus can be used to distinguish animal species, populations and individuals. The advantages are that, RFLPs give highly reproducible band patterns and are codominant markers, hence, heterozygotes are distinguishable. The limitations of RFLPs are: It is labour intensive and relatively expensive because a good supply of probes is needed and the blotting and hybridisation steps are time-consuming and difficult to automate. The extent of detectable polymorphism may be limited for intraspecies comparison. And sufficient quantities (e.g. 10μg per digestion) of good quality DNA are required. Because of the last limitation, RFLPs are not applicable where very limited amount of source material or preserved tissues are available.

2.4.2.2 Mitochondrial DNA (mtDNA) analysis

Mitochondrial DNA (mtDNA) sequence differences among individuals or populations can be assayed indirectly through restriction site analysis (Loftus *et al.*, 1994) or directly through DNA sequencing (Cunningham *et al.* 1994; Bradley *et al.*, 1996; Kikkawa *et al.*, 1997). The former is often less expensive and allows rapid screening of large samples. The

latter is more laborious but provides data in their most elemental form. In restriction site analysis, isolated mtDNA is digested with a restriction enzyme and the resulting fragments separated by gel electrophoresis. Genotypes are comprised of the composite restriction fragment pattern for a band of 10 - 20 restriction enzymes. The pattern of restriction site loss or gain among genotypes for a panel of restriction enzymes can be used to estimate genetic distinction and evolutionary history of populations or species (Avise, 1994). In sequencing, mtDNA genes are directly sequenced via the PCR using universal primers that allow amplification of sequences as much as several thousand base pairs in length. Since different genes in the mtDNA genome evolve at different rates, rapidly evolving genes can be analysed to determine the relationship of recently diverged populations, whereas slowly evolving genes can be used to answer for systematic questions involving distantly related species.

The reasons for increasingly widespread use of mtDNA in the studies of animal populations are: the matrilineal inheritance and concomitant hyplotype transmission, general conservation of gene order and composition (especially among closely related forms), rapid rate of sequence divergence (especially in vertebrates), small size as compared to the nuclear genome and relatively ease of isolation and purification. The drawbacks of using mtDNA for population studies are, first, the lack of recombination makes the mitochondrial genome a single heritable unit (effectively a single locus with multiple alleles) potentially producing gene diversity estimates that have larger standard

errors than those determined using nuclear loci. Second, there appears to be more within-individual copy number variation of certain tandemly repeated lengths (more rarely sites) variants than formerly realized. This variation may reflect *de novo* replication errors and not necessarily heritable changes. Third, a small chance of paternal inheritance of mtDNA has been suggested. This possible paternal contribution of mtDNA inheritance can complicate some population level analysis.

2.4.2.3 Random Amplified Polymorphic DNA (RAPD)

The ability of this technique to examine DNA sequences is based on the polymerase chain reaction (PCR) described by Mullis and Faloona (1987). The PCR is a method used to exponentially amplify a specific DNA sequence. The basis of the PCR consists, in repetitive cycles, of three thermal steps, denaturation of the target DNA at a high temperature, hybridisation of specific oligonucleotides on the target DNA at a determined temperature and extension of these oligonucleotides used as primers by DNA polymerase at an elevated temperature usually between 70 and 72°C (Sambrook *et al.*, 1989). RAPD involve the use of a single arbitrary primer (purchasable from commercial companies) in a PCR and result in the amplification of several discrete DNA products. Each product is derived from a region of the genome that contains two short segments in inverted orientation, on opposite strands that are complementary to the primer and sufficiently close together for amplification to work. In RAPD, the amplification products are separated on

agarose gels in the presence of ethidium bromide and the resultant bands are visualized under ultraviolet light (Williams *et al.*, 1990; Welsh and McClelland, 1990). The presence or absence of bands can be scored and the data converted into similarity matrices for calculation of genetic distances (Gwakisa *et al.*, 1994).

The advantages of RAPD are that: there is no requirement for DNA probes or sequence information for the design of primers, the procedure is quick, simple and can be automated, and very small amount of DNA (e.g. 10 ng per reaction) is required. Disadvantages of RAPD are many. First, RAPD methods are very sensitive to the reaction conditions, DNA quality and PCR temperature profiles (Karp *et al.*, 1998). It is absolutely critical to maintain strictly constant PCR conditions in order to achieve reproducible results. Second, the markers are dominant, thus heterozygotes cannot be detected. Third, in the absence of pedigree analysis, the identity of individual bands in the multi-band profiles is not known and there can be uncertainty in assigning markers to specific loci. This makes it difficult to use RAPD in interpopulation or interspecific comparisons (Dowling *et al.*, 1996). Fourth, single bands on the gel can sometimes be comprised of several co-migrating amplification products. This makes it difficult to distinguish many of the polymorphisms apparent from PCR artifacts.

2.4.2.4 Amplified Fragment Length Polymorphism (AFLP)

The AFLP method is a more recent PCR based technique (Vos *et al.*, 1995) which is essentially intermediate between RFLPs and RAPDs. In AFLPs, the first step involves restriction digestion of the genomic DNA and then oligonucleotide adapters are ligated to the ends of the restricted fragments. The second step involves selective amplification of sets of restriction fragments. Either, a pre-selection step is performed using magnetic beads followed by a round of selective PCR or two selective rounds of PCR amplification are applied (Vos *et al.*, 1995). In the last step, the amplified products are separated by polyacrylamide gel electrophoresis and can be visualised using radioactive or fluorescent labelling.

An additional advantage over RAPDs is that AFLPs are reproducible (Vos *et al.*, 1995). Compared with RFLPs, AFLPs are faster, less labour intensive and provide more information (Powell *et al.*, 1996). Because of their large genome coverage (on average they give 100 bands per gel compared with 20 for RAPDs), AFLPs are particularly good for mapping and fingerprinting and genetic distances can be calculated between genotypes. However, AFLPs share many of the limitations, with respect to band homologies and identities, as RAPDs. Furthermore, AFLP bands have been shown to cluster on genetic maps around the centromere (Karp *et al.*, 1998).

2.4.2.5 Variable Number of Tandem Repeats (VNTRs)

Throughout the genome there are many regions comprised of tandemly repeated simple sequences. These repeat sequences vary in number and hence in length and are generally called variable number of tandem repeats (VNTRs), although the terms minisatellite and microsatellite are used depending on the size of the sequence of base pairs which gets repeated. The repeated sequences in minisatellite are 10 to 60 bp long (Haberfeld *et al.*, 1991, Armour and Jeffreys, 1992) whereas in microsatellite the repeated sequences are only 2 to 5 bp long (Tauz, 1989; Weber and May, 1989; Queller *et al.*, 1993). The number of times a given sequence of bases is repeated varies between animals and the rate of mutation at these loci can be very high, up to 10^{-2}/gamete/generation (Jeffreys, *et al.*, 1988) as can the number of alleles per locus. Because of this variation these repeat sequences can be used as markers for studying genetic diversity. Although variation in both systems is the result of changes in copy number, the underlying mechanisms of mutation and the chromosomal locations differ. Microsatellite loci are randomly distributed and subject to replication slippage (Weber and May, 1989; Schlötterer and Tautz, 1992) while minisatellite loci tend to be concentrated near telomeres and vary because of intramolecular or intrallelic recombination and gene conversion (Jeffreys *et al.*, 1994).

Minisatellites are assayed by digestion of DNA with restriction enzymes which do not cleave within the tandem repeats and transfer hybridisation using minisatellite sequences,

an entire hypervariable sequence or synthetic oligonucleotide probes. Two strategies can be applied. In one approach, the mixture of restriction fragments can be separated by size on agarose gels in the presence of an electric field and visualized by hybridisation with a radioactively labelled probe based on the core sequence of the repeat (Jeffreys *et al.*, 1985). The result is complex patterns of bands that are usually unique to an individual. The complex band patterns are widely known as DNA fingerprint. DNA fingerprinting has become an extremely powerful tool for testing parentage (Burke, 1989), relatedness of individuals as well as overall population variability (Wayne *et al.*, 1991). A major advantage of this method is that many probes can be applied across a wide spectrum of plants and animals. An apparent problem with DNA fingerprinting is that it is a multilocus approach, which reveals fragments not all of which segregate independently of each other. Thus, there are correlations among loci due to linkage. Assigning specific fragments to a particular locus is difficult and hence, identifying alleles and determining genotypes is not possible. Moreover, there is potential co-migration of non-homologous fragments. This leads to inflated band sharing coefficients between non-relatives (Lynch, 1988). In the alternative approach, hypervariable loci are assayed one at a time using synthetic oligonucleotides or cloned sequences as single locus probes (Nakumara *et al.*, 1987). The advantage of this method is that alleles can be assigned to specific loci and genotypes can be identified. However, cloning of minisatellite loci is a relatively complex process. This method requires more effort than multilocus fingerprinting and the variation at individual loci tends to be taxonomically restricted (Hanotte *et al.*, 1991).

Microsatellite loci are usually examined one at a time via PCR (several loci can be examined simultaneously by multiplexing). The PCR amplification products can be visualized using radioactive or fluorescence methods. Under radioactive methods, PCR amplified microsatellites can be detected either by direct incorporation of a single labelled deoxynucleoside triphosphate (dNTP) like [α-^{32}P]dCTP during thermal cycling (internal labelling) or a single 5' [γ-^{32}P]dATP end-labelled primer in the PCR mix (end labelling). The products are resolved on acrylamide gels, fixed, dried and autoradiographed. Under the fluorescence methods, fluorescently labelled dNTPs are used for internal labelling in the PCR. The PCR products, fluorescently labelled, are separated on polyacrylamide gel and detected when excited to fluoresce by a laser and analysed by computer software. Although the initial cost is high, this system offers considerable advantage in terms of sizing accuracy and enhances throughput. An advantage of microsatellite over minisatellite is that the total size of the arrays is much smaller and single loci can be amplified using PCR. Thus, there is a great potential for studying allele frequencies at single hypervariable loci in contemporary and historic populations. Moreover, codominant genotypes can be scored and exact allele sizes can be determined. The technique has great potential for speed and accuracy once the appropriate PCR primers are known and a great deal of polymorphism can be determined. Part of the appeal of microsatellites over RFLPs and RAPDs is that their genetic basis of variability is readily apparent: unique primers amplify a genomic region including a well defined repeat structure that is responsible for the observed variation. The disadvantage is the work required to develop primers for each new

species examined and that only a few allelic states are possible, hence, increasing the chance of parallel evolution of a particular sequence repeat. However, with the increasing number of microsatellite maps for economically important species and due to the fact that primers developed for one particular species can be applied across a wide range of related taxa (Moore *et al.*, 1991), cloning of microsatellites for development of primers for new species will be unnecessary in the near future.

2.5 Application of microsatellite markers for the analysis of populations

2.5.1 Microsatellites as genetic markers for population studies

Microsatellites are segments of DNA which consist of simple tandem repeats of very short nucleotide motifs of 1 - 6 bp long, repeated up to ~ 100 times (Tauz, 1989). The dinucleotide repeat cytosine - adenine (CA) (or guanine - thymine (GT) on the other strand) is the most common in mammalian genomes (Meghen *et al.*, 1994). Polymorphism of microsatellites takes the form of variation in the number of repeats at any given locus. Microsatellite polymorphism has advantages over other techniques in that it is a very reliable, highly accurate and repeatable method and has the potential of being standardised all over the world. The desirable characteristics are that: microsatellite loci are found in large numbers and are relatively evenly distributed throughout mammalian genomes. For instance, in haploid human genome there are at least 35,000 CA repeats. That is, they

occur at least every 100,000 bp (Weber, 1990). Tri- and tetra-nucleotides have been shown to occur at a frequency of 1 every 300 - 500 kb on X-chromosome (Edwards *et al.*, 1991). Due to exceptionally high rate of mutation, the majority of microsatellite loci are highly polymorphic in most mammalian species (Jeffreys *et al.*, 1988; Weber, 1990). Alleles at microsatellite loci conform to Hardy-Weinberg principle and segregate in a Mendelian fashion. Most microsatellite loci are selectively neutral. This makes them compatible with the assumptions of most population genetic theories. One potentially very large advantage of microsatellites in population genetics is the fact that primers developed for a particular species can be applicable across a wide range of related taxa. For example, primers developed for cattle can be used for sheep, goats, buffalo, yak, etc.

Microsallites as genetic markers have, recently, become the markers of choice for gene mapping and population studies (Bruford and Wayne, 1993; Wall *et al.*, 1993). They have been extensively used for linkage mapping in diverse organisms from humans to mosquitoes (Weissenbach *et al.*, 1992; Crawford *et al.*, 1995; Vaiman *et al.*, 1996; Barendse, *et al.*, 1997) and their use has enabled the identification of quantitative trait loci in major livestock species (Georges *et al.*, 1993). Trinucleotide repeats have been used for linkage analyses in association with disease susceptibility genes in humans (Richards and Sutherland, 1994). Microsatellites have proven useful in the analysis of paternity and kinship (Queller *et al.*, 1993) and in the probability of sample identity at both the individual (Edwards *et al.*, 1992) and population levels (Paetkau *et al.*, 1995).

Microsatellite variation has been used to study the amount of hybridisation between closely related species (Gottelli *et al.*, 1994; Roy *et al.*, 1994). Spatial distributions of alleles have been used to study local gene flow and population sub-structure (Allen *et al.*, 1995; Gottelli *et al.*, 1994). Effective population sizes and inbreeding have been estimated from microsatellite data (Edwards *et al.*, 1992). Microsatellites have been increasingly used for the study of genetic variation between and within animal populations. They have been successfully applied in the study of genetic variation in a variety of vertebrates (Kashi *et al.*, 1990) and in livestock species such as poultry, sheep and bovines (Haberfeld *et al.*, 1991). Microsatellite DNA variation has been employed to assess genetic distances between strains of poultry (Kuhnlein *et al.*, 1989) and within and among European and African cattle (MacHugh *et al.*, 1994; MacHugh *et al.*, 1997; Okomo, *et al* 1998) and sheep breeds (Buchanan *et al.*, 1994; Arranz *et al.*, 1998).

2.5.2 Estimation of genetic distances from microsatellite data

One of the most important problems in the conservation of genetic diversity in domestic animals is how to choose appropriate breeds or populations for conservation, among many available now, such that maximum genetic diversity could be maintained. One way could be to compute the genetic distances between pairs of breeds which are likely to be useful and to choose breeds that show the widest genetic differences among them (Barker, 1994). Genetic distances are used either for estimating divergent time or for construction of

phylogenies, which in turn can be used to decide which breeds should be kept (Nei and Takezaki, 1994). There are many different measures of genetic distances (Nei, 1987) and the merits and demerits of each measure have been investigated (Takezaki and Nei, 1996).

Many of the earlier measures of genetic distance are geometric distances, which are based explicitly on geometric representations of gene frequencies rather than on any particular population genetic models. These distance measures obey the axioms of Euclidean geometry. The most commonly used distance is Cavalli-Sforza and Edward's (1967) chord distance (D_c). In this distance measure, populations are represented as points on a multidimensional hypersphere with dimension equal to the number of alleles. D_c is the chord distance between two points representing the gene frequencies in two different populations. This distance incorporates an angular transformation of gene frequencies and the principle of triangle inequality is fulfilled. The transformation has the effect of standardising the distance with respect to random drift, so that the rate of increase in genetic distance under drift is nearly independent of the initial gene frequencies. However, little attention is paid to the relationship between genetic distance and evolutionary change of populations. Another widely used measure of genetic distance which is close to D_c is the Nei *et al.*'s (1983) angular distance (D_A). The problem of geometric distances is that the absolute values of these measures do not have any particular biological meaning (Nei *et al.*, 1983), and only the relative values are important for finding the genetic relationship among populations.

The most commonly used genetic distances are based on population genetic models and are supposed to be linearly proportional to time: either to the average time to a common ancestor (coalescence time) of pairs of genes or to the time since two populations diverged from one another. There are two most important genetic models, which have been proposed to explain the mode of mutation as the cause of evolutionary divergence between copies of homologous genes since the time of their common ancestor. The models are infinite alleles model (IAM) and stepwise mutational model (SMM). The IAM model assumes that every new mutation gives rise to an allele that does not already exist in the population while the SMM assumes that mutations increase or decrease allele sizes by single units. IAM seems to apply to classical genetic markers such as protein and blood polymorphisms (Nei, 1987) and SMM apply to microsatellite loci (Goldstein *et al.*, 1995a) since alleles at microsatellites are thought to evolve by a stepwise mutation. Edwards *et al.* (1992) and Valdes *et al.* (1993) have shown that the pattern of mutation for microsatellite loci follow a kind of stepwise mutation model which is close to the IAM of mutation when a relatively short evolutionary time is considered. Thus microsatellite data can be analysed in a manner similar to protein polymorphisms. The most widely used distance measure based on IAM is Nei's (1972) standard genetic distance (D_s). D_s is intended to measure the number of codon substitutions per locus that have occurred after divergence between a pair of populations and is expected to increase linearly with time (Nei and Takezaki, 1994). D_s assumes that the rate of gene substitution per locus is uniform across loci and lineages. This assumption is violated under microsatellite loci. According to Farris (1981) D_s is not

appropriate for making a phylogenetic tree because it is not a metric measure and does not obey the triangle inequalities. Another widely used distance measure under IAM is Rogers' (1972) distance (D_R). This measure has the virtues of simplicity and it satisfies the principle of triangle inequality. However, D_s is better than D_R with respect to the linear relationship with time. Both D_s and D_R have the undesirable property of being influenced by within-taxon heterozygosity. That is, the distance between two taxa that are fixed for alternate alleles exceeds that between two taxa in which one or both are heteroallelic but have no alleles in common.

Although a large number of genetic distances could conceivably be applied to microsatellites, estimators of population parameters based on the IAM are unlikely to apply to microsatellites (Goldstein *et al.*, 1995a). This is because the majority of mutations at microsatellite loci are stepwise in nature, changing allelic sizes up or down by one or a very few number of repeats. Thus, distance measures which apply to microsatellites generally assume the SMM model. Three distances have recently been developed specifically for application to microsatellite evolution assuming the SMM model: Goldstein *et al.*'s (1995a) genetic distance $(\delta\mu)^2$, Goldstein *et al.*'s (1995b) and Slatkin's (1995) average square distance (ASD) and Shriver *et al.*'s (1995) stepwise distance (DSW). The expected values of $(\delta\mu)^2$ and ASD increase linearly with time under the SMM model, but DSW seems to be non-linear with time. Although these distances apply to microsatellites, their efficiencies for obtaining the correct topology of the phylogenetic tree

are generally low (Takezaki and Nei, 1996). This is because ASD, $(\delta\mu)^2$ and DSW have large sampling errors due to their dependence on the variation within populations and on the sensitivity to fluctuation in effective population size (Takezaki and Nei, 1996).

Other widely used measures of genetic distances are various estimators of population subdivision (F_{ST}) such as, G_{ST} (Nei, 1973), θ (Weir and Cockerham, 1984) and R_{ST} (Slatkin, 1995). Different measures of genetic distances are appropriate for different purposes. Nei and Takezaki (1994) and Forbes *et al.* (1995) have reported that classical distances based on the IAM model and the distances based upon multidimensional geometric considerations are more effective in obtaining the correct tree topology than the SMM-based genetic distances. Nei and Takezaki (1994) and Takezaki and Nei (1996) have shown that D_c and D_A are more efficient in obtaining the correct tree topology than other distances whereas D_s and $(\delta\mu)^2$ are more appropriate for estimating evolutionary time than other distances.

2.5.3 Methods for the reconstruction of phylogeny

Phylogenetic trees are graphical representations consisting of nodes (taxonomic units) and branches (pathways connecting nodes) that summarize the evolutionary relationships among organisms (Avise, 1994). Reconstruction of a phylogenetic tree is considered as statistical inference of a true phylogenetic tree which is unknown (Nei, 1996). For the

reconstruction of phylogenetic trees, a large number of methods have been described for tree building. In principle there are two ways of building trees, either by clustering methods or by exhaustive search methods. Clustering methods use a specific algorithm for constructing the best tree. These methods have the advantage of being easy to perform, resulting in very fast computer programs (Saitou and Imanishi, 1989). The biggest limitation is that cluster methods do not provide a ranking criterion for evaluating trees. Exhaustive search methods use optimality criteria for choosing the best tree among the set of all possible trees. The criterion is used to assign to each tree a score or a rank, which is a function of the relationship between the tree and data. Search methods have the advantage that they provide a ranking criterion under which all trees can be evaluated (Page and Holmes, 1998). However, they are computationally demanding, resulting in computer programmes which are much slower to perform and cannot give exact results when there are more than 12 taxa (Nei, 1996).

Based on the above ways of building trees, the methods of phylogenetic inference currently used in molecular phylogenetics can be classified into three major groups: distance methods, parsimony methods and likelihood methods.

2.5.3.1 Distance methods

In distance methods, an evolutionary distance is computed for all pairs of populations and a phylogenetic tree is constructed from the pairwise distances. The most commonly used methods are the 'unweighted pair group method with arithmetic averages' (UPGMA) of Sneath and Sokal (1973) and the Neighbour-Joining (NJ) of Saitou and Nei (1987).

The UPGMA is a clustering method, which operates as follows: A distance matrix is scanned for the smallest distance element and the corresponding operation taxonomic units (OTUs) are joined. This distance element in the matrix is then discarded. The matrix is again scanned for the smallest remaining distance element and the process is repeated until all OTUs are finished. In each cycle of the clustering procedure, OTUs or previously formed clusters are grouped according to the smallest mean distance between the populations involved. Each OTU contributes equally to the calculation of these mean distances. UPGMA produces a dendrogram which is rooted, implicitly at the point where the last clusters join. The major assumption of UPGMA clustering is the equal rate of evolution along all dendrogram branches (Nei, 1987). UPGMA is considered as a heuristic for finding the least squares ultrametric tree for a distance matrix. UPGMA method performs well in recovering tree topologies (Tateno *et al.*, 1982; Nei *et al.*, 1983). This is due to the fact that genetic distance estimates are subject to large

stochastic error and the distance averaging aspects of UPGMA reduces the effects of this error (Nei, 1987).

The NJ is a clustering method but assumes unequal rates of molecular change among the branches (Nei, 1987). In the NJ method the raw data are provided as a distance matrix and the construction of a tree begins by assuming that the populations are all related to one another by a star phylogeny, thus the initial tree is a star tree. The next step is to begin a procedure that groups certain populations together. In order to do this, first, branch lengths for the tree are estimated by least square and the sum of the branch lengths for the entire tree is calculated. Then pairs of populations that give the minimum sum of branch lengths are considered as neighbours and are joined and considered as single entity. At each step of the analysis a transformed distance matrix is constructed in which the separation between each pair of nodes is adjusted on the basis of their average divergence from all other nodes. The tree is constructed by linking the least distant pair of nodes as defined by the modified matrix. When two nodes are linked, their common ancestral node is added to the tree and the terminal nodes with their respective branches are removed from the tree. This pruning process converts the newly added common ancestor into terminal node on a tree of a reduced size. The process of considering all possible pairs is repeated and is completed when two nodes remain, separated by a single branch. According to Nei (1996) NJ is a heuristic method for estimating the minimum evolution tree. NJ produces a reasonably good phylogenetic tree when the extent of

population differences is large and a large number of loci are examined (Rzhetsky and Nei, 1992). The efficiency of the NJ method in obtaining the true tree is due to the fact that in each step of population clustering the principle of minimum evolution is applied and according to Nei (1996) the repeated application of this principle reduces the effects of sampling errors in topology construction.

2.5.3.2 Maximum parsimony methods

Maximum parsimony (MP) methods fall under optimality criteria methods. Under MP the best tree is the one that requires the smallest number of evolutionary changes to explain the observed differences among OTUs. A number of related methods have been described for inferring phylogenies by parsimony approach, the most commonly employed are Fitch and Wagner parsimony (Fitch, 1971; Farris, 1970), Dollo parsimony (Farris, 1981), Camin-Sokal parsimony (Farris, 1971) and generalised parsimony (Swofford and Olsen, 1990). These methods differ in their underlying evolutionary assumptions but they are united by the goal to reconstruct the evolution of OTUs on a tree based on the constraint of invoking the fewest possible evolutionary changes (Page and Holmes, 1998). Among the advantages of parsimony methods are that the approach apparently makes few assumptions about the evolutionary process and it is based on implicit assumption about evolution that evolutionary changes are rare (Swofford *et al.*, 1996). The rarity of evolutionary changes implies that the tree that minimises the change

is likely to be the best estimate of the actual phylogeny (Pages and Holmes, 1998). According to Swofford *et al.*, (1996), if there are no multiple nucleotide substitutions at each site, MP is expected to generate the correct topology as long as enough parsimony-informative sites are examined. The main disadvantage of MP is that under some models of evolution it is not consistent. Indeed, when the true tree has a special type of topology and branch lengths, MP may generate an incorrect topology even if an infinite number of loci are examined (Nei, 1996). Takezaki and Nei (1996) have shown that this can happen even if the rate of nucleotide substitution is constant for all evolutionary lineages. Another disadvantage of MP is that in parsimony analysis it is difficult to treat the phylogenetic inference in a statistical framework because there is no natural way to compute the means and variances of minimum number of substitutions obtained by the parsimony procedure. However, under certain circumstances, MP is quite efficient in obtaining the correct topology and it is the only method that can easily take care of insertions and deletions of nucleotides, which sometimes give important phylogenetic information (Nei, 1996).

2.5.3.3 Maximum likelihood

Maximum likelihood (ML) methods of phylogenetic inference evaluate the likelihood that a given tree will yield the observed data. The data are observed gene frequencies or nucleotide sequences, the unknowns are the branching order and branch lengths of the

tree (Felsenstein, 1981). In ML approach, a model of the evolutionary process that accounts for the conversion of one sequence or allele into another must first be specified. The next step is to calculate the probability that the observed data would have been generated by a particular tree topology under that model. The tree that makes the observed data the most probable evolutionary outcome (i.e. the tree that gives the highest likelihood) is chosen as the ML estimate of the phylogeny. ML is an appealing method of inference as it can incorporate explicit models of evolution and also allows statistical tests of evolutionary hypotheses. The objection to likelihood methods is that the dependence on a model raises the question of which model to use as the pattern of nucleotide substitution varies from site to site and with evolutionary time (Tateno *et al.*, 1982). The biggest problem with ML is a technical difficulty for computing the likelihood itself. This is computationally time-consuming, and recently it has been shown that more than one maximal likelihood value may exist for a given tree (Nei, 1996), making it difficult to guarantee that the likelihood value for that tree is actually maximal.

2.6 Summary of the review

It is believed that the domestic goat, *Capra hircus* is derived from the wild goat, *Capra aegagrus*. Archeological evidence suggests that the domestic goat was first domesticated 10,000 years B.P. in the Fertile Crescent on the borders of present-day Iran and Iraq. Immediately after domestication, differentiation of goats into types must have begun. In the process of domestication, genetically unique breeds and strains developed in different parts of the world mainly through mutation, genetic drift and selection, both natural and artificial. Most breeds have evolved in geographical isolation. However, mixing and interbreeding might have occurred during human movements, trade and wars (livestock raids).

In sub-Saharan Africa there exist populations of goats, in some locations, which seem to be relatively uniform in phenotypic characteristics. Based on phenotypic characters the goats of sub-Saharan Africa can be classified into five major types:- The Northern long lop-eared goats, the Intermediate types or Short-eared twisted horns, the Short-eared dwarf goats, the Small short-eared goats and the Southern long lop-eared goats. However, in between these major types there are many intermediate and less distinguishable subtypes. The history of these types and subtypes (how they were established and if they were later crossed with other types) is not well documented and, thus, the genetic relationships among these types and subtypes is unclear. Unplanned

crossbreeding, ravages of drought, civil conflicts/wars and overexploitation through indiscriminate slaughter and trade pose a risk of loss to these important genetic resources. Genetic characterisation of the indigenous goats in sub-Saharan Africa with the aim of determining the genetic uniqueness of different goat types and subtypes and clarifying the phylogenetic relationships among these types and sub-types is of paramount importance. This information could be used to prioritise breeds/strains for conservation programmes and for establishing improvement programmes to enhance the utilisation of indigenous goats for meat and milk to improve food security and the livelihoods of the people.

In order to help to understand the genetic relationships among the breeds/strains and to quantify the genetic diversity within breeds/strains, this study was undertaken to genetically characterise the goats of sub-Saharan Africa using microsatellite markers. Genetic distances estimated from microsatellite data will give insight into the uniqueness of the different goat breeds/strains in the present study. Genetic distances are used in the reconstruction of phylogenetic relationships of the breeds/strains. The phylogenetic tree, in turn, could be used in prioritising breeds/strains for conservation.

CHAPTER 3

3.0 MATERIALS AND METHODS

The study was divided into two parts, part one investigated the genetic relationships among the major goat types of sub-Saharan Africa and between sub-Saharan African goat breeds and goat breeds from Europe, Middle East and Asia. Part two studied the genetic diversity and structure of goats from eastern African region. The genotyping of samples for part one was done at the Molecular genetics and genomics laboratory, International Livestock Research Institute (ILRI), Nairobi, Kenya, while for part two it was done at the Animal genetics laboratory, Donnan laboratories, School of Biological Sciences, University of Liverpool, Liverpool, UK.

3.1 Breeds/populations and sampling strategy

3.1.1 Breeds representing the major goat types of sub-Saharan Africa

The aim of this study was to assess the genetic diversity and relationships among the goat breeds in sub-Saharan Africa and to compare sub-Saharan African goats with breeds of the other continents. For this purpose, a total of 10 breeds representing the major types of goats in sub-Saharan Africa (i.e. Long lop-eared goats, Small short-eared

goats, Dwarf short-eared goats and Sahelian goats) were sampled. The sampled animals were selected to be as representative of the breed as possible based on the phenotypic characteristics and geographical distribution of the breeds as pre-determined through consultation with local authorities and scientists in the respective countries where the breeds were sampled. For each breed, animals were sampled from distantly located villages, sampling 1 – 4 flocks which are far apart (approximately 1 km) in each village and 1 – 5 animals per flock. In order to avoid sampling related individuals, farmers were asked about the origins and familial relationships of individual animals. Additional five breeds from outside Africa were sampled to serve as reference breeds. Two breeds were sampled from Asia, one from Middle East and two from Europe. For the reference breeds, animals were sampled from research institutions and breeding records were consulted to avoid sampling related individuals. For all the breeds, approximately equal number of females and males were sampled. The breeds, number of animals per breed, country of origin and phenotypic characterisation are shown in Table 1.

Samples collected were either blood (for all African breeds) or hair roots (for the breeds from outside Africa). Blood was collected by jugular vein puncture using 10 ml EDTA vacutainer tubes. To avoid cross-contamination between samples, a new needle was used for each animal. The blood in the tube was then gently inverted about five times to mix it with EDTA. After collecting blood at the field site, all necessary precautions were taken to avoid exposure to extreme temperatures (hot or cold) and direct sunlight. Following

field collection, blood samples were centrifuged within 48 h. Samples were centrifuged at 2000 rpm for 20 minutes. The plasma was discarded and the buffy-coat containing the peripheral blood lymphocytes transferred with a wide-mouth pasteur pipette into a 2 ml cryotube. Then 1 ml of 8 M urea was added into the tube, the content of the tube mixed by gentle inversions five times, the tube properly labelled and stored in upright position at room temperature, ready for DNA extraction.

Hair samples were collected by plucking 50 – 100 hairs from each animal, making sure that the hairs had hair roots. The plucked hairs for each individual animal were packed into a new, clean envelope and the envelope was then sealed and properly labelled. Collecting plucked hairs is an attractive alternative for collecting materials for DNA extraction as it requires less skill, time, money and it is much easier to obtain a sampling permit than collecting blood (Goossens *et al.*, 1998). For this reason there were situations whereby only hair roots were sampled.

Table 1: Breeds used in the analysis of genetic diversity of sub-Saharan African goats and the breeds from outside Africa

Breeds	Number genotyped	Origin	Type	Sub-type
African breeds				
Maasai	48	Tanzania	Short-eared small horns	East African small meat goat
Kigezi	45	Uganda	Short-eared small horns	East African small meat goat
Mubende	47	Uganda	Short-eared small horns	East African small meat goat
Ndebele	40	Zimbabwe	Long lop-eared goats	Southern lops
Pafuri	40	Mozambique	Long lop-eared goats	Southern lops
North West Highland	42	Ethiopia	Short-eared small horns	East African intermediate meat goat
Arsi-Bale	40	Ethiopia	Short-eared small horns	East African intermediate meat goat
West African Dwarf	40	Nigeria	Short-eared small horns	Dwarf African meat goat
Maure	40	Mali	Short-eared twisted horns	Intermediate West African twisted horn
Djallonke	40	Senegal	Short-eared small horns	Dwarf African meat goat
Reference breeds				
Bandipur	40	Nepal		
Mongolian Cashmere	40	Mongolia	Cashmere goats	
Arab	33	United Arab Emirates	Long lop-eared goats	
Toggenburg	20	Switzerland	Short-eared small horns	European dairy goat
Grisons Striped	40	Switzerland	Short-eared small horns	European dairy goat

3.1.2 Populations sampled for the study of eastern African goat populations

The aim of this study was to quantify the genetic diversity, both within- and between-populations of goats from eastern Africa (Ethiopia, Somalia, Kenya, Uganda, Tanzania) as a case study for a sub-regional genetic diversity study. The eastern African region is dominated by small short-eared goats, the major types being the Small East African, the Somali and the Ethiopian short-eared goats (sometimes classified as Small East African goats). It has been very difficult to classify the various populations of eastern African goats into breeds based on phenotypic characters. This is because the goats of this region are very variable in terms of body size, coat colours but homogeneous in most characteristics used for phenotypic characterisation e.g. horns and ears. A total of thirteen (13) populations of eastern African goats were sampled. Reference breeds were sampled from southern part of Africa (2 breeds), West Africa (2 breeds) and one European breed was also included (sampled in Kenya). Table 2 shows the populations sampled, country of origin and their classification based on physical characteristics. The sampling strategy was the same as that explained above in section 3.1.1.

3.2 DNA extraction

DNA was extracted from either peripheral blood lymphocytes (PBL) or hair roots. The extraction procedures applied are described in the following sections.

3.2.1 DNA extraction from peripheral blood lymphocytes (PBL)

Two hundred (200) µl of blood PBL were put into an eppendorf tube and 200 µl of 2 M urea were added, then the mixture vortexed. To this mixture 500 µl of phenol-chloroform-isoamyl-alcohol were added and the contents of the tube were mixed for about 1 h, then centrifuged at 14,000 rpm for 10 minutes and the supernatant transferred to a new eppendorf tube. To the supernatant 30 µl 6 M of sodium chloride were added and the contents mixed by inverting the tube slowly. Then 500 µl of chloroform were added, the contents mixed for 10 minutes, centrifuged at 14,000 rpm for 10 minutes and the supernatant transferred to a new tube. To this supernatant, 800 µl of ice-cold absolute ethanol were added, the contents were mixed and then allowed to settle for 30 minutes. Thereafter, the contents were centrifuged at 14,000 rpm for 10 minutes to precipitate the DNA and the supernatant was decanted. To the precipitate 500 µl of 70% ethanol was added, the contents were shaken to wash the precipitates and then spinned at 14,000 rpm for 10 minutes and the supernatant was decanted. The DNA was air-dried for three days and then dissolved with 100 µl of Tris-EDTA (TE) buffer. After extraction, the concentration of DNA was determined by spectrophotometry and then the DNA concentration was adjusted to 20 ng/µl.

Table 2: Breeds used in the analysis of genetic diversity and structure of eastern African goat populations

Populations	Number genotyped	Origin	Type	Sub-type
Eastern African populations				
Ugogo	48	Tanzania	Short-eared small horns	East African small meat goat
Maasai	50	Tanzania	Short-eared small horns	East African small meat goat
Sukuma	48	Tanzania	Short-eared small horns	East African small meat goat
Newala	50	Tanzania	Short-eared small horns	East African small meat goat
Mbeya	48	Tanzania	Short-eared small horns	East African small meat goat
Ujiji	48	Tanzania	Short-eared small horns	East African small meat goat
Tanzanian Coastal	45	Tanzania	Short-eared small horns	East African small meat goat
Kenyan Small East African	41	Kenya	Short-eared small horns	East African small meat goat
Boran	40	Kenya	Short-eared small horns	East African intermediate meat goat
Galla	20	Kenya	Short-eared small horns	East African intermediate meat goat
North East Highland	46	Ethiopia	Short-eared small horns	East African intermediate meat goat
Afar	45	Ethiopia	Short-eared small horns	East African intermediate meat goat
Landim	36	Mozambique	Short-eared small horns	East African small meat goat
Reference breeds				
Tswana	40	Botswana	Long lop-eared goats	Southern lops
Venda	15	S. Africa	Long lop-eared goats	Southern lops
Red Sokoto	41	Nigeria	Short-eared twisted horns	Intermediate West African twisted horn
West African Dwarf	40	Nigeria	Short-eared small horns	Dwarf African meat goat
Toggenburg	24	Kenya	Short-eared small horns	European dairy goat

3.2.2 DNA extraction from hairs

Fifteen (15) to twenty (20) hairs with hair roots were cut at about 5 mm from the hair root base and put into an eppendorf tube. One hundred (100) µl of lysis buffer (10 mM Tris (pH 8.3), 50 mM KCl and 0.5% Tween 20) were added into the tube. Then 20 µl of 20 µg/µl solution of Proteinase K in 10 mM Tris-HCl (pH 7.5) were added. The contents were vortexed for one minute and centrifuged at 14,000 rpm for one minute. Thereafter, the contents were incubated overnight in a water-bath set at 56°C. Then the contents were incubated for 10 minutes in a water-bath set at 94 °C (to destroy the proteinase K), cooled down to room temperature and then centrifuged at 14,000 rpm for one minute. The resultant sample was used as PCR template.

3.3 Microsatellite markers, PCR conditions and genotyping

Nineteen (19) microsatellite loci were chosen from the goat genome database maintained at INRA (Institut National de Recherché Agronomique) (http://locus.jouy.inra.fr/) and from the sheep genome database maintained at Roslin Research Institute (http://www.ri.bbsrc.ac.uk). The characteristics of the 19 microsatellites are shown in Table 3.

Table 3: Microsatellite markers studied and their characteristics

Locus	Chromosome	Primer sequences 5' – 3'	Annealing temp.°C	Labelling	
				ABI (Colour)	LI-COR (IRD)
BM1818	23	F- AGCTGGGAATATAACCAAAGG R- AGTGCTTTCAAGGTCCATGC	55	TET	800
BMC1222	13q12	F-CCAATTTTGCAGATAAGAAAACA R-CCTGAGTGTTCCTCCTGAGT	55	HEX	700
BMS357	Unknown	F-TCCAAACAAGTCTTCTCTATTTACC R-CCAAATAATTGCTGGTCAGG	58	HEX	800
BMS1494	Unknown	F-TCTGGAGCTTGCAAAAGACC R-AATGGATGACTCCTGGATGG	55	HEX	800
ILSTS005	10	F-GGAAGCAATGAAATCTATAGCC R-TGTTCTGTGAGTTTGTAAGC	55	FAM	-
ILSTS11	14	F-GCTTGCTACATGGAAAGTGC R-CTAAAATGCAGAGCCCTACC	55	HEX	-
ILSTS17	Unknown	F-GTCCCTAAAATCGAAATGCC R-GCATCTCTATAACCTGTTCC	57	HEX	800
ILSTS44	Unknown	F-AGTCACCCAAAAGTAACTGG R-ACATGTTGTATTCCAAGTGC	57	TET	700
ILSTS87	Unknown	F-AGCAGACATGATGACTCAGC R-CTGCCTCTTTTCTTGAGAGC	58	TET	700
INRA5	12	F-CAATCTGCATGAAGTATAAATAT R-CTTCAGGCATACCCTACACC	57	FAM	700
INRA63	18q22	F-ATTTGCACAAGCTAAATCTAACC R-AAACCACAGAAATGCTTGGAAG	55	FAM	800
INRA132	23	F-AACATTTCAGCTGATGGTGGC R-TTCTGTTTTGAGTGGTAAGCTG	57	HEX	700
MAF35	Unknown	F-TCAAGAATTTTGGAGCACAATTCTGG R-AGTTACAAATGCAAGCATCATACCTG	55	HEX	700
MAF65	15	F-AAAGGCCAGAGTATGCAATTAGGAG R-CCACTCCTCCTGAGAATATAACATG	50	TET	-
MAF209	Unknown	F-GATCACAAAAAGTTGGATACAACCGTGG R-TCATGCACTTAAGTATGTAGGATGCTG	57	FAM	800
OarAE129	7	F-AATCCAGTGTGTGAAAGACTAATCCAG R-GTAGATCAAGATATAGAATATTTTTCAACACC	58	FAM	800
OarFCB304	Unknown	F-CCCTAGGAGCTTTCAATAAAGAATCGG R-CGCTGCTGTCAACTGGGTCAGGG	55	HEX	700
SRCRSP3	10q36	F-CGGGGATCTGTTCTATGAAC R-TGATTAGCTGGCTGAATGTCC	55	TET	700
SRCRSP7	6	F-TCTCAGCACCTTAATTGCTCT R-GTCAACACTCCAATGGTGAG	57	FAM	800

3.3.1 Study of genetic diversity within and among sub-Saharan African goats

The 19 microsatellite loci shown in Table 3 were analysed at ILRI. The PCR protocol applied was as follows:- All PCR amplifications were performed in a total volume of 10 µl on either PTC-100™ Programmable Thermal Controller (MJ Research, Inc.) or PTC-200 Peltier Thermal Cycler (MJ Research, Inc.) or GeneAMP$^{\ominus}$PCR System 9700 (*PE* Applied Biosystems). Each PCR reaction contained 1.5 µl of 20 ng template DNA, 0.05 µl of 200 mM each primer, 0.55 µl of 2.5 mM each dNTP (i.e. dATP, dCTP, dGTP, dTTP), 1.10 µl of PCR buffer (2.0 mM $MgCl_2$, 100 mM Tris-HCl (pH 8.3), 50 mM KCl, 0.01% gelatin and 0.25% Tween 20) and 0.5 units *Taq* DNA polymerase (Promega). All amplifications included an initial denaturing step of 4 min at 95 °C, followed by 35 cycles of 45 sec at 94 °C, 1 min at an appropriate annealing temperature (Table 3) and 1 min at 72 °C. Final extension was for 20 min at 72 °C. The PCR products were analysed by electrophoresis on a 4.25% denaturing polyacrylamide gel using a 377 ABI automatic DNA sequencer (PERKIN-ELMER). For this purpose, the forward primer for each locus was labelled with a fluorescent dye, either FAM (blue) or TET (green) or HEX (yellow). Three PCR products (with primers either labelled with different colours or have different sizes) were co-loaded in a single lane of the gel. In preparation for loading, the sample was prepared by mixing 0.5 µl of each PCR products and 2.0 µl of loading buffer (made of GENESCAN™ TAMRA 350, loading dye and formamide in the ratio of 1:1:4). The mixture was heated to 94 °C for 4 min, immediately put on ice and then 1.5 µl was

loaded in each lane of the gel. The GENESCAN™ TAMRA 350 served as an internal size standard in each lane. For further standardisation of fragment sizes, in addition to the GENESCAN™ TAMRA 350, a control sample was included in each gel. Sizes of amplified fragments were analysed using the 672 GENESCAN™ analysis software (version 2.0) and the GENOTYPER™ software (version 2.0).

3.3.2 Study of genetic diversity and structure of eastern African goat populations

In this study 3 microsatellite loci (ILSTS5, ILSTS11 and MAF65) were analysed at ILRI using the same protocol as that explained above (section 3.3.1). The remaining 16 microsatellite loci (BM1818, BMC1222, BMS357, BMS1494, ILSTS17, ILSTS44, ILSTS87, INRA5, INRA63, INRA132, MAF35, MAF209, OarAE129, OarFCB304, SRCRSP3 and SRCRSP7) were analysed at the University of Liverpool. The PCR protocol adopted for the 16 loci was as follows:- All PCR amplifications were performed in a total volume of 5 µl on the GeneAMP©PCR System 9700 (*PE* Applied Biosystems). Each PCR reaction contained 1.5 µl of 20 ng template DNA, 0.05 – 0.1 µl of 200 mM each primer and 2.5 µl Reddy-load PCR master mix (ABgene) (containing 75 mM Tris-HCl (pH 8.8), 20 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 0.01% (v/v) Tween 20, 0.2 mM each dNTP and 1.25 units *Taq* DNA polymerase). All amplifications included an initial denaturation step at 95 °C for 1 min followed by a phase of 5 cycles at 94 °C for 30 sec, annealing temperature (Table 3) for 30 sec and 72 °C for 30 sec. The 5 cycles were

followed by 30 cycles of 30 sec at 92 °C, 45 sec at the annealing temperature (Table 3) and 55 sec at 72 °C. Final extension was for 30 min at 72 °C. The PCR products were analysed on a 6% denaturing polyacrylamide gel using a 4200 LI-COR automatic DNA sequencer (MWG-BIOTECH). For this purpose, the forward primers were labelled with infrared radiations, either 700 or 800 IRD. Four PCR products (with different radiation wave length or different sizes) were co-loaded in a single lane of the gel. To prepare the loading mixture, 1.0 µl from each of the PCR products was pooled, 2.0 µl of sterile double distilled water added, then 1.5 µl of the mixture taken and mixed with an equivalent amount of loading buffer. The mixture was heated for 4 min at 94 °C, immediately put on ice and 1.0 µl was loaded in each lane. External size markers consisting of a mixture of samples with known alleles for each of the microsatellite loci were loaded on the first, middle and last lane of each gel. Scoring of the allele sizes was done using the Gene ImagIR™ software.

3.4 Statistical analyses

Observed allele frequencies were computed using the Excel macros (a programme kindly provided by Stephen Park, (park@tcd.ie)). The allele frequencies were used in the determination of both within- and between-population genetic variation.

3.4.1 Determination of genetic variation within populations and test for Hardy-Weinberg equilibrium

Genetic diversity within each population was determined as the mean number of alleles per locus and average observed and expected heterozygosities. Number of alleles per locus per population was obtained by direct counting. Average observed and expected heterozygosities were calculated using the GENEPOP package version 3.1 (Raymond and Rousset, 1995a). Deviations from Hardy-Weinberg equilibrium (HWE) were tested in two ways: (a) for each locus per population by an exact test using Guo and Thompson's (1992) Markov chain Monte Carlo algorithm implemented in the GENEPOP package version 3.1 (Raymond and Rousset, 1995a); and (b) for all loci and populations using Fisher's method in the GENEPOP package version 3.1 (Raymond and Rousset, 1995a). In addition, the correlation between pairs of genes within individuals within populations (inbreeding coefficient) (*f*) (Weir and Cockerham, 1984), analogous to Wright's F_{IS}, was estimated using the GENEPOP package version 3.1.

3.4.2 Measurement of population sub-division

The drift-based classical measures of population differentiation (F-statistics): the correlation of genes between individuals in the same population (θ) (Weir and Cockerham, 1984) and the coefficient of gene differentiation (G_{ST})(Nei, 1973),

analogous to Wright's F_{ST}, were used to measure the genetic differentiation among the populations. These measures were considered more appropriate for this analysis because genetic drift is said to be the main factor in genetic differentiation among closely related populations or for short-term evolution (Reynolds *et al.*, 1983; Weir, 1990; Takezaki and Nei, 1996). The parameter θ was estimated using the GENEPOP package version 3.1 (Raymond and Rousset, 1995a) while G_{ST} was estimated using the DISPAN program (Ota, 1993). The significance of breed differences was tested using an exact test of population differentiation (Raymond and Rousset, 1995b) implemented in the GENEPOP package version 3.1. The exact test of differentiation (genic differentiation) is designed to assess population differentiation based on whether allelic composition is independent of population assignment. The null hypothesis (H_o) tested is that " the allelic distribution is identical across populations". As an additional measure of interbreed differentiation, breed assignment test was performed using the WHICHRUN programme (Banks and Eichert, 2001). It has been demonstrated that the set of alleles an organism carries for a panel of microsatellites can be used to identify its source population. WHICHRUN uses multilocus genotypic data to allocate individuals to their most likely source population. The likelihood that an individual sample may come from each of the source populations is presumed to be equal to the Hardy-Weinberg frequency of its specific genotype at each locus in each respective source population. Likelihood values for each locus are multiplied to give a series of multi-locus likelihood functions for assignment of individuals to each of the source populations. Alternative hypotheses

that individual samples in question may come from each source population were evaluated using the jacknife method. A jacknife method provides a more empirical means for evaluating baseline data and the chances of correct allocation.

The genetic structure of eastern African goat populations was studied by assessing the similarity between individual animals. The similarity between individual animals was measured as the proportion of alleles shared over all loci and a distance measure between pairs of individuals was calculated as 1- Ps (Bowcock *et al.*, 1994) using the MICROSAT program (Minch, 2000). Where Ps is the proportion of alleles that the individuals share averaged over loci. A distance matrix for all individuals was generated using the MICROSAT program (Minch, 2000). A neighbour-joining tree (NJ) (Saitou and Nei, 1987) using individual animals as operational taxonomic units (OTUs) was constructed with the distance matrix derived from the allele-sharing statistic. The TREEVIEW program (Page, 1996) was used to build the tree.

3.4.3 Determination of genetic variation and relationships among populations

Genetic diversity among populations was measured using three genetic distances: Nei's (1972) standard genetic distance (D_S), Nei *et al.'s* (1983) angular distance (D_A) and Goldstein *et al.'s* (1995b) distance $(\delta\mu)^2$. The D_S distance is based on the infinite alleles model (IAM) and is said to be more appropriate for estimating evolutionary time. The

D_A distance is based on genetic drift model and has been shown to be efficient in recovering the true topology of an evolutionary tree of closely related populations. The $(\delta\mu)^2$ distance is based on stepwise mutational mode (SMM) and has been shown to be appropriate for estimating evolutionary time for microsatellite loci. Both D_S and D_A distance measures were computed using the DISPAN program (Ota, 1993). The NJ methodology (Saitou and Nei, 1987) implemented in the DISPAN program was used to construct the phylogenetic tree of breed relationships. The NJ method is known to be more efficient than many other methods in obtaining the correct tree (Rzhetsky and Nei, 1992). The reliability of the tree topology was examined by performing 1000 bootstrap resampling. The $(\delta\mu)^2$ genetic distance was computed using the MICROSAT program (Minch, 2000) and the NJ tree was built using the TREEVIEW program (Page, 1996).

3.4.4 Multidimensional scaling and principal component analysis

Multidimensional scaling (Manly, 1986) was used to produce graphical representation of the genetic distances (D_A distances) between the populations. The D_A distance, compared to D_S and $(\delta\mu)^2$, revealed better the relationships of the breeds and hence, it was used here. Multidimensional scaling analysis was done using the XLSTAT program (version 4.3) (Fahmy, 2000). Multidimensional scaling is a method that reduces the distance matrix to a set of co-ordinates representing the populations in a small number of dimensions (e.g. two or three). The purpose of multidimensional scaling is to determine

both the dimensionality needed to represent the information in the distance matrix adequately and the position of the points, so that there is maximum correspondence between the observed distances and the inter-point distances. The co-ordinates selected minimise the difference between the observed genetic distances and the fitted distances. The goodness of fit of the distances between co-ordinates (fitted distances) to the observed distances is measured by a stress statistic. The stress statistic varies between 0 and 1, with small values of stress (0 to 0.1) indicating a better fit.

A principal component analysis (PCA)(Manly, 1986) was carried out on the allele frequencies to determine breed relationships based directly on the gene frequencies. PCA involves a linear transformation of the observed allele frequencies into a new set of variables (in geometrical terms, a rotation of the co-ordinate axes). The new variables are uncorrelated with each other. The coefficients defining the linear transformations are chosen so as to minimise the variation of the transformed data measured along each new co-ordinate axis. PCA was performed using the XLSTAT program (version 4.3) (Fahmy, 2000). The first three principal components are the most informative and, in the present study, these were plotted on a three-dimensional scatter diagram to allow visual inspection of the relationships of all populations.

CHAPTER 4

4.0: RESULTS

4.1 Genetic diversity within and among sub-Saharan African goats

A total of 592 individuals, from 10 African breeds and five breeds from outside Africa, were analyzed by determining genotypes at 19 microsatellite loci. All loci were polymorphic and a total of 263 alleles were detected. The numbers of alleles at each locus for each breed are presented in Appendix 1. Table 4 shows the size range, number of alleles and gene diversity for each locus analyzed. The number of alleles across all breeds varied from 8 for locus MAF35 to 23 for locus OarFCB304. Mean gene diversity was lowest for the locus MAF35 (0.321 ± 0.044) and highest for MAF65 (0.796 ± 0.019).

4.1.1 Hardy-Weinberg equilibrium test

A summary matrix of the tests performed to test the deviations from Hardy-Weinberg equilibrium (HWE) is shown in Appendix 2. In total, 95 out of 285 locus-population combination tests conformed to HWE proportions. The genotypes for loci ILSTS17 and OarFCB304 showed significant deviation from HWE expectations in all sub-Saharan

Table 4: Size ranges, number of alleles and gene diversity of the nineteen (19) loci analysed

Locus	Size range (bp)	Number of alleles	Mean gene diversity ± s.e.
BM1818	250 - 300	13	0.759 ± 0.020
BMC1222	273 - 297	13	0.685 ± 0.029
BMS357	102 - 138	18	0.734 ± 0.033
BMS1494	252 - 284	15	0.695 ± 0.019
ILSTS5	164 - 194	12	0.480 ± 0.049
ILSTS11	250 – 316	19	0.705 ± 0.024
ILSTS17	122 - 142	11	0.658 ± 0.039
ILSTS44	135 - 175	17	0.493 ± 0.062
ILSTS87	139 - 179	19	0.599 ± 0.037
INRA5	133 - 167	9	0.569 ± 0.032
INRA63	141 - 179	10	0.624 ± 0.031
INRA132	139 - 159	11	0.512 ± 0.045
MAF35	99 - 119	8	0.321 ± 0.044
MAF65	104 - 138	14	0.796 ± 0.019
MAF209	101 - 131	9	0.443 ± 0.063
OarAE129	130 - 176	22	0.713 ± 0.034
OarFCB304	128 - 180	23	0.651 ± 0.044
SRCRSP3	107 - 127	11	0.585 ± 0.045
SRCRSP7	115 - 131	9	0.582 ± 0.025

African breeds. Table 5 shows the total number of loci showing departure from HWE and the estimates of *f* (analogous of Wright's F_{IS}) in each breed analyzed. The highest number of loci showing departure from HWE were observed in the Ndebele goats; the breed revealed significant departure ($P \leq 0.05$) from HWE at all loci. The Arsi-Bale goats had the minimum number of loci (9 loci) showing deviation from HWE. When the results were pooled across loci and breeds using Fisher's combined probability method, all populations showed significant deviation ($P \leq 0.001$) from HWE expectations. The average values for *f* estimates across all loci were all positive and ranged from 0.126 ± 0.082 to 0.500 ± 0.046 in Grisons Striped and Ndebele, respectively, indicating a deficiency of heterozygosity from Hardy-Weinberg expectations. Genotypic disequilibrium was assessed for all pair-wise combinations of loci in each population. The exact test for non-random association of genotypes across loci in each population revealed 168 pairs (out of 1065 pairs) of loci with significant values ($P \leq 0.05$) in all the 15 breeds.

4.1.2 Genetic variation within breeds

Genetic variability within breeds was measured by the mean number of alleles per locus and average heterozygosity (both observed and expected) (Table 6). Among the sub-Saharan African breeds, the mean number of alleles ranged from 5.26 ± 0.464 to 7.05 ± 0.516 in Djallonke and Mubende, respectively. When all breeds were considered, the

lowest mean number of alleles (4.05 ± 0.442) was observed in Toggenburg while the Bandipur from Nepal had the highest (7.47 ± 0.646). Observed average heterozygosity per breed ranged from 0.342 ± 0.038 in Ndebele to 0.547 ± 0.048 in Mongolian Cashmere. For the expected heterozygosity, the lowest value (0.542 ± 0.036) was found in Pafuri while the highest value was observed in Ndebele (0.672 ± 0.031). In all cases, average observed heterozygosities were lower than that expected under HWE.

4.1.3 Measures of population differentiation

Table 7 shows two measures of population subdivion, the G_{ST} (Nei, 1973) and θ (Weir and Cockerham, 1984) (analogous of Wright's F_{ST}), values across breeds. The values for the two measures of population differentiation were more or less the same. The overall values were 0.157 and 0.146 for G_{ST} and θ, respectively. Locus MAF35 had very high values (0.481 for G_{ST} and 0.504 for θ) and locus INRA63 had the lowest values, 0.095 and 0.076 for G_{ST} and θ, respectively. The exact test for population differentiation for all pairs of breeds across all loci showed that all breeds were significantly ($P \leq 0.001$) different from each other. When breed pairs were tested at each locus (Appendix 4), only 164 tests (out of 1995 tests) showed that the two breeds being compared were not different from each other.

The results of the breed assignment test (Table 8) showed that on average 92.1% of the animals could be assigned correctly to their original population. In each breed the proportion of individuals assigned to their source population ranged from 79% for Mubende goats to 100% for Arsi-Bale, Arab, Toggenburg and Bandipur.

4.1.4 Pair-wise genetic distances among sub-Saharan African breeds

Three measures of genetic distances, D_S, D_A and $(\delta\mu)^2$ were used to estimate the genetic distance for each pair of breeds. The data are shown in Tables 9, 10 and 11 for D_S, D_A and $(\delta\mu)^2$, respectively. For all measures of genetic distances, the lowest distances were observed between Arsi-Bale and North West Highland (Ethiopian breeds) and it was 0.0819, 0.0874 and 0.1590 for D_S, D_A and $(\delta\mu)^2$, respectively. Among the sub-Saharan African goat breeds, the largest distance for D_S was found between Pafuri and West African Dwarf (WAD) (0.4222) while for D_A the largest distance was between Pafuri and Djallonke (0.2731). For the $(\delta\mu)^2$ distances, the largest distance was observed between Pafuri and Maure (7.106). For all the 15 breeds analyzed, the largest distances were found between Mongolian Cashmere and Djallonke (0.6942), Toggenburg and Arab (Emirates) (0.3824) and Pafuri and Maure (7.106) for D_S, D_A and $(\delta\mu)^2$, respectively. Generally, the distance values for $(\delta\mu)^2$ were higher than those of D_S and D_A distances.

Table 5: Number of loci showing deviations from HWE and estimates of inbreeding coefficients (f) in sub-Saharan African goats

Breed	Number of loci significantly deviating from HWE	Average f estimates ± s.e.
African breeds		
Maasai	12	0.266 ± 0.056
Kigezi	18	0.429 ± 0.036
Mubende	17	0.464 ± 0.048
Arsi-Bale	9	0.218 ± 0.057
North West Highland (NWH)	10	0.306 ± 0.057
West African Dwarf (WAD)	17	0.457 ± 0.067
Maure	12	0.303 ± 0.079
Djallonke	15	0.300 ± 0.078
Pafuri	12	0.320 ± 0.070
Ndebele	19	0.500 ± 0.046
Reference breeds		
Arab	14	0.311 ± 0.057
Grisons Striped	10	0.126 ± 0.082
Toggenburg	7	0.162 ± 0.066
Mongolian Cashmere	8	0.144 ± 0.049
Bandipur	10	0.242 ± 0.059

Table 6: Within-breed genetic diversity of sub-Saharan African goats as indicated by mean number of alleles per locus and average heterozygosity

Breed	Mean number of alleles ± s.e.	Average observed heterozygosity ± s.e.	Average expected heterozygosity ± s.e.
African breeds			
Maasai	5.84 ± 0.327	0.420 ± 0.040	0.588 ± 0.037
Kigezi	6.00 ± 0.452	0.376 ± 0.032	0.649 ± 0.032
Mubende	7.05 ± 0.516	0.356 ± 0.034	0.656 ± 0.035
Ndebele	6.84 ± 0.318	0.342 ± 0.038	0.672 ± 0.031
WAD	6.11 ± 0.483	0.366 ± 0.051	0.637 ± 0.036
NWH	5.79 ± 0.570	0.406 ± 0.041	0.597 ± 0.040
Pafuri	5.74 ± 0.523	0.376 ± 0.048	0.542 ± 0.036
Maure	6.05 ± 0.521	0.421 ± 0.053	0.609 ± 0.041
Arsi-Bale	5.95 ± 0.623	0.475 ± 0.038	0.623 ± 0.034
Djallonke	5.26 ± 0.464	0.423 ± 0.056	0.569 ± 0.045
Reference breeds			
Arab	5.53 ± 0.564	0.391 ± 0.044	0.587 ± 0.050
Grisons Striped	5.58 ± 0.608	0.509 ± 0.060	0.559 ± 0.056
Toggenburg	4.05 ± 0.442	0.473 ± 0.068	0.553 ± 0.060
Mongolian Cashmere	6.53 ± 0.707	0.547 ± 0.048	0.634 ± 0.044
Bandipur	7.47 ± 0.646	0.498 ± 0.055	0.645 ± 0.049

Table 7: Measures of population differentiation in sub-Saharan African goats

Locus	G_{ST}	θ
BM1818	0.111	0.096
BMC1222	0.125	0.108
BMS357	0.168	0.159
BMS1494	0.104	0.097
ILSTS5	0.109	0.094
ILSTS11	0.151	0.136
ILSTS17	0.099	0.082
ILSTS44	0.242	0.225
ILSTS87	0.205	0.207
INRA5	0.102	0.092
INRA63	0.095	0.076
INRA132	0.227	0.201
MAF35	0.481	0.504
MAF65	0.125	0.115
MAF209	0.148	0.333
OarAE129	0.164	0.157
OarFCB304	0.167	0.154
SRCRSP3	0.099	0.082
SRCRSP7	0.110	0.097
All loci	0.157	0.146

Table 8: Percentage of animals from each breed assigned to each of the 15 breeds analysed using 19 microsatellite loci

Original breed	Assigned breed														
	Mas	Kig	Mub	Ndeb	WAD	NWH	Paf	Mau	Asib	Djal	Emi	Gris	Tog	Mong	Band
Mas (46)	86.8	2.2	2.2	2.2	2.2	2.2	0	0	2.2	0	0	0	0	0	0
Kig (43)	7.0	88.4	0	2.3	0	2.3	0	0	0	0	0	0	0	0	0
Mub (43)	2.3	4.7	79.0	0	0	0	4.7	0	4.7	2.3	0	0	2.3	0	0
Ndeb (40)	2.5	0	0	90.0	2.5	2.5	2.5	0	0	0	0	0	0	0	0
WAD (40)	0	0	0	0	90	0	0	2.5	2.5	2.5	0	0	2.5	0	0
NWH (42)	2.4	0	0	0	0	90.4	0	2.4	4.8	0	0	0	0	0	0
Paf (40)	0	0	0	0	0	0	95.0	0	0	0	0	5.0	0	0	0
Mau (40)	0	2.5	0	0	5.0	0	0	92.5	0	0	0	0	0	0	0
Asib (40)	0	0	0	0	0	7.5	2.5	0	90.0	0	0	0	0	0	0
Djal (40)	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0
Emi (33)	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0
Gris (40)	0	0	0	0	0	0	0	0	0	0	0	92.5	7.5	0	0
Tog (20)	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0
Mong (38)	0	0	0	0	0	0	0	0	0	0	0	0	2.6	94.8	2.6
Band (40)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100

The numbers in the brackets are the number of animals for each breed.
Mas – Maasai, Kig – Kigezi, Mub – Mubende, Emi – Arab Emirates, Ndeb – Ndebele, WAD – West African Dwarf, NWH – North West Highland, Paf – Pafuri, Mau – Maure, Asib – Asi-Bale, Djal – Djallonke, Gris – Grisons Striped, Tog – Toggenburg, Mong – Mongolian Cashmere, Band – Bandipur.

4.1.5 Phylogenetic relationships among breeds

The neighbour-joining (NJ) tree shown in Figure 2 was constructed using the D_S genetic distances. The phylogeny of all breeds revealed five clusters. The first cluster comprised of West African breeds (WAD, Maure and Djallonke) and the second cluster was made up of eastern African breeds (Maasai, Kigezi, Mubende, Arsi-Bale and North West Highland). The Arab (Emirates) from the United Arab Emirates formed the third cluster. The fourth cluster consisted of the breeds of Southern Africa (Pafuri and Ndebele) and the fifth cluster was made up of the European breeds (Grisons Striped and Toggenburg) and the Asian breeds (Mongolian Cashmere and Bandipur). Within the fifth cluster, two subclusters could be identified, one for the European breeds and the other for the Asian breeds. The bootstrap values (number of times a node was observed in 1000 replicates of resampled loci) ranged from 15 to 95%. The NJ tree generated using the D_A distances (Figure 3) indicated six clusters. The first cluster consisted of East African breeds (Maasai, Kigezi, Mubende) and the second cluster consisted of West African breeds. Unlike the phylogeny based on D_S in which the Ethiopian breeds were on the same cluster with the other Eastern African goats, the D_A based phylogenetic tree separated these breeds from the East African breeds (Maasai, Kigezi, Mubende) and they formed the third cluster. The fourth cluster was made up of breeds from Southern Africa and the fifth cluster consisted of one breed, the Arab (Emirates). Finally, the sixth cluster consisted of the European – Asian breeds. As was the case with the phylogeny generated

using the D_S genetic distances, the two sub-clusters for European and Asian breeds were revealed in this phylogeny. The bootstrap values ranged from 38 to 97%. The phylogeny constructed from $(\delta\mu)^2$ distances indicated three major clusters. The first cluster was made up of one Ethiopian goat population (Arsi-Bale) and the second cluster consisted of Pafuri, Ndebele, North West Highland, Grisons Striped and Kigezi. The third cluster had three subclusters, the first comprised of East African goats (Mubende and Maasai), the second subcluster consisted of Asian and Middle East breeds (Bandipur, Mongolian Cashmere and Arab (Emirates)), and the third subcluster consisted of West African breeds (Maure, WAD and Djallonke) grouped together with a European breed (Toggenburg).

Table 9: Nei's standard genetic (D_S) distances matrix for 10 breeds of sub-Saharan Africa and 5 breeds from outside Africa

	Mas	Kig	Mub	Emi	Ndeb	WAD	NWH	Paf	Mau	Asib	Djal	Gris	Tog	Mong
Kig	0.1070													
Mub	0.1181	0.1180												
Emi	0.3543	0.3875	0.3223											
Ndeb	0.2329	0.2754	0.1943	0.3788										
WAD	0.2964	0.3447	0.2978	0.5549	0.4008									
NWH	0.0907	0.1330	0.1305	0.2462	0.2337	0.2860								
Paf	0.2115	0.1940	0.2073	0.3446	0.1222	0.4222	0.1333							
Mau	0.2610	0.2548	0.3101	0.4998	0.3418	0.1558	0.2233	0.3173						
Asib	0.1574	0.1854	0.0945	0.3179	0.1731	0.3177	0.0819	0.1836	0.2851					
Djal	0.2491	0.2836	0.2425	0.4949	0.3714	0.2416	0.2475	0.3990	0.2299	0.2750				
Gris	0.4234	0.3837	0.4474	0.5230	0.2849	0.4866	0.3055	0.2468	0.2500	0.3650	0.5604			
Tog	0.5136	0.5058	0.5866	0.5654	0.3309	0.5134	0.4263	0.3652	0.3338	0.4592	0.5985	0.1491		
Mong	0.5517	0.5203	0.4530	0.5318	0.3443	0.5311	0.4300	0.3756	0.3501	0.3895	0.6942	0.1895	0.2918	
Band	0.6040	0.5752	0.5353	0.4783	0.3147	0.4963	0.4478	0.3958	0.3116	0.4067	0.6347	0.2644	0.3239	0.1423

Mas – Maasai, Kig – Kigezi, Mub – Mubende, Emi – Arab (Emirates), Ndeb – Ndebele, WAD – West African Dwarf, NWH – North West Highland, Paf – Pafuri, Mau – Maure, Asib – Asi-Bale, Djal – Djallonke, Gris – Grisons Striped, Tog – Toggenburg, Mong – Mongolian Cashmere, Band – Bandipur.

Table 10: Nei's D_A genetic distances matrix for 10 breeds of sub-Saharan Africa and 5 breeds from outside Africa

	Mas	Kig	Mub	Emi	Ndeb	WAD	NWH	Paf	Mau	Asib	Djal	Gris	Tog	Mong
Kig	0.1347													
Mub	0.1366	0.1490												
Emi	0.2841	0.3004	0.2666											
Ndeb	0.1812	0.1931	0.1764	0.2792										
Wad	0.1951	0.2389	0.2167	0.3327	0.2430									
NWH	0.1299	0.1680	0.1504	0.2309	0.1830	0.2183								
Paf	0.1698	0.1771	0.1695	0.2569	0.1252	0.2563	0.1453							
Mau	0.1689	0.1993	0.1944	0.3034	0.2137	0.1285	0.1635	0.2076						
Asib	0.1362	0.1727	0.1374	0.2541	0.1568	0.2165	0.0874	0.1627	0.1708					
Djal	0.1985	0.2358	0.2024	0.3493	0.2612	0.1672	0.2038	0.2731	0.1599	0.2167				
Gris	0.2924	0.2897	0.2953	0.3566	0.2381	0.2942	0.2294	0.1974	0.1944	0.2485	0.3149			
Tog	0.3619	0.3627	0.3551	0.3824	0.2841	0.3386	0.3006	0.2664	0.2532	0.2963	0.3512	0.1571		
Mong	0.3384	0.3433	0.3101	0.3099	0.2511	0.3280	0.2929	0.2605	0.2562	0.2715	0.3819	0.2044	0.2715	
Band	0.3141	0.3528	0.3193	0.2950	0.2174	0.2964	0.2710	0.2444	0.2529	0.2635	0.3512	0.2395	0.2913	0.1666

Table 11: $(\delta\mu)^2$ genetic distances matrix for 10 breeds of sub-Saharan Africa and 5 breeds from outside Africa

	Mas	Kig	Mub	Emi	Ndeb	WAD	NWH	Paf	Mau	Asib	Djal	Gris	Tog	Mong
Kig	0.878													
Mub	0.632	1.142												
Emi	2.724	3.997	3.038											
Ndeb	1.983	1.408	2.093	2.788										
Wad	1.346	3.412	1.642	3.572	3.890									
NWH	1.262	0.983	1.462	2.505	0.538	3.148								
Paf	2.491	1.468	2.852	4.208	0.727	5.891	0.752							
Mau	2.931	3.914	2.717	4.894	4.464	0.960	4.113	7.106						
Asib	0.638	0.692	1.338	2.492	1.019	2.438	0.159	1.080	3.636					
Djal	0.614	2.209	0.700	2.975	2.401	0.393	1.898	3.681	1.970	1.455				
Gris	1.899	1.171	1.564	3.702	1.045	3.278	0.667	1.423	3.268	1.020	2.112			
Tog	2.481	3.570	2.394	3.023	3.060	2.075	1.912	4.724	2.246	2.154	2.214	1.782		
Mong	2.738	3.366	1.982	2.927	1.438	3.707	1.866	2.240	4.303	2.500	2.330	1.892	3.341	
Band	2.575	3.659	1.792	2.198	1.519	2.909	1.670	3.100	3.825	2.415	1.762	2.146	2.125	0.515

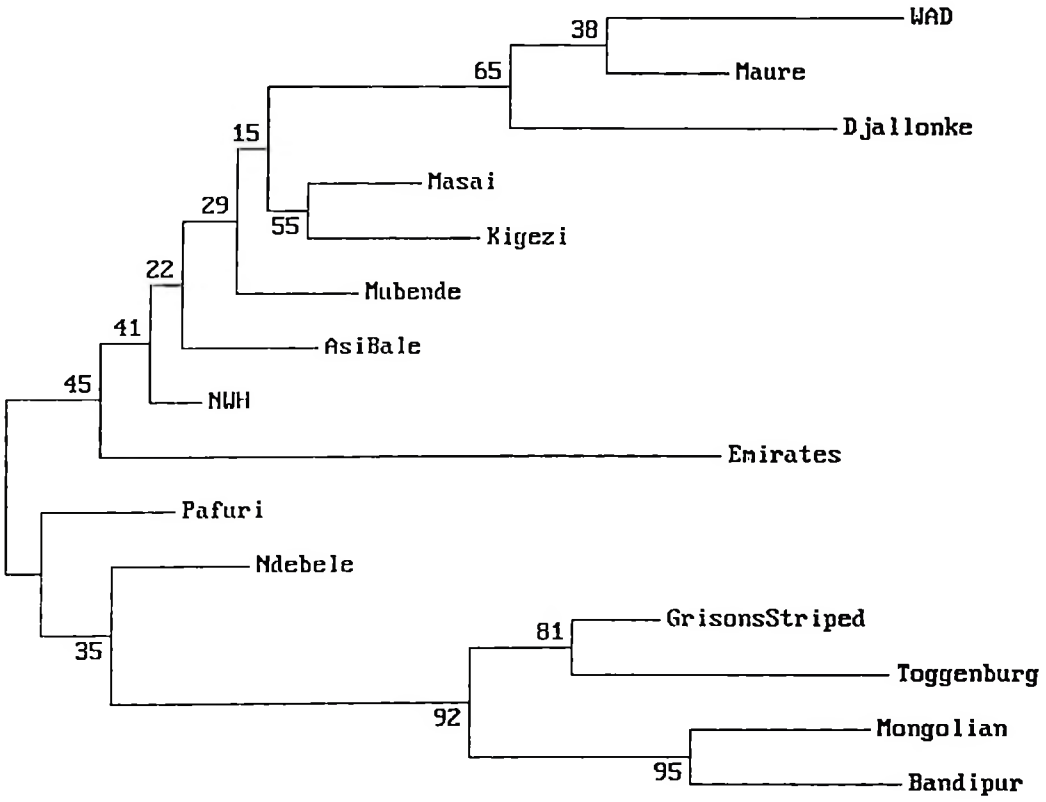


Figure 2: Unrooted neighbour-joining tree generated from D_S distances showing genetic relationships among 10 goat breeds of sub-Saharan Africa and 5 breeds from outside Africa

The numbers at each node are percentage boostrap values obtained after 1000 replications of resampled loci.

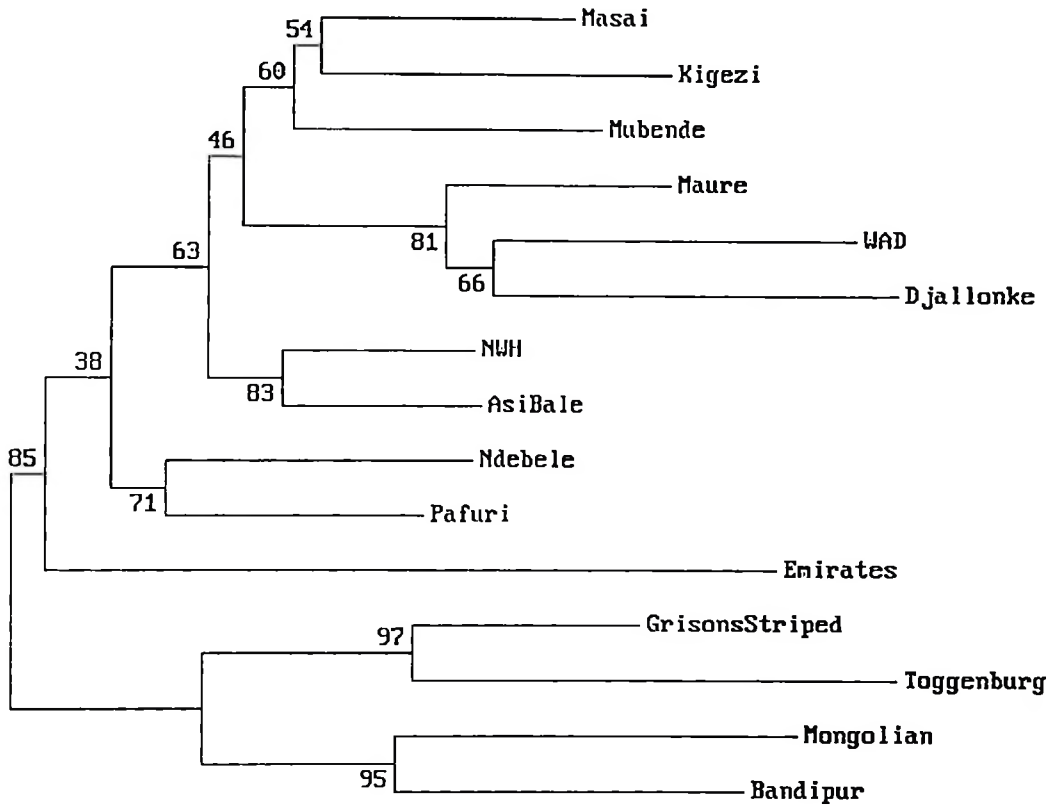


Figure 3: Unrooted neighbour-joining tree showing genetic relationships of 10 breeds from sub-Saharan Africa and 5 breeds from outside Africa based on D_A genetic distances

Numbers at the nodes are bootstrapping values from 1000 replicates.

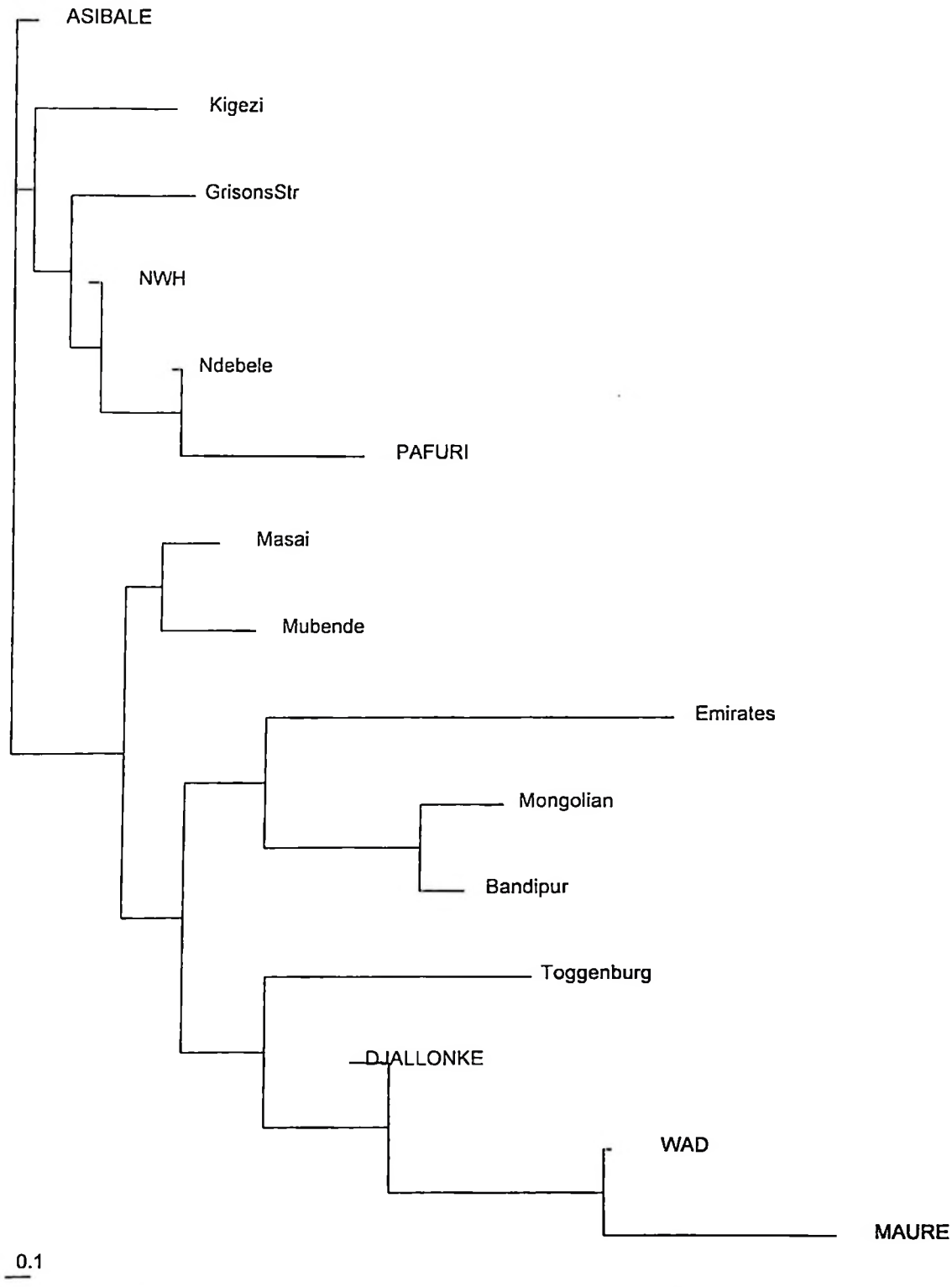


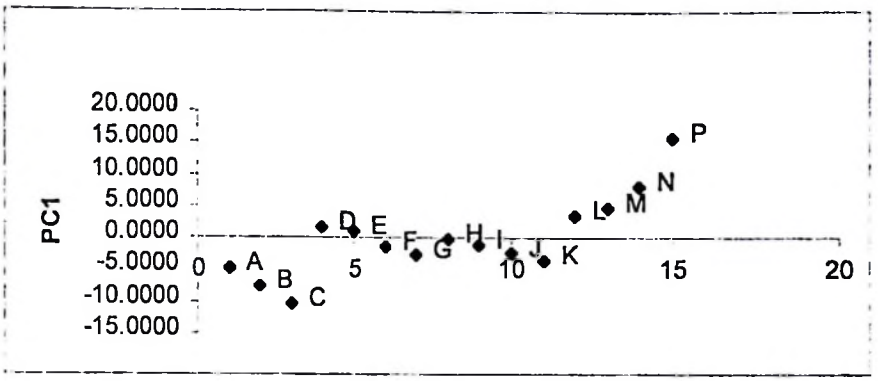
Figure 4: Unrooted Neighbour-joining tree showing genetic relationships of 10 breeds from sub-Saharan Africa and 5 breeds from outside Africa based on $(\delta\mu)^2$ genetic distances
0.1 is the units on the linear scale which relates the branch lengths to $(\delta\mu)^2$ units.

4.1.6 Results of principal component analysis and multidimensional scaling

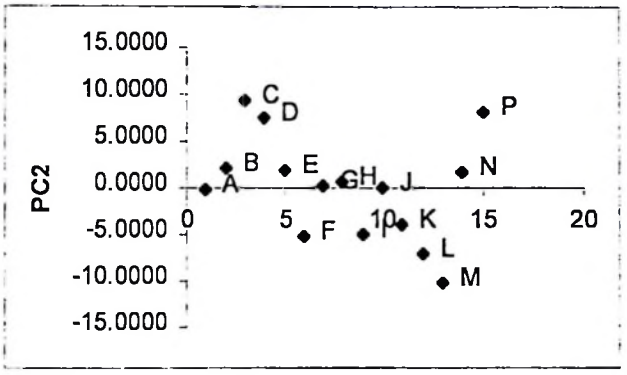
Figure 5 (a) illustrates the plots of the first, second and third principal component scores obtained from the analysis of allele frequencies. The first principal component accounted for 16.1% of the total variability and separated the European-Asian breeds and the East African breeds (Maasai, Kigezi and Mubende) from the rest of the breeds. The second principal component accounted for 12.4% of the variation and clearly separated the European breeds from the Asian breeds. The Arab (Emirates) from the Middle East was also separated from the African, European and Asian breeds with this principal component. The third principal component, which separated the West African breeds from the other African breeds, accounted for 9.1% of the total variation. The 3-dimensional diagram (Figure 5 (b)) obtained by combining the first three principal components indicated three separate groupings of the breeds of sub-Saharan Africa, the first group being for the West African breeds, the second for the eastern African breeds and the third for the breeds of southern Africa. The Mubende goats were well separated from all other African breeds. When all the breeds were considered, the PCA revealed that the European breeds were well separated from the Asian breeds. The African breeds were found to be in between the Asian breeds and the European breeds. The Mongolian Cashmere was found to be closer to the breeds of Southern Africa than to the Bandipur from Nepal. The first ten principal components accounted for 86.1% of the variation and, in total 14 principal components were required to account for all the variation.

Figure 6 illustrates the graphical representation of the D_A genetic distance matrix using multidimensional scaling. The plot has a stress statistic of 0.11. The plot indicated clearly the grouping of the sub-Saharan African goats into the three groups revealed by principal component analysis i.e. a group of eastern African breeds (Maasai, Kigezi, Mubende, North West Highland and Arsi-Bale), a group of Southern African breeds (Ndebele and Pafuri) and a group of West African breeds (WAD, Maure and Djallonke). Similarly, the European breeds were separated from the Asian breeds. The Arab (Emirates) was well separated from all other breeds.

(i)



(ii)



(iii)

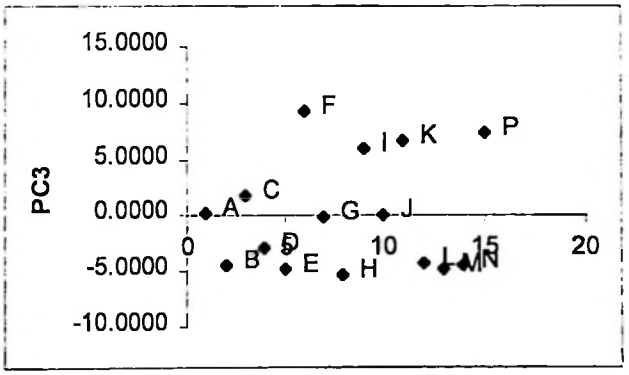


Figure 5 (a): Plots (i), (ii) and (iii) show the separation and associations of the 10 African breeds and 5 reference breeds from outside Africa for the first, second and third principal components

The letters stand for the breeds: A – Maasai, B – Kigezi, C – Mubende, D – Arab (Emirates), E – Ndebele, F – West African Dwarf, G – North West Highland, H – Pafuri, I – Maure, J – Arsi-Bale, K – Djallonke, L – Grisons Striped, M – Toggenburg, N- Mongolian Cashmere and P – Bandipur.

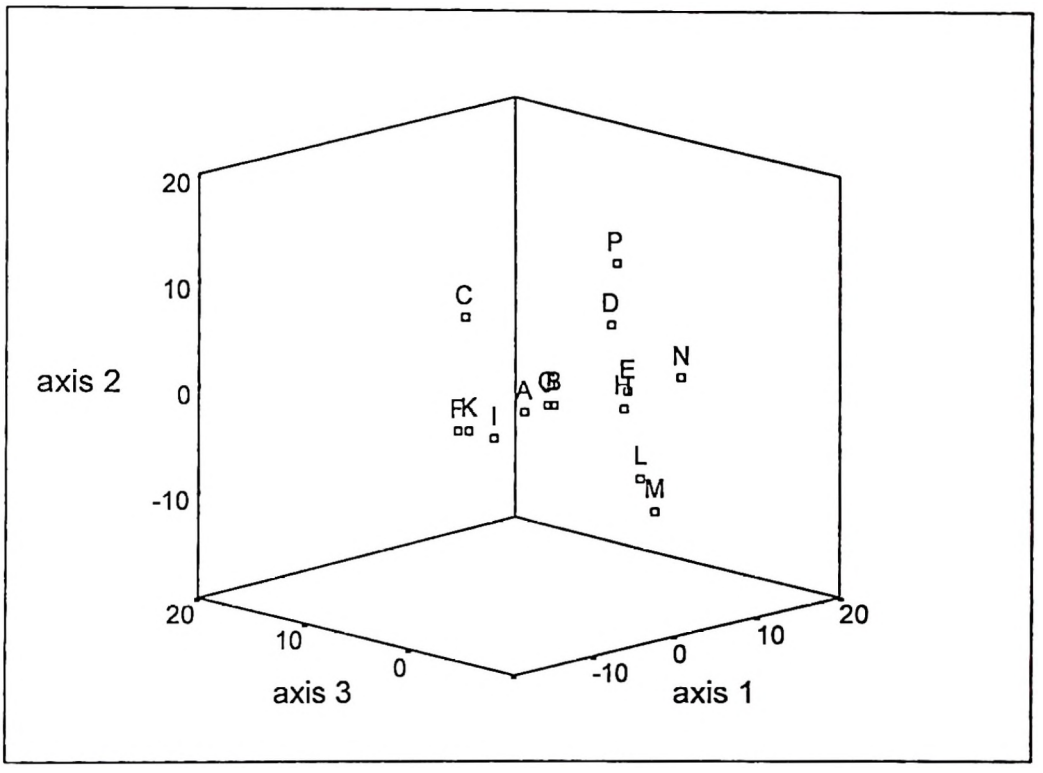


Figure 5(b): Three-dimensional scatterplot for the 10 African breeds and 5 reference breeds from outside Africa generated by combining the first three principal components

The first three components were plotted with the first, second and third principal components represented by axes 1, 2 and 3, respectively. The letters stand for the breeds: A – Maasai, B – Kigezi, C – Mubende, D – Arab (Emirates), E – Ndebele, F – West African Dwarf, G – North West Highland, H – Pafuri, I – Maure, J – Arsi-Bale, K – Djallonke, L – Grisons Striped, M – Toggenburg, N- Mongolian Cashmere and P – Bandipur.

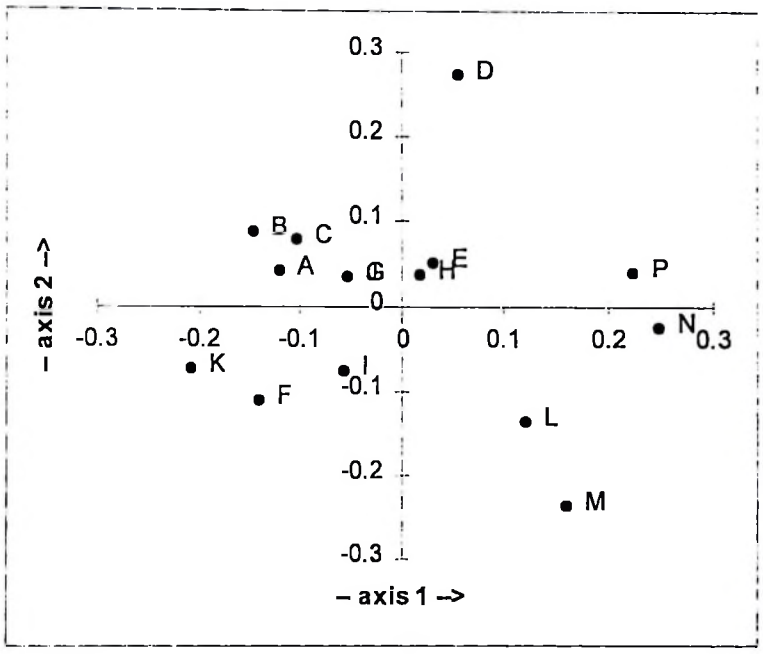


Figure 6: Graphical representation of the D_A genetic distances among 10 African breeds and 5 reference breeds using multidimensional scaling

The letters stand for the breeds: A – Maasai, B – Kigezi, C – Mubende, D – Arab Emirates, E – Ndebele, F – West African Dwarf, G – North West Highland, H – Pafuri, I – Maure, J – Arsi-Bale, K – Djallonke, L – Grisons Striped, M – Toggenburg, N- Mongolian Cashmere and P – Bandipur.

4.2 Genetic diversity and structure of eastern African goats

4.2.1 Hardy-Weinberg equilibrium test

A summary of the results for the HWE tests is presented in Appendix 5. Out of 247 locus-population combinations, 140 tests showed significant ($P \leq 0.05$) deviations from HWE expectation. At locus BMS357, all populations revealed genotype frequencies which significantly ($P \leq 0.05$) deviated from HWE proportions while at locus MAF209, 10 out of 13 populations did not show significant ($P \geq 0.05$) departure from HWE expectation. Table 12 indicates the total number of loci showing deviations from HWE and *f* estimates per population. The Galla had the minimum number of loci (5 loci) with genotype frequencies significantly ($P \leq 0.05$) different from HWE proportions. On the other hand, Ugogo and Mbeya goats showed significant ($P \leq 0.05$) departures from HWE at 14 and 13 loci, respectively, out of the 19 loci analysed. The average *f* estimates were all positive and ranged from 0.066 ± 0.061 in Galla to 0.238 ± 0.040 in Tanzanian Coastal goats.

4.2.2 Within population genetic diversity

The measures of genetic variation within populations are shown in Table 13. In total, 201 alleles were detected in the 13 eastern African goats across the 19 loci analyzed. The

numbers of alleles at each locus in each population are given in Appendix 5. Table 13 shows that, among the eastern African goats, the lowest mean number of alleles per locus was observed in Boran (5.53 ± 0.698) and Galla (5.53 ± 0.599) goats while the highest number of alleles was found in Afar goats (6.53 ± 0.646). Correspondingly, the Afar goats had the highest observed (0.613 ± 0.038) and expected (0.667 ± 0.035) heterozygosities. The lowest heterozygosities, both observed (0.455 ± 0.037) and expected (0.553 ± 0.036), were found in Newala goats.

4.2.3 Population differentiation

Estimates of population subdivision (G_{ST} and θ) are given in Table 14. It is evident that the two measures of population subdivision were approximately equivalent. The G_{ST} values ranged from 0.06 for locus INRA132 to 0.70 for locus MAF35 while θ ranged from 0.04 to 0.72 for the same loci. The overall values were 0.12 and 0.11 for G_{ST} and θ, respectively. The exact test for population differentiation between pairs of populations at each locus revealed that out of 1482 tests, only 363 tests indicated no significant differentiation. The overall test by Fisher's method indicated that all pair-wise comparisons for eastern African goat populations were significantly differentiated from each other in terms of allele frequencies at the loci analyzed. The results for the assignment of individual animals to their source populations are presented in Table 15.

Table 12: Number of loci that showed significant departure of genotype proportions from HWE and inbreeding coefficient (*f*) estimates in eastern African goat populations

Population	Number of loci showing deviations from HWE	*f* estimates ± s.e.
Kenyan Small East African (KSEA)	8	0.071 ± 0.061
Boran	11	0.092 ± 0.065
Galla	5	0.066 ± 0.061
North East Highland (NEH)	11	0.087 ± 0.032
Afar	10	0.076 ± 0.047
Tanzanian Coastal	13	0.238 ± 0.040
Ugogo	14	0.196 ± 0.048
Ujiji	12	0.151 ± 0.047
Maasai	12	0.168 ± 0.043
Sukuma	10	0.146 ± 0.039
Newala	12	0.161 ± 0.048
Mbeya	13	0.169 ± 0.046
Landim	10	0.143 ± 0.051
Tswana	11	0.191 ± 0.051
Venda	7	0.194 ± 0.061
Red Sokoto	10	0.071 ± 0.052

Table 13: Mean number of alleles per locus, average observed and expected heterozygosities in eastern African goats

Population	Mean number of alleles per locus ± s.e.	Average observed heterozygosity ± s.e	Average expected heterozygosity ± s.e
Eastern African			
KSEA	5.63 ± 0.553	0.542 ± 0.035	0.609 ± 0.036
Boran	5.53 ± 0.698	0.536 ± 0.041	0.602 ± 0.035
Galla	5.53 ± 0.599	0.566 ± 0.046	0.615 ± 0.038
NEH	6.00 ± 0.519	0.565 ± 0.041	0.618 ± 0.038
Afar	6.53 ± 0.646	0.613 ± 0.038	0.667 ± 0.035
Tanzanian Coastal	6.26 ± 0.425	0.481 ± 0.042	0.627 ± 0.036
Ugogo	5.95 ± 0.527	0.507 ± 0.042	0.630 ± 0.042
Ujiji	6.16 ± 0.520	0.542 ± 0.040	0.642 ± 0.032
Masai	5.95 ± 0.449	0.485 ± 0.039	0.595 ± 0.040
Sukuma	6.26 ± 0.670	0.524 ± 0.040	0.624 ± 0.044
Newala	5.74 ± 0.545	0.455 ± 0.037	0.553 ± 0.036
Mbeya	6.21 ± 0.487	0.541 ± 0.039	0.646 ± 0.028
Landim	6.21 ± 0.629	0.524 ± 0.045	0.610 ± 0.048
Reference breeds			
Tswana	6.53 ± 0.574	0.518 ± 0.044	0.633 ± 0.034
Venda	4.79 ± 0.416	0.488 ± 0.053	0.603 ± 0.044
Red Sokoto	6.21 ± 0.712	0.594 ± 0.051	0.639 ± 0.042

Among the eastern African goats, the lowest proportion of individuals correctly assigned to their source population was found in Ugogo goats (64.6%) and Galla goats (65%) while the highest proportion was observed in Newala goats (100%). The overall percentage of animals assigned to their source population was 79.1%.

4.2.4 Genetic structure of eastern African goats

The genetic structure of eastern African goats was assessed using the phylogenetic analysis of individual animals. The phylogenetic trees using individual animals as the taxonomic units for the Tanzanian and Ethiopian-Kenyan goat populations are shown in Figures 7 and 8, respectively. For the Tanzanian goats (Figure 7), the pattern of grouping can be divided into 6 clusters. Table 16 shows the distribution of animals into these clusters for each population. The Table shows that 83% (15 out of 18 animals) of the Tanzanian Coastal goats were found in cluster I at the top of the tree. Cluster II consisted of a mixture of Ujiji (89%), Ugogo (50%), Sukuma (44%), Maasai (22%) and Mbeya (17%) goat populations. Cluster III at the middle of the tree was mainly made up of a mixture of Mbeya (61%), Sukuma (44%) and Ugogo (17%) goats. The Newala goats (94%) were grouped in cluster IV. Cluster V consisted of the Tswana goats (reference breed) and cluster VI at the bottom of the tree was mainly made up of the Maasai goats (50%).

Table 14: Estimates of genetic differentiation between eastern African goat populations

Microsatellite Locus	G_{ST}	θ
BM1818	0.094	0.083
BMC1222	0.086	0.069
BMS357	0.103	0.088
BMS1494	0.096	0.089
ILSTS5	0.124	0.106
ILSTS11	0.128	0.122
ILSTS17	0.071	0.055
ILSTS44	0.143	0.106
ILSTS87	0.082	0.073
INRA5	0.072	0.061
INRA63	0.090	0.084
INRA132	0.057	0.038
MAF35	0.701	0.720
MAF65	0.122	0.123
MAF209	0.100	0.079
OarAE129	0.078	0.058
OarFCB304	0.079	0.070
SRCRSP3	0.068	0.061
SRCRSP7	0.130	0.114
All loci	0.120	0.110

Table 15: Percentages of individual animals from each population assigned to the source populations

Source population	Assigned population												
	KS	LD	BR	GA	NE	AF	TC	UG	UJ	MS	SK	NW	MB
KS (41)	78.0	0	9.8	7.3	4.9	0	0	0	0	0	0	0	0
LD (36)	0	88.9	2.8	5.5	2.8	0	0	0	0	0	0	0	0
BR (40)	5.0	0	67.5	17.5	0	5.0	0	0	0	0	0	0	0
GA (20)	10.0	0	25.0	65.0	0	0	0	0	0	0	0	0	0
NE (46)	4.4	0	0	8.7	80.4	6.5	0	0	0	0	0	0	0
AF (45)	4.4	0	4.4	11.1	6.7	71.1	0	0	0	0	0	0	2.2
TC (45)	0	0	0	0	0	0	86.7	2.2	0	6.7	2.2	0	2.2
UG (48)	0	0	0	0	0	0	0	64.6	8.3	14.6	8.3	0	4.2
UJ (48)	0	0	0	0	0	2.1	0	4.2	87.5	4.2	2.1	0	0
MS (50)	0	0	0	0	0	2.0	4.0	4.0	2.0	82.0	6.0	0	0
SK (48)	0	0	0	0	0	0	0	4.2	4.2	6.2	70.8	0	14.6
NW (50)	0	0	0	0	0	0	0	0	0	0	0	100	0
MB (48)	0	0	0	0	0	0	2.1	10.4	2.1	0	6.2	2.1	77.1

The letters used in Table 15 represent the populations as follows; KS – Kenyan Small East African, BR – Boran, GA – Galla, AF –Afar, NE – North East Highland, TC – Tanzanian Coastal, UG – Ugogo, UJ – Ujiji, MS – Maasai, SK – Sukuma, NW – Newala, MB – Mbeya. The numbers in the brackets are the total number of animals included in the analysis for each population.

In the phylogenetic tree for the Ethiopian-Kenyan populations (Figure 8), seven clusters could be identified. Table 17 shows the distributions of animals into these clusters. The Kenyan Small East African were found mainly in cluster III (68%) and VI (12%), the Boran were distributed in cluster I (20%), II (12%), III (16%), V (28%) and VI (12%). The Galla appeared to be spread in clusters III (45%) and V (30%). The Afar were found in cluster I (68%) and II (12%). The North East Highland were distributed in cluster II (20%), III (16%), VI (40%) and VII (20%). The Tswana goats (reference breed) were found in cluster IV (88%).

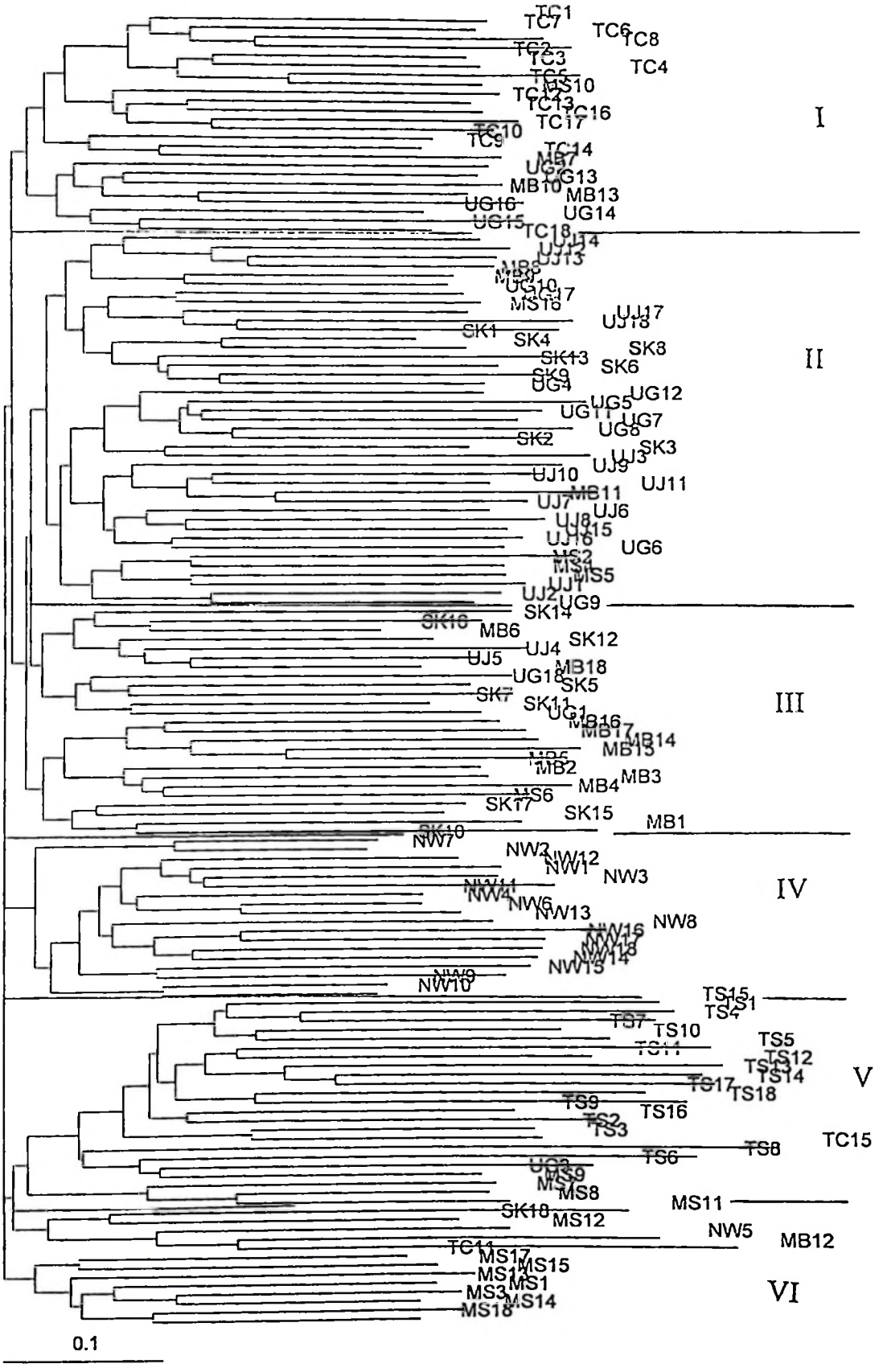


Figure 7: Neighbour-joining dendrogram constructed from allele-sharing distances among 144 individuals from 7 Tanzanian and Tswana goat populations

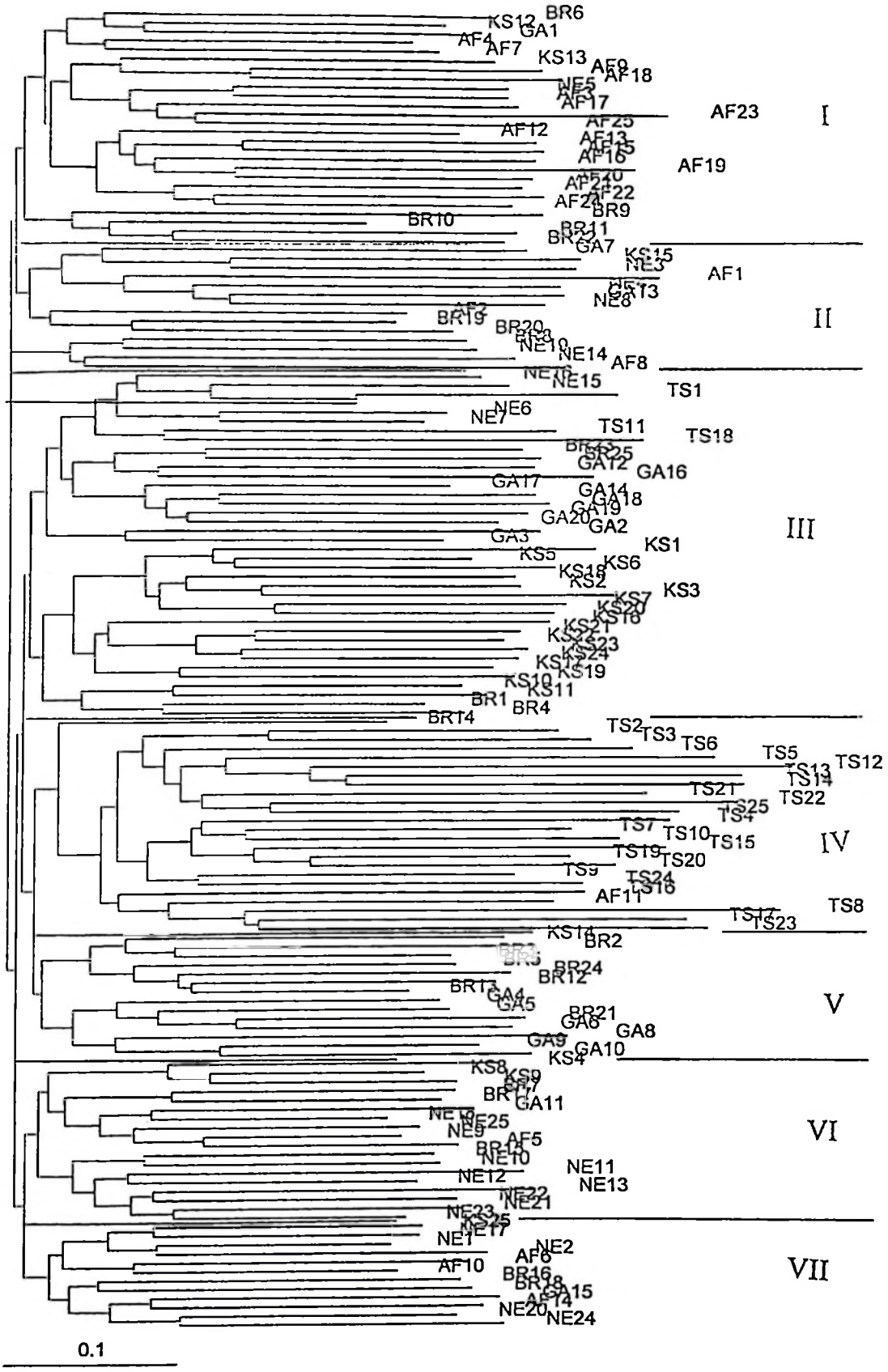


Figure 8: Neighbour-joining dendrogram constructed from allele-sharing distances among 145 individual animals from 2 Ethiopian, 3 Kenyan and Tswana goat populations

Table 16: Allocation of animals into clusters identified from the individual animals dendrogram of the Tanzanian goat populations

Populations	Clusters						Total
	I	II	III	IV	V	VI	
Tanzanian Coastal	15	1	0	0	1	1	18
Ugogo	5	9	3	0	1	0	18
Ujiji	0	16	2	0	0	0	18
Sukuma	0	8	8	1	0	1	18
Mbeya	3	3	11	0	0	1	18
Maasai	1	4	1	0	3	9	18
Newala	0	0	0	17	0	1	18
Tswana	0	0	0	0	18	0	18
Total	24	41	25	18	23	13	144

Table 17: Allocation of animals into clusters identified from the individual animals dendrogram of the Ethiopian - Kenyan goat populations

Populations	Clusters							Total
	I	II	III	IV	V	VI	VII	
KSEA	2	1	17	0	1	3	1	25
Boran	5	3	4	1	7	3	2	25
Galla	1	2	9	0	6	1	1	20
NEH	1	5	4	0	0	10	5	25
Afar	17	3	0	1	0	1	3	25
Tswana	0	0	3	22	0	0	0	25
Total	26	14	37	24	14	18	12	145

4.2.5 Pair-wise genetic distances among eastern African goat populations

Genetic distances between pairs of populations are shown in Tables 18 and 19 for D_S and D_A genetic distances, respectively. Among the eastern African populations, the lowest distance was found between Boran and Galla (0.0381) while the highest distance was observed between Ugogo and Landim (0.3859) followed by Tanzanian Coastal and Landim (0.3839) for the D_S genetic distance. When all populations were considered, the highest distance was observed between WAD and Newala goats (0.5897). For the D_A distance, the pair of populations that had the lowest distances were Sukuma – Ugogo (0.0676), Afar – North East Highland (0.0684) and Boran – Galla (0.0687). The highest distance was observed between Ujiji and Landim (0.2260), followed by Ugogo and Landim (0.2171). For all the populations, the highest D_A distance was found between Newala and Toggenburg (0.3017).

4.2.6 Phylogenetic relationships among eastern African goat populations

The dendrogram showing the genetic relationships based on D_S genetic distances among the eastern African goat populations is shown in Figure 9. This dendrogram indicates three groupings of eastern African goats. The first group, which also contains the reference breeds from West Africa (West African Dwarf and Red Sokoto), consisted of Boran and Galla. The second group consisted of North East Highland, Afar and Kenyan

Small East African. However, the Kenyan Small East African formed their own subgroup separated from the Ethiopian goats (North East Highland and Afar). The third group was made up of the Tanzanian goat populations. This could be divided into two subgroups, one containing Newala, Mbeya and Tanzanian Coastal goats and the other containing Sukuma, Maasai, Ugogo and Ujiji goats.

Figure 10 shows the D_A- based dendrogram of genetic relationships among eastern African goat populations. This dendrogram clustered the goats of eastern Africa into two groups, a group of Kenyan – Ethiopian populations and a group of Tanzanian populations. The first group had two subgroups, the first being for Ethiopian populations (North East Highland and Afar) while the second subgroup consisted of Kenyan populations (Boran, Galla and Kenyan Small East African). In the second group, three subgroups could be identified. The first subgroup was made up of Tanzanian Coastal and Maasai, the second subgroup consisted of Newala and Mbeya, while the last subgroup consisted of Sukuma, Ugogo and Ujiji.

Both dendrograms (Figures 9 and 10), showed that the Landim goats are genetically closer to the goats of southern Africa (Tswana and Venda) than to the eastern African goats. Also the European breed (Toggenburg) is genetically close to the southern African goats but well separated from eastern African goats.

Table 18: Matrix of D_S genetic distances between 13 eastern African goat populations and 5 refernce goat populations

	Tsw	Ven	KSEA	Tog	Lad	Bor	Gal	NEH	Afar	Tco	Ugo	Uji	Mas	Suk	New	Mby	Rsk
Ven	0.0506																
KSEA	0.2120	0.2004															
Tog	0.2219	0.2029	0.2710														
Lad	0.1054	0.0626	0.2025	0.1828													
Bor	0.2047	0.1775	0.0658	0.1721	0.1843												
Gal	0.2052	0.1911	0.0764	0.1597	0.1559	0.0381											
NEH	0.1878	0.1757	0.0555	0.2308	0.1520	0.0866	0.0551										
Afar	0.2020	0.1554	0.0767	0.2068	0.1704	0.0729	0.0756	0.0454									
Tco	0.4648	0.3305	0.2781	0.4986	0.3839	0.3038	0.3141	0.2980	0.2579								
Ugo	0.3756	0.3303	0.2371	0.4224	0.3859	0.2206	0.2321	0.2341	0.2315	0.1433							
Uji	0.3584	0.3235	0.2254	0.3888	0.3799	0.1957	0.2369	0.2384	0.2307	0.1904	0.0555						
Mas	0.3827	0.3769	0.1968	0.4672	0.3685	0.2090	0.2302	0.2232	0.2424	0.1484	0.0596	0.0998					
Suk	0.3392	0.2963	0.2143	0.3603	0.3151	0.1746	0.2053	0.2262	0.2210	0.1456	0.0494	0.0858	0.0838				
New	0.3166	0.2888	0.2926	0.5089	0.2690	0.3629	0.3229	0.2813	0.3467	0.2478	0.1998	0.2691	0.1944	0.1775			
Mby	0.3162	0.2891	0.1733	0.4810	0.3240	0.2210	0.2261	0.2035	0.2028	0.1346	0.0702	0.1362	0.1101	0.0865	0.1328		
Rsk	0.2564	0.2428	0.1327	0.2462	0.2957	0.0934	0.1274	0.1509	0.1158	0.3182	0.2730	0.2723	0.3350	0.2608	0.4622	0.2588	
WAD	0.3085	0.2819	0.2260	0.3316	0.4229	0.1818	0.2776	0.2673	0.1921	0.3978	0.3707	0.3455	0.3993	0.3479	0.5897	0.3566	0.1037

Tsw – Tswana, Ven – Venda, KSEA – Kenyan Small East African, Tog – Toggenburg, Lad – Landim, Bor – Boran, Gal – Galla, NEH – North East Highland
Tco – Tanzanian Coastal, Ugo – Ugogo, Uji – Ujiji, Mas – Maasai, Suk – Sukuma, New – Newala, Mby – Mbeya, Rsk – Red Sokoto, WAD – West African Dwarf.

Table 19: Matrix of D_A genetic distances between 13 eastern African goat populations and 5 reference goat populations

	Tsw	Ven	KSEA	Tog	Lad	Bor	Gal	NEH	Afa	Tco	Ugo	Uji	Mas	Suk	New	Mby	Rsk
Ven	0.1130																
KSEA	0.1486	0.1586															
Tog	0.1946	0.2111	0.1928														
Lad	0.1329	0.1311	0.1511	0.1893													
Bor	0.1775	0.1796	0.0793	0.1675	0.1654												
Gal	0.1696	0.1779	0.0812	0.1695	0.1632	0.0687											
NEH	0.1498	0.1603	0.0949	0.1997	0.1538	0.1033	0.0926										
Afa	0.1587	0.1620	0.0948	0.1869	0.1611	0.0945	0.0846	0.0684									
Tco	0.2210	0.2011	0.1489	0.2647	0.2067	0.1681	0.1700	0.1658	0.1373								
Ugo	0.2042	0.1949	0.1515	0.2541	0.2171	0.1500	0.1657	0.1667	0.1410	0.0981							
Uji	0.1978	0.2019	0.1574	0.2362	0.2260	0.1437	0.1767	0.1593	0.1476	0.1278	0.0700						
Mas	0.2206	0.2310	0.1311	0.2790	0.2131	0.1469	0.1560	0.1561	0.1413	0.0938	0.0753	0.0968					
Suk	0.1922	0.2073	0.1307	0.2273	0.1958	0.1223	0.1498	0.1584	0.1418	0.1203	0.0676	0.0953	0.0937				
New	0.1994	0.1903	0.1661	0.3017	0.1795	0.2033	0.2085	0.2111	0.2067	0.1646	0.1410	0.1806	0.1459	0.1248			
Mby	0.1609	0.1727	0.1090	0.2606	0.1795	0.1344	0.1486	0.1469	0.1278	0.0993	0.0740	0.1065	0.1003	0.0792	0.0998		
Rsk	0.1659	0.1773	0.1226	0.1807	0.2150	0.1215	0.1317	0.1538	0.1309	0.1909	0.1538	0.1571	0.1933	0.1585	0.2497	0.1549	
WAD	0.1939	0.2163	0.1790	0.2223	0.2840	0.1700	0.2054	0.2015	0.1751	0.2145	0.1921	0.1849	0.2162	0.1956	0.2962	0.1983	0.0996

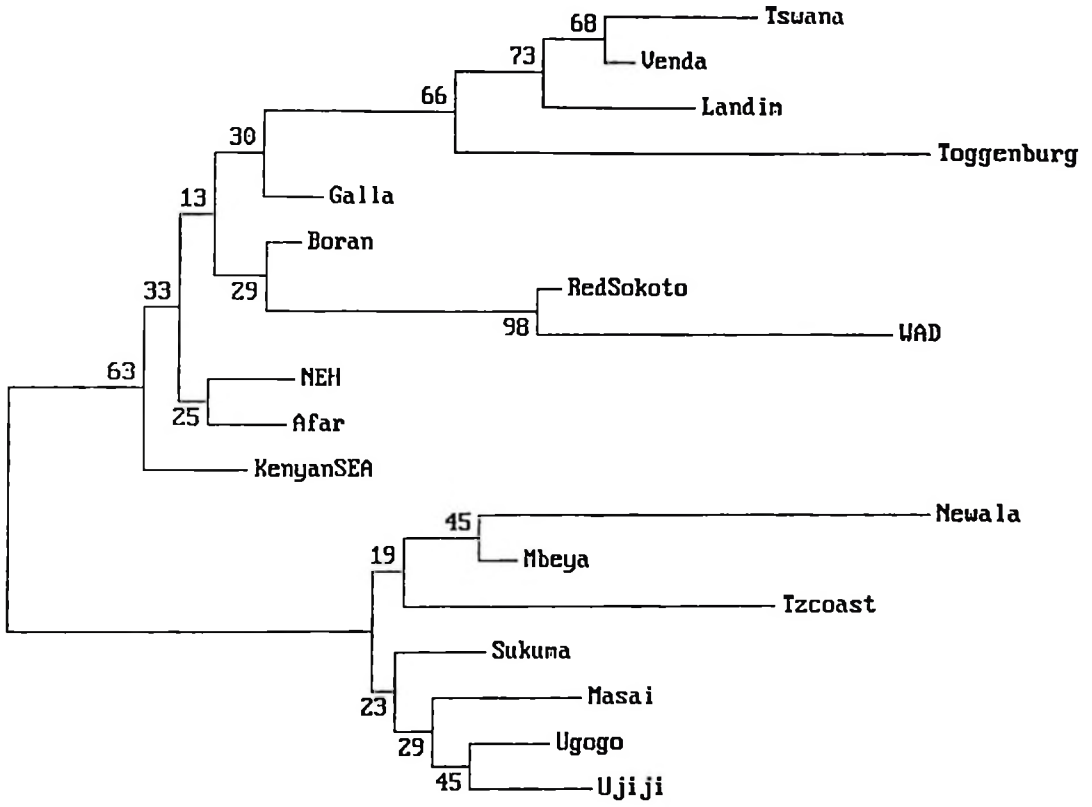


Figure 9: Unrooted neighbour-joining tree constructed from D_S distances showing genetic relationships of 13 eastern African goat populations and five reference goat populations from southern Africa and West Africa and one European breed

The numbers at the nodes are bootstrap values for 1000 bootstrap resamplings.

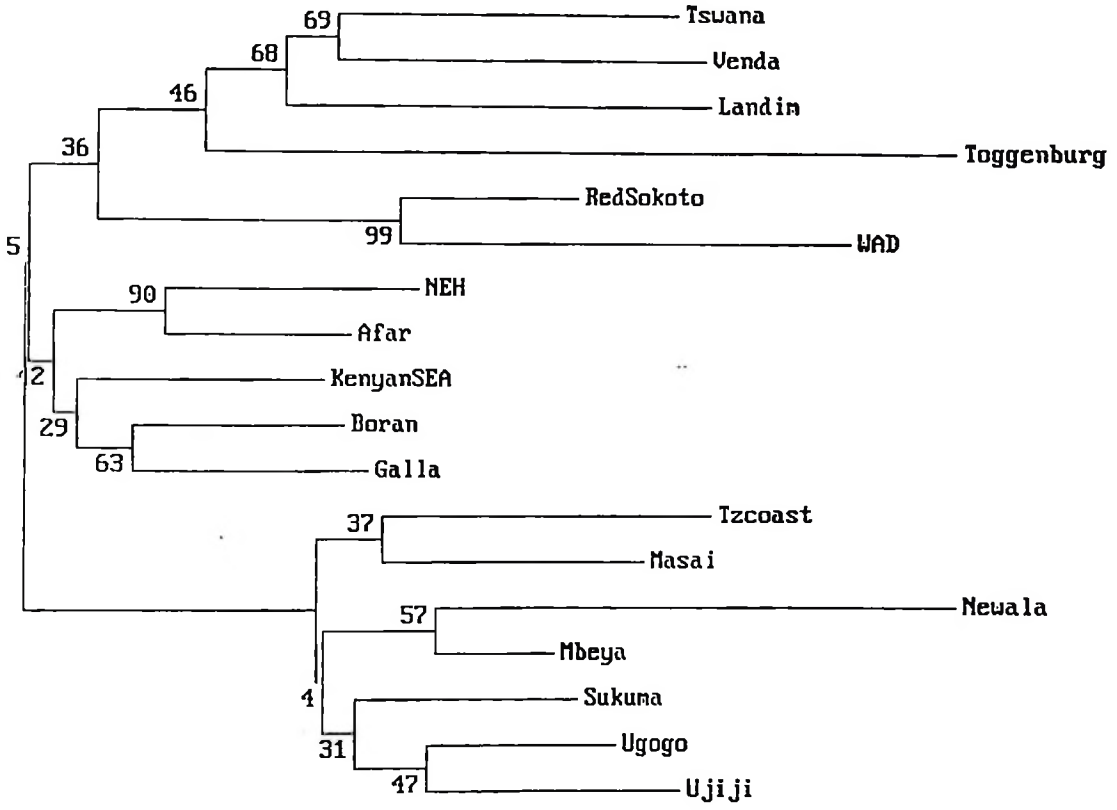


Figure 10: Unrooted neighbour-joining tree constructed from D_A genetic distances showing genetic relationships of 13 eastern African goat populations and five reference goat populations from West Africa and southern Africa and one European breed

Bootstrap values are shown at each node.

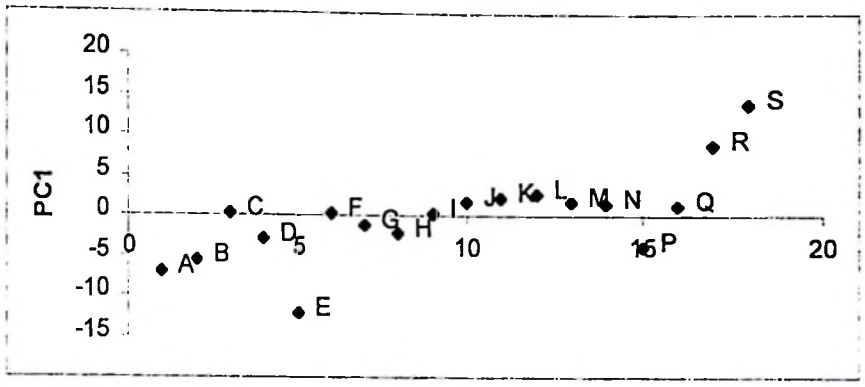
4.2.7 Principal components analysis and multidimensional scaling

The plots of the first three principal components are shown in Figure 11 (a). The first, second and third components accounted for 13.9%, 11.8% and 9.8% of the total variations. The first principal component separated the breeds from West Africa while the second principal component separated the European breed from eastern African goats. The breeds from southern Africa were separated from the eastern African goats by the third principal component. The first ten principal components accounted for 77.7% of the variability. Figure 11 (b) shows the three-dimensional PCA plot constructed using the first three principal components simultaneously. The 3-dimensional plot indicated that the eastern African goats are very close to each other. However, close inspection of the plot revealed two clusters of eastern African goats, the cluster of Ethiopian – Kenyan goats and the cluster of Tanzanian goats. Unlike the dendrograms, the PCA plot showed that the West African goat populations are well separated from all eastern African goat populations while the goats of southern Africa are closely related to eastern African goats. It can also be noted that the European breed (Toggenburg) is well separated from all African goat populations.

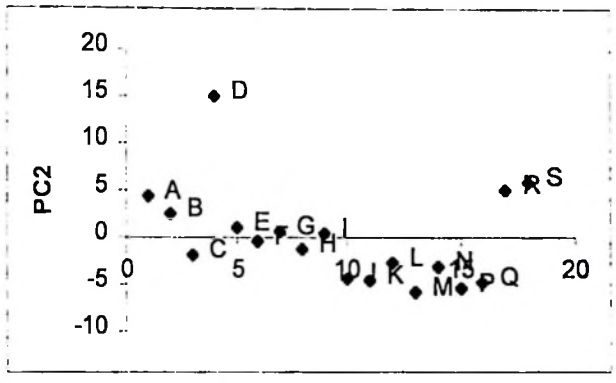
The graphical representation of the D_A genetic distance matrix using multidimensional scaling (Figure 12) showed that the eastern African goat populations were clustered into two groups, a group of Tanzanian populations and a group of Ethiopian-Kenyan

populations. Among the Tanzanian populations, the Newala goats were separated from the other goat populations. The reference breeds from southern Africa (Tswana and Venda) as well as from West Africa (WAD and Red Sokoto) were separated from the two groups of eastern African goats. The European breed (Toggenburg) was well separated from all African breeds. The value of the stress statistic was 0.13.

(i)



(ii)



(iii)

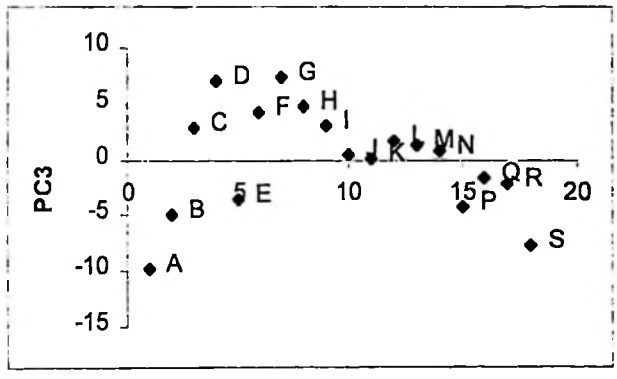


Figure 11(a): Plots for the first (i), second (ii) and third (iii) principal components to show the separation and associations of eastern African goats and five reference breeds

The letters represent the breeds: A – Tswana, B – Venda, C – Kenyan Small East African, D – Toggenburg, E – Landim, F – Boran, G – Galla, H – North East Highland, I – Afar, J – Tanzanian Coastal, K – Ugogo, L – Ujiji, M – Maasai, N – Sukuma, P – Newala, Q – Mbeya, R – Red Sokoto and S – WAD.

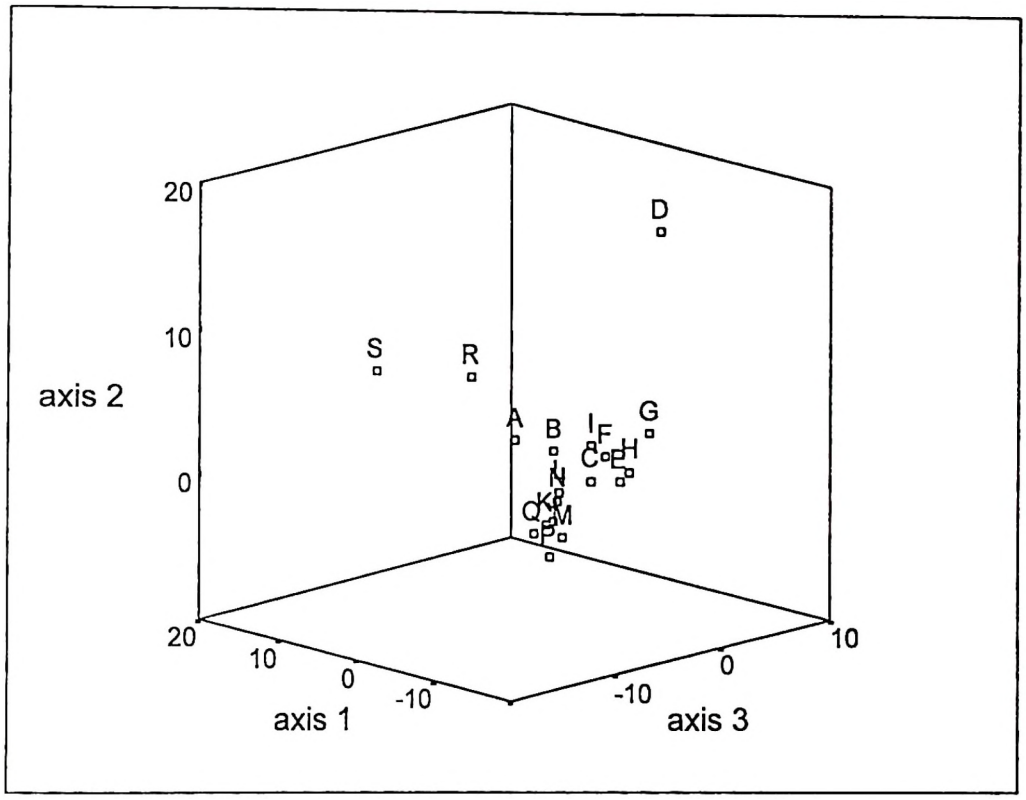


Figure 11 (b): Three-dimensional scatterplot showing the genetic relationships of 13 eastern African goat populations and five reference goat populations from West Africa and southern Africa and one European breed

The first three principal components are plotted with the first, second and third components represented as axes 1, 2 and 3, respectively. The letters represent the breeds: A – Tswana, B – Venda, C – Kenyan Small East African, D – Toggenburg, E – Landim, F – Boran, G – Galla, H – North East Highland, I – Afar, J – Tanzanian Coastal, K – Ugogo, L – Ujiji, M – Maasai, N – Sukuma, P – Newala, Q – Mbeya, R – Red Sokoto and S – WAD.

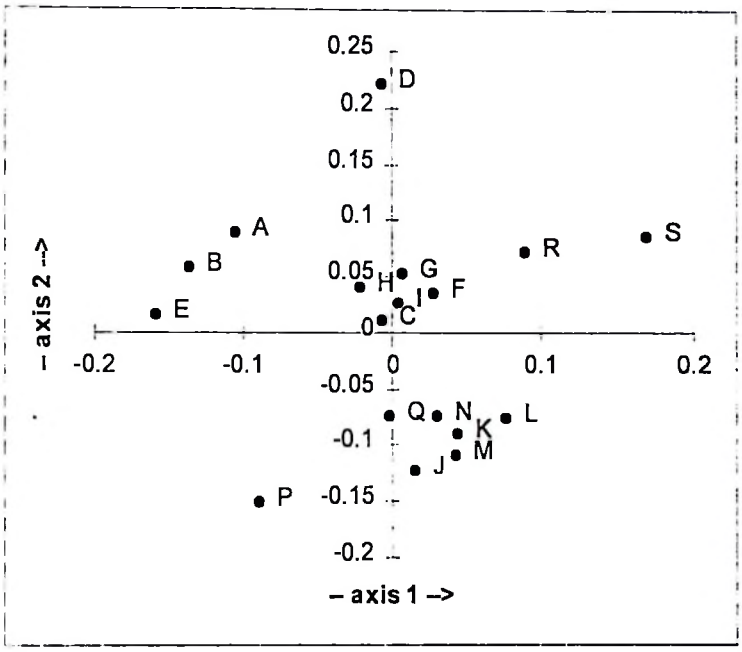


Figure 12: Graphical representation of the D_A genetic distances among 13 eastern African goat populations and five reference goat populations using multidimensional scaling

The letters represent the breeds: A – Tswana, B – Venda, C – Kenyan Small East African, D – Toggenburg, E – Landim, F – Boran, G – Galla, H – North East Highland, I – Afar, J – Tanzanian Coastal, K – Ugogo, L – Ujiji, M – Maasai, N – Sukuma, P – Newala, Q – Mbeya, R – Red Sokoto and S – WAD.

CHAPTER 5

5.0: DISCUSSION

5.1 Genetic diversity within and among sub-Saharan African goats

The main focus of this study was to examine the genetic diversity within- and between-populations of goats from sub-Saharan Africa using allele frequencies data obtained from microsatellite analysis. Microsatellite data can be used to produce genetic distances among the breeds, the raw material for phylogenetic analyses. Barker (1994) recommended that, for studies of genetic distance, microsatellite markers to be used should have more than four alleles to reduce the standard error of the distance estimates. In this study the total number of alleles per locus ranged from 8 (MAF35) to 23 (OarFCB304). This observed number of alleles for each locus demonstrates that all the microsatellite loci were sufficiently polymorphic in the goat populations studied. Mean gene diversity for each locus ranged from 0.321 (MAF35) to 0.796 (MAF65) indicating that the markers used were a mixture of slightly less polymorphic markers and very polymorphic markers in most of the populations. The use of a mixture of very polymorphic and less polymorphic microsatellites reduces the danger of overestimating genetic variability which might occur if only very polymorphic loci are used (Wimmers *et al.*, 2000).

The first step in studying genetic differentiation among populations within a species is to test for HWE conformity. HWE testing is done in order to see if populations can be characterised by allelic frequency instead of genotypic frequency. In this study, the exact test for HWE indicated that out of 285 locus-population combinations, 190 tests showed significant ($P \leq 0.05$) departures from HWE. This was more than what would be expected to occur by chance alone. Since the genotype frequencies deviated from HWE proportions, the expected average heterozygosity was not equal to the average proportions of heterozygotes in the populations studied, but rather, it should be considered as the probability that two randomly chosen genes from a population are different (Nei, 1987). According to Nei (1987), this probability is called gene diversity. Gene diversity is equivalent to heterozygosity in a randomly mating diploid population and can be defined in terms of gene frequencies in the same way as that of heterozygosity (Nei, 1987).

Deviations from expected HWE values may be caused by many factors: inbreeding or outbreeding, selection, population subdivision, migration or gene flow from an external population and existence of null alleles. The estimates for the coefficient of inbreeding (f) gave positive average values in all populations. Positive f values indicate deficit of heterozygotes, thus providing evidence for inbreeding to be common in the populations studied. However, this conclusion does not hold, since by investigating all the tests at all

loci (Appendix 3), all populations showed both positive and negative *f* values. Negative *f* indicates excess heterozygotes thus revealing outbreeding. In addition, except the Ndebele goats, none of the loci analysed showed systematic deviations across all populations. Though no problems were encountered in amplifying PCR fragments at any locus, the problem of null alleles cannot be excluded as some of the markers (MAF35, MAF209 and ILSTS44) were monomorphic in some of the populations. Non-amplifying alleles due to mutations at PCR primer sites have been frequently reported in microsatellite analysis (Arranz *et al.*, 1998), especially when suitable markers in one species are used in studies of another species. In this study, most of the primers were bovine and ovine and the overall deficit of heterozygotes observed in all populations suggests that this was probably the true situation in these cases. However, in a large-scale population survey of genetic variation like this where there is no pedigree information, it is difficult to say which factor caused the deviations. According to Callen *et al.* (1993), if the genotype proportions at a locus are not in HWE for a number of populations, perhaps it is an evidence of selection or an indication of null alleles. Conversely, if a population deviates significantly from HWE at a number of independent loci, it suggests that the population may actually be composed of discrete demes, subject to undergoing nonrandom mating. Therefore, it is possible that the populations were actually subdivided as samples were collected from more than one area (i.e. administrative districts) for each breed. However, it should be remembered that non-compliance with HWE does not indicate the reason for non-compliance. A

demonstration that the populations did not have HWE genotypic frequencies only means that at least one of the assumptions (i.e. large population, random mating, with no mutation, selection and migration) for HWE did not hold for those populations.

5.1.1 Genetic variation within breeds

Mean number of alleles observed over a range of loci in different populations is considered to be a reasonable indicator of genetic variation within the populations, provided that the populations are at mutation-drift equilibrium and that sample size is more or less the same (MacHugh *et al.*, 1997). Generally, all populations studied revealed high level of allelic diversity, the mean number of alleles ranged from 5.26 ± 0.464 in Djallonke to 7.05 ± 0.516 in Mubende goats. When the mean number of alleles per locus and gene diversity (expected heterozygosity) are compared among the goat breeds studied, it is evident that the African (except Djallonke for the number of alleles and Pafuri for the expected heterozygosity) and Asian populations had slightly higher values than the European breeds. This phenomenon, of lower level of genetic variation in European breeds, as measured by allelic and gene diversities, has also been observed in cattle (MacHugh *et al.*, 1997) and has been attributed to the recent evolutionary history of European populations. Reproductive isolation and artificial selection, with the concomitant increase in the level of inbreeding, may have contributed to the slightly lower levels of genetic variability within the European breeds.

A more appropriate measure of genetic variation within a population is gene diversity (average expected heterozygosity) (Nei,1987). Gene diversity for each breed ranged from 54.2% in Pafuri to 67.2% in Ndebele. This level of genetic diversity is similar to the values reported in Swiss goat breeds (51 – 58%) for 20 microsatellite loci (Saitbekova *et al.*, 1999), but is slightly lower than those reported in Chinese goat breeds (77.7 – 82.3%) for 6 microsatellite loci (Yang *et al.*, 1999). It is possible that the small number of markers (6) used in the case of the Chinese breeds were all very polymorphic and thus overestimated the expected heterozygosity. The levels of gene diversity observed in the goats of sub-Saharan Africa are comparable to the values reported in African cattle (MacHugh *et al.*, 1997) and chicken (Wimmers et al., 2000) but below the values observed in humans (Bowcock *et al.*, 1994). This is probably, because of strictness of cultural norms for avoiding mating between closely related individuals in humans, which is not commonly the case in livestock.

5.1.2 Population differentiation

The two estimators of population subdivision indicated that the levels of population differentiation were considerable as all values for G_{ST} and θ differed from zero and the exact test of population differentiation indicated significant differentiation between pair-wise comparisons. The breed assignment test confirmed the differentiation among the breeds. In the assignment test 92.1% of all animals genotyped for 19 microsatellite loci

were correctly assigned to their original breed. Reproductive isolation due to geographic barriers or socio-political reasons had caused the genetic differentiation observed between the breeds. Since there have been no deliberate efforts undertaken to create distinct goat breeds in sub-Saharan Africa, founder effects and genetic drift may have played a major role in the differentiation of African populations. The multi-locus G_{ST} and θ values indicated that between 14.6 and 15.7 % of the total genetic variation was due to population differences, the remaining 85.4 (for G_{ST}) and 84.3% (for θ) corresponded to differences among individuals within the populations. These values for the between-breed variation are close to that reported in Swiss goats (17%) (Saitbekova *et al.*, 1999) using microsatellites and higher than the amount of variation observed among goat breeds (10.7%) from Africa, Middle East, Asia and Europe using mtDNA analysis (Luikart *et al.*, 2001). Furthermore, the values observed in this study are slightly higher than that reported in European cattle breeds (10.5%) (MacHugh *et al.*, 1998) and within the range found in humans (10 – 20% (Cavalli-Sforza *et al.*, 1994). Analysis of the estimators of population differentiation can provide an indication of the evolutionary forces acting on individual markers. In this study, considerable breed differentiation was observed at locus MAF35 (48.1 and 50.4% for G_{ST} and θ, respectively, of the total variation was due to between breed differences). This is because locus MAF35 was monomorphic for different alleles in most of the breeds; that is, there were few overlapping alleles between populations. This suggests that genetic drift is operating at this locus and thus, causing fixation or loss of alternative alleles in different populations.

Probably the locus is close to a region of the genome that is under strong influence of natural selection.

5.1.3 Genetic relationships

Three different measures of genetic distance (D_S, D_A and $(\delta\mu)^2$) between pairs of breeds were estimated. These measures were selected due to their superior performance in phylogeny reconstruction when using microsatellites (Takezaki and Nei, 1996). The D_S distance is based on the infinite alleles model and the D_A is based on genetic drift model. The $(\delta\mu)^2$ distance assumes the stepwise mutational model which is thought to be appropriate for microsatellite data. The data in Tables 9, 10 and 11 for the three genetic distance measures showed that the values for $(\delta\mu)^2$ distance were higher than those for D_S and D_A distances. A close examination of the three measures of genetic distances, revealed that the distances between eastern African breeds and southern African breeds were smaller than the distances between either West African breeds and southern African breeds or West African breeds and eastern African breeds. The largest distance values were those between West African breeds and the breeds from southern Africa. This indicates that the breeds of southern Africa were more closely related to the eastern African breeds than to the breeds of West Africa.

The phylogenies constructed from the three distance measures indicated that D_S and D_A grouped the breeds according to their geographic locations of origin but $(\delta\mu)^2$ distances did not show a clear pattern of population clustering. Unlike the phylogenies constructed from D_S (Figure 2) and D_A (Figure 3), the phylogeny generated from $(\delta\mu)^2$ distances (Figure 4) showed that the two European breeds did not cluster together: Toggenburg was grouped together with the West African goats while Grisons Striped was in a cluster containing southern African populations and eastern African populations. Also the East African goats (Maasai, Mubende and Kigezi) were separated into different groups. From this observation it can be said that the genetic distances estimated from $(\delta\mu)^2$ were less effective in revealing the true genetic relationships of the populations compared to those estimated from D_S and D_A. Similar observation has been reported by Kantanen (1999) who found the values for $(\delta\mu)^2$ distances to be less reliable, compared to D_A distances, in recovering the evolutionary history of North European cattle breeds.

Although the phylogenies constructed from D_S (Figure 2) and D_A (Figure 3) distances were slightly different, there was a general accordance in their topology. A clear separation was observed between the group of European-Asian breeds and the group of African breeds indicating that European breeds are more closely related to Asian breeds (breeds from Mongolia and Nepal) than to African breeds. This is probably due to admixture or introgression of Europeans breeds into Asian breeds. For a long time European breeds have been exported to Asia to improve milk production of local breeds

through crossbreeding but no widespread crossbreeding has been practised between European and African breeds. However, the multidimensional scaling and principal component analyses showed that the European and Asian breeds are well separated from each other. Hence, it is conceivable that the close relationships between European and Asian breeds evident in the phylogenetic trees may be a result of the limitation of phylogenetic analysis procedure to show the relationships of distantly related breeds when using microsatellite data.

A close investigation of the two phylogenies (D_S-and D_A-based dendrograms) showed that the phylogeny based on D_A distances clearly separated all the breeds from outside Africa from the African breeds. In the phylogeny constructed from the D_S distances the Arab (Emirates) breed (from the Middle East) was grouped together with the African breeds. Bootstrap values (percentage of occurrence of a node in 1000 bootstrap resampling of loci) ranged from 15 to 95% for D_S-based phylogeny and 38 to 97% for D_A-based phylogeny. This indicates that the topology for the phylogeny constructed from D_A distances was more reliable than that of the phylogeny based on the D_S distances. This is in agreement with the findings of Arranz *et al.* (1998) who, in a study of Spanish sheep, reported higher bootstrap values for the dendrogram constructed from D_A distances, compared to that constructed from D_S distances. Similarly, in a simulation study to compare different genetic distance measures, Takezaki and Nei (1996) concluded that D_A genetic distance gives the best performance in obtaining the correct

tree topology from microsatellite data. In view of this, it can be said that the phylogeny based on D_A distances recovered the evolutionary relationships of the populations studied better than that based on the D_S distances. From the D_A-based phylogeny it is evident that the breeds were grouped according to their geographic origins. A similar observation of populations clustering according to their geographic origins has been reported in humans (Bowcock *et al.*, 1994), cattle (MacHugh *et al.*, 1997), and chickens (Wimmers *et al.*, 2000). This implies that geographically adjacent populations are more genetically related, probably because of founder effects and interbreeding, especially around bordering areas.

The three-dimensional plot for principal component analysis (PCA) (Figure 5) and the multidimensional scaling (Figure 6) support the grouping of the African populations according to their geographic origins, i.e. West Africa, eastern Africa and southern Africa. The eastern African breeds were found to be in between the West African breeds and the southern African breeds. This confirmed the observation that the largest genetic distances were those between southern African breeds and West African breeds. Hence, it can be said that the eastern African breeds are more closely related to either of the two groups (West African and southern African breeds) than the two groups are related to each other. Unlike the results of the phylogenies, the plots for PCA and the multidimensional scaling indicated the separation between the European and Asian breeds more clearly. This is in agreement with the results of the study of domestication

and phylogeography of Taurine and Zebu cattle by MacHugh *et al.* (1997) which reported the PCA, using microsatellite allele frequencies, to be a powerful tool for revealing the underlying evolutionary history among distantly related populations of cattle from Africa, Europe and Asia. While the PCA results of the present study provide no evidence for the Mongolian Cashmere breed to cluster together with the Bandipur breed, the multidimensional scaling grouped the two breeds together, but the interpoint distance between the two breeds was large. This implies that, though grouped together by the phylogenies, Mongolian Cashmere and Bandipur are somewhat distinct breeds.

5.2 Genetic variation within and among eastern African populations

5.2.1 Within population genetic diversity

The mean number of alleles per locus ranged from 5.53 to 6.53 and the average level of gene diversity ranged from 55.3% to 61.3% indicating that all populations had relatively high within-population genetic variability. The narrow range of these values indicates that the differences, in terms of within-population genetic variability, are not large among the eastern Africa goat populations. Intercrossing between adjacent populations may be the main factor homogenising the level of genetic variation and thus retaining this relatively high level of within-population genetic diversity in all populations. The levels of genetic variations observed in eastern African populations are comparable to

those observed in other goats of sub-Saharan Africa (section 5.1.1). This is most likely a direct reflection of the absence of strict breed standards in Africa similar to the western definition in which animal registration ensures breed boundaries. The mean number of alleles and the gene diversity observed in eastern African goat populations corresponds to that reported in East African cattle from microsatellite analysis (Okomo *et al.*, 1998). This implies that the ruminant livestock in eastern Africa are, in large part, quite outbred. Deviations from HWE were observed in all eastern African goat populations and average observed heterozygosity was always lower than expected heterozygosity. Moreover, the values of f in all populations were positive suggesting some inbreeding within the populations. However, deviation from HWE can be caused by other factors such as assortative mating, population subdivision, selection, migration or gene flow from an external population and presence of experimentally undetectable alleles (null alleles). As explained above, in a genetic variation survey like this, it is not possible to determine which factors caused these deviations.

5.2.2 Population differentiation

The overall values of the measures of population genetic differentiation (G_{ST} and θ) indicated that the proportion of genetic variations due to interpopulation subdivision among eastern African goat populations was 12% and 11% for G_{ST} and θ, respectively. As expected, these values are slightly lower than those reported for the first data set on

the whole of sub-Saharan Africa (section 5.1.2), indicating that the extent of genetic differentiation is low for populations which are geographically in close proximity to each other compared to those which are far apart. This could be attributed to admixture between adjacent populations or recent separation from a common ancestral population. However, the exact test for population differentiation indicated that the populations are significantly differentiated in terms of allelic composition across all the 19 loci analysed. The results of the assignment test indicated that on average, 79.1% of all animals from the 13 eastern African goat populations were correctly assigned to their source populations. This may reflect a relatively low level of gene flow between populations probably due to limited movements of populations across national borders.

5.2.3 Individual animals genetic relationships

The neighbour-joining dendrograms in which each individual animal from the eastern African goat populations was treated as a taxonomic unit are shown in Figures 7 and 8. The dendrogram for the Tanzanian goats (Figure 7) indicated that, among the populations analysed, Newala and Tanzanian Coastal goats appeared to be tightly clustered together. However, close investigation revealed that, only Newala goats formed a distinct cluster. This is because 94% of the animals from the Newala population were found in a single clade and only 6% of the animals from other populations were found in this clade. In the breed assignment test 100% of the animals

from the Newala goats were correctly assigned to their source population. Also the plot (Figure 12) for multidimensional scaling confirmed the separation of the Newala goats from other Tanzanian goats. This indicates that the Newala goats are very distinct from the other Tanzanian goats, probably, because of little mixing of the Newala goats with other Tanzanian goats as a result of limited movement and trade between the Newala district and other parts of Tanzania. Although most of the Tanzanian Coastal goats appeared in a single clade, in this clade there were an appreciable number of animals from the other populations. Out of the 24 animals in this clade, 9 (38%) were from other populations (Ugogo, Mbeya, Maasai). This was expected as animals from the goat populations in the interior are normally brought to the Coastal region through trade resulting in substantial interbreeding. The rest of the populations (Ugogo, Ujiji, Sukuma, Mbeya and Maasai), did not show a clear pattern of clustering. The lack of population-specific clustering for these populations is probably a consequence of a recent evolutionary history or of intermixing and interbreeding as communities keeping these populations intermingle. The dendrogram showing the relationships between individual animals from the Ethiopian and Kenyan populations (Figure 8) indicated that, only Kenyan Small East African and Afar goat populations appeared in a distinct clade (17 out of 25 animals clustered together). The rest of the populations showed a very low degree of population structuring: Animals from North East Highland, Boran, and Galla populations were distributed both at the top and at the bottom of the tree. The small tendency for clustering together for individuals from the same population may be a result

of a recent separation from a common ancestral population or population admixture. An examination of the two dendrograms (Figures 7 and 8), shows that the reference breed (Tswana goats) from southern Africa was differentiated from the eastern African goat populations.

5.2.4 Genetic relationships among populations

The phylogenetic relationships of eastern African goat populations is shown in Figures 9 and 10. There were some differences in the topology of the two phylogenies because the position of some populations differed between the two dendrograms. However, both phylogenies clearly showed the differentiation of the Tanzanian populations from the Ethiopian and Kenyan populations. This is probably because of relatively longer period of separation of the Tanzanian goats from the Ethiopian-Kenyan goats due to geographical distances. Also, the continuing movements of the Ethiopian and Kenyan pastoralists between southern Ethiopia and North-east Kenya (the place where the goats were sampled) may be responsible for admixture between Ethiopian and Kenyan populations than is experienced between Kenyan and Tanzanian populations. As reported above, in the study of sub-Saharan African goats, the phylogeny constructed from D_A distances represented better the evolutionary relationships of the eastern African populations. In Figure 10 the eastern African populations are clearly separated from other African goats (used as reference breeds). In this phylogeny (Figure 10) the

populations are grouped according to countries of origin indicating that founder effects and genetic drift may have played a major role in the differentiation of the populations between countries. This is supported by the fact that, among the Kenyan populations, the Small East African goats, sampled from the south-west, are separated from the Boran and Galla sampled from the north-east. Also, the Tanzanian populations were grouped according to their locations of origin. However, the bootstrap values are low for most of the nodes. This indicates that the topology of eastern African goat population clade is not robust implying that there is low level of genetic differentiation between populations. The low level of differentiation between populations is also reflected in the branching orders, which vary depending on the genetic distance measure used.

The principal component analysis and the multidimensional scaling support the separation of the Tanzanian populations from the Ethiopian-Kenyan populations. Although the principal component analysis was carried out on allele frequencies and the multidimensional scaling was based on genetic distances, the two plots of breed relationships were in good agreement. The small inter-point distances between the two groups may be the result of recent separation from a common ancestral population, coupled with lack of selection to create standardised populations or breeds. Also, as was the case for sub-Saharan African goats, the PCA plot showed that the goats from southern Africa are genetically closer to eastern African goat populations than to the goats of West Africa.

5.3 Evolutionary relationships and origins of domestic goats

The current classification of goat breeds is based on body size, ear shape and length, horn types and functions. It is assumed that the differences in such traits reflect distinct origins or genetic identity. The classification of the breeds studied according to these criteria is shown in Tables 1 and 2. Investigation of the phylogenies indicated that the phylogeny constructed from D_A distances best grouped the African breeds according to their phenotypic characterisation. Among the West African breeds, the Maure goats, classified as "Intermediate twisted horn" breed type, were separated from the Dwarf small horn breeds, the WAD and Djallonke. Also the PCA plot indicated that the WAD and the Djallonke are more close to each other than to the Maure. Similarly, for the eastern African goat populations, there is clear separation of the Ethiopian - Kenyan (North East Highland, Afar, Boran and Galla) breeds, classified as "Intermediate East African goats", from the Tanzanian populations which are classified as "Small East African goats". However, the Landim goats, classified as Small East African goats, clustered together with the long lop-eared breeds (Tswana and Venda). This made the conclusion that the breeds were related according to phenotypic characteristics not to hold. The lack of long lop-eared goats from Northern Africa in this study made it impossible to show that the genetic relationships of the breeds were according to physical types rather than geographic locations of origin as explained above. However, it has been reported that morphological differences between populations are not taken into

account by neutral molecular markers such as microsatellite (Hartl and Clark, 1997). It is genetic drift process that causes the genetic differentiation between populations detected by the use of neutral molecular markers. Hence, breed relationships were more related to geographical locations of the breeds than to the morphological differences between the breeds.

Domestication of goats is thought to have occurred in the highlands of western Iran 10,000 years before present (Zeder and Hesse, 2000). Archaeological evidence based on changes in skeletal morphology, particularly population-wide reduction in body size following human-controlled management, changes in age and sex profiles resulting from controlled breeding and selective slaughter of young males, and metrical analyses of skeletal collections from modern wild goats, have been used to locate the place of early domestication. It is believed that goats entered Africa from the Middle East and were first domesticated in Egypt. From Egypt goats moved westwards to West Africa and southwards through Sudan and Ethiopia to East Africa. Thus, the present-day goat populations may indicate a clue to their ancient ancestors. The phylogenetic relationships among the breeds in Figure 3 support the evolutionary split of African, European and Asian breeds from the Middle East breeds represented by the Arab (Emirates) breed. Multidimensional scaling and the analysis of allele frequencies by PCA revealed the separation of African, European and Asian breeds more clearly, thus, providing a strong indicator of different ancestry. Luikart *et al.* (2001) suggested three

distinct origins of domestic goats from mtDNA analysis. According to Luikart *et al.* (2001) domestic goats arose from three independent domestication locations, the first being in the southern Turkish region, the second place is the Zagros Mountains and the third place is the Indus Basin. However, the intercontinental mixing and interbreeding of goat breeds through human movements and trade may have disturbed the phylogenetic relationships of the breeds. Population admixture has been shown to affect the relationships among breeds (Blott *et al.*, 1998). For the African breeds, the phylogeny (Figure 3) indicates the differentiation of eastern African populations, West African populations and southern African populations from Ethiopian populations. This supports the contention that, from the centre of domestication in north-east Africa, two movements were followed in the spread of goats in Africa, one westwards to west Africa and another one southwards to East Africa. With regard to the goats of southern Africa, there are two hypotheses which have been suggested to explain the origins of all ruminant livestock in southern Africa (Smith, 2000). The first hypothesis is that the African Iron Age people moved southwards with their livestock from Cameroon and the second hypothesis is that East African herders moved far southwards to southern Africa. The second hypothesis is supported by the widespread distribution of Small short-eared goats, like those of East Africa, in southern Africa – Malawi, Mozambique, Zambia, Zimbabwe and some parts of South Africa. The present study supports the second hypothesis. This is because, the genetic distances between West African breeds and southern African breeds were larger than those between eastern African breeds and

southern African breeds. This was reflected in the plot for multidimensional scaling. The figure presented (Figure 6) had a stress statistic of 0.11 implying that the plot provides a good representation of the population relationships. The PCA plot also indicates that the southern African breeds are more closely related to eastern African breeds than to West African breeds.

5.4 Implication for conservation programmes

Resources are limited while large numbers of indigenous breeds are in danger of extinction. Thus, making a decision about which breeds will be prioritised for conservation programmes is unlikely to be an easy task. One criterion used to classify a specific population as an important genetic resource is its distinctiveness from other populations within the species. Hence, genetic uniqueness is an important input into priority setting for conservation programmes. It is assumed that conservation of breeds/populations that are genetically different is a means of ensuring that the species will have sufficient genetic variation to adapt to environmental changes and respond to selection in future generations. This is due to the fact that breeds/populations that are genetically divergent from each other are more likely to possess different alleles and gene combinations affecting a range of traits that may be important for adaptation for future production environments.

Looking at the phylogenies (Figures 2 and 3), it seems that there is sufficient differentiation among the goat populations of West, eastern and southern Africa. Thus, it is important to consider separate conservation programmes for the three regions of sub-Saharan Africa. Within each region, populations can be selected based on their within-population genetic variability, giving priority to the populations with the highest number of alleles and gene diversity. The phylogenies for eastern African populations revealed that there is also differentiation of goat populations between countries, hence, necessitating that each country should have its own conservation programme for priority breeds. Since the difference in within population diversity, as indicated by the mean number of alleles per locus and gene diversity, was not large, a few populations can be selected without losing much of the genetic diversity. However, in selecting populations for conservation, other factors should be considered apart from the genetic uniqueness of the populations and within population variability. According to Hammond (1993), the following factors should be considered: degree of endangerment, adaptation to a specific environment, traits of economic importance, unique traits and cultural or historical values. For instance, among the goats of sub-Saharan Africa, the Pafuri of Mozambique has been listed by FAO as an endangered breed, the West African Dwarf from the coast of West Africa is adapted to the tsetse infested areas, the Red Sokoto of Nigeria is valued for its skin and the Mubende of Uganda is useful for leather production. Although, some of these populations (e.g. Pafuri) may have low genetic variability, because of their special attributes, they deserve to be conserved.

The use of microsatellite loci is regarded as the best available method to assess genetic variability among closely related populations. However, it is questionable whether by using microsatellites, which are non-coding loci, the populations carrying unique traits that may be beneficial now or in the future can be identified. It has been suggested that the final target for genetic characterisation should be to detect variations within genomic regions having effects on phenotypes of economic importance. In recent years, efforts have been put on the identification and mapping of genes in economically important livestock. Thus, in the near future it will be possible to select breeds for conservation based on the presence of unique alleles or allelic combinations coding for specific production or adaptability related traits. Furthermore, it will be possible to select breeds based on microsatellite loci associated with genomic region coding for economically important traits as currently microsatellites are the markers of choice for gene mapping studies in livestock species. Nonetheless, microsatellite marker technology will remain useful in genetic diversity studies as a first step in short-listing candidate breeds for conservation before other data (e.g. economic value or resistance genes) become available or can be used in combination with such data.

CHAPTER 6

6.0 CONCLUSSIONS AND RECOMMENDATIONS

The two studies have shown that microsatellite analysis is very useful in examining the genetic relationships of closely related goat populations in sub-Saharan Africa and in assigning individual animals to their source populations. Specifically the two studies have indicated that:-

(i) The level of within-population genetic diversity, as measured by the mean number of alleles per locus and gene diversity, in African goat populations is relatively high and that the differences in terms of number of alleles and gene diversity among the populations are small.

(ii) The two estimators of population subdivision, G_{ST} and θ indicated that in the goats of sub-Saharan Africa, between 14.6% and 15.7% of the total genetic variation is between breeds/populations, the rest being among individuals within the breeds/populations. When only eastern African goats are considered, the level of genetic variation between breeds/populations decreases to 11% - 12%.

(iii) The phylogenetic trees grouped the goats of sub-Saharan Africa according to their geographic origins indicating that the goats of eastern Africa, West Africa and southern Africa are substantially differentiated from each other. Also, the case study of eastern African goats showed that there is substantial

population differentiation according to country or even regions within country.

(iv) In this study, the phylogeny based on D_A distance was found to be more reliable compared to those based on other distance measures($(\delta\mu)^2$ and D_S) in revealing the evolutionary relationships of closely related goat populations of sub-Saharan Africa. Furthermore, D_S-based phylogeny differentiated the breeds according to their geographic origins more clearly than the phylogeny based on $(\delta\mu)^2$distance.

(v) The principal component analysis and multidimensional scaling revealed better the genetic relationships of distantly related populations (African, European and Asian breeds) than did the phylogenetic trees. On the other hand, the phylogenetic trees based on D_A and D_S distances revealed more clearly the genetic relationships of closely related populations (populations within the same country) compared to the principal component analysis .

(vi) The study has shown the necessity of having conservation programmes for goat populations of each region of sub-Saharan Africa (i.e eastern Africa, West Africa and southern Africa). Moreover, given the between country differentiation, it is important for each country to have its own conservation programme. Since the differences among the populations in terms of within-population variability is not large and the genetic distances between pairs of

populations within the country are small, few populations per country can be selected without losing much of the genetic diversity.

(vii) It is suggested that the socio-economic value and performance of the different indigenous populations be established in order to facilitate and rationalize the establishment of conservation and sustainable utilisation programmes. Information on traits such as longevity, fertility, mortality, disease resistance, feed and management requirements should be gathered. These characters might significantly contribute in making a decision about which breeds will be prioritised for conservation programmes.

REFERENCES

Aboagye, G. (1992). Characterisation and evaluation of Sanga and West African shorthorn cattle on the Accra plains. In: *African Animal Genetic Resources: Their characterisation, conservation and utilisation. Proceedings of the research planning workshop* held at ILCA, Addis Ababa, Ethiopia 19 - 21 February, 1992. ILCA (International Livestock Centre for Africa) Addis Ababa, Ethiopia. pp 73 – 75.

Aboagye, G.S.; Tawah, C.L. and Rege, J.E.O. (1994). Shorthorn cattle of West and Central Africa. III. Physical, adaptive and special genetic characteristics. *World Animal Review* 78 (1): 22 -32.

Allen, P.J.; Amos, W.; Pomeroy, P.P. and Twiss, S.D. (1995). Microsatellite variation in grey seals (*Halichoerus grypus*) shows evidence of genetic differentiation between two British breeding colonies. *Molecular Ecology* 4 : 653 – 662.

Allendorf, F.W.; Knudson, K.L. and Leary, R.F. (1983). Adaptive significance of the differences in the tissue-specific expression of a phosphoglucomutase gene in rainbow trout. In: *Proceedings of the National Academy of Science of the USA* 800 : 1397 – 1400.

Allendorf, F.W. and Leary, R.F. (1986). Heterozygosity and fitness in natural populations of animals. In: *Conservation Biology: The science of scarcity and diversity.*(Edited by M.E. Soulé) Sinauer, Sunderland, Massachusetts. pp 57 – 76.

Armour, J.A.L. and Jeffreys, A.J. (1992). Biology and applications of human minisatellite loci. *Current Opinion in Genetics and Development* 2 : 850 – 856.

Arranz, J.J.; Bayón, Y. and San Primitivo, F. (1998). Genetic relationships among Spanish sheep using microsatellites. *Animal Genetics* 29: 435 – 440.

Avise, J.C. (1994). *Molecular markers, Natural history and Evolution.* Chapman & Hall. New York, London. pp 511.

Baker, A.J. and Moeed, A. (1987). Rapid genetic differentiation and founder effect in colonizing populations of common Mynas (*Acridotheres tristis*). *Evolution* 41 : 525 – 538.

Banks, M. and Eichert, W. (2001). WhichRun: a computer program for population assignment of individuals based on multilocus genotype data. http://www-bml.ucdavis.edu/whichrun.htm

Barbancho, M.; Llanes, D.E.; Morera, L.; Garzon, R. and Rodero, A. (1984). Genetic markers in the blood of Spanish goat breeds. *Animal Blood Groups and Biochemical Genetics* 15 : 207 – 212.

Barendse, W.; Vaiman, D. and Kemp, S.J. (1997). A medium-density genetic linkage map of the bovine genome. *Mammalian Genome* 8 : 21 – 28.

Barker, J.S.F. (1994). A global protocol for determining genetic distances among domestic livestock breeds. In: *Proceedings of the 5th World Congress of Genetics Applied to Livestock Production.* 7 – 12 August, 1994 Guelph, Canada. Vol 21: 501 - 508.

Blott, S.C.; Williams, J.L. and Haley, C.S. (1998). Genetic variation within the Hereford breed of cattle. *Animal Genetics* 29 : 202 – 211.

Bowcock, A.M.; Ruiz-Linares, A.; Tomfohrde, J.; Minch, E.; Kidd, J.R. and Cavalli-Sforza, L.L. (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* 368: 455 – 457.

Bradley, D.G.; MacHugh, D.E.; Cunningham, P. and Loftus, R.T. (1996). Mitochondrial diversity and the origins of African and European cattle. In : *Proceedings of the National Academy of Science of the USA* 93 : 5131 – 5135.

Bruford, M.W. and Wayne, R.K. (1993). Microsatellites and their application to population genetics studies. *Current Opinion in Genetics and Development* 3 : 939 – 943.

Buchanan, F.C.; Adams, L.J.; Littlejohn, R.P.; Maddox, J.F. and Crawford, A.M. (1994). Determination of evolutionary relationships among sheep breeds using microsatellites. *Genomics* 22 : 397 – 403.

Burke, T. (1989). DNA fingerprinting and other methods for the study of mating success. *Trends in Ecology and evolution* 4: 139 – 144.

Callen, D.F.; Thompson, A.D. and Shen , Y. (1993). Incidence and origin of 'null' alleles in the $(AC)_n$ microsatellite markers. *American Journal of Human Genetics* 52: 922 – 927.

Carson, H.L. (1983). The genetics of the founder effect. In: *Genetics and Conservation: A reference for managing wild animal and plant populations.*(Edited by C.M.

Schonewald-Cox; S.M. Chambers; B. MacBryde and L. Thomas). Benjamin/Cummings, menlopark, CA. pp 189 – 200.

Cavalli-Sforza, L.L. and Edwards, A.W.F. (1967). Phylogenetic analysis, models and estimation procedures. *American Journal of Human Genetics* 19 : 233 – 257.

Cavalli-Sforza, L.L.; Piazza, A., and Menozzi, P. (1994). *The history and geography of human genes*. Princeton University press, New Jersey

Charlesworth, D. and Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics* 18 : 237 – 268.

Crawford, A.M.; Dodds, K.G. and Pierson (1995). An autosomal genetic linkage map of the sheep genome. *Genetics* 140 : 703 – 724.

Cunningham, E.P.; Loftus, R.T.; MacHugh, D.E. and Bradley, D.G. (1994). Molecular evolution of African, European and Asian cattle. In: *Proceedings of the 5th World Congress of Genetics Applied to Livestock Production.* 7 – 12 August, 1994 Guelph, Canada. Vol 21: 86 - 89.

Devendra, C. and Burns, M. (1983). *Goat production in the tropics.* Commonwealth Agricultural Bureau. Farnham house, Farnham royal, Slough SL2 3BN UK. pp183.

Dowling, T.E.; Moritz, C.; Palmer, J.D. and Rieseberg, L.H. (1996). Nucleic Acids III: Analysis of fragments and restriction sites. In: *Molecular systematics* 2nd edition (Edited by D.M. Hills; C. Moritz, and B.K. Mable). Sinauer Associates, Inc. Publishers. Sunderland, Massachusetts, USA. pp 249 – 320.

Edwards, A.; Civitello, H.; Hammond, H.A. and Caskey, C.T. (1991). DNA typing and genetic mapping with trimeric tandem repeats. *American Journal of Human Genetics* 49: 746 – 756.

Edwards, A.; Hammond, H.A.; Jin, L.; Caskey, C.T. and Chakraborty, R. (1992). Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 12: 241 – 253.

Ellegren, H. (1993). Genome analysis with microsatellite markers. *PhD dissertation.* Swedish University of Agricultural Sciences.

Epstein, H. (1971). *The origin of the domestic animals of Africa.* Vol. II. African publishing corporation. New York, London, Munich. pp 719.

Fahmy, T. (2000). XLSTAT program version 4.3. http://www.xlstat.com

FAO (1995). *World watch list for domestic animal diversity.* 2nd edition. Food and Agriculture Organization of the United Nations, Rome, Italy. pp 769.

FAO (1996). *FAO production year book*, Vol. 50. Food and Agriculture Organization of the United Nations, Rome, Italy. pp 235.

Farris, J.S. (1970). Methods for computing Wagner trees. *Systematic Zoology* 34: 21 – 34.

Farris, J.S. (1971). The hypothesis of nonspecificity and taxonomic congruence. *Annual Review of Ecology and Systematics* 2: 227 – 302.

Farris, J.S. (1981). Distance data in phylogenetic analysis. In: *Advances in Cladistics. Proceedings of the first meeting of the Willi Henning Society* (V.A. Funk and D.R. Brooks edition). New York Botanical garden, New York. pp 3 – 23.

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

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

Franklin, I.R. (1980). Evolutionary change in small populations. In: *Conservation Biology: An evolutionary-Ecological perspective.* (Edited by M.E. Soulé and B.A. Wilcox). Sinauer Associates, Sunderland, MA. pp 135 – 149.

Frankham, R. (1994). Conservation of genetic diversity for animal improvement. In: *Proceedings of the 5th World Congress of Genetics Applied to Livestock Production.* 7 – 12 August, 1994 Guelph, Canada. Vol 21: 385 - 392.

Forbes, S.H.; Hogg, J.T.; Buchanan, F.C.; Crawford, A.M. and Allendorf, F.W. (1995). Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. *Molecular Biology and Evolution* 12 : 1106 – 1113.

Georges, M.; Dietz, A.B.; Mishra, A.; Nielsen, D.; Sargeant, L.S.; Sorensen, A.; Steele, M.R.; Zhao, X.Y.; Leipold, H.; Womack, J.E. and Lathrop, M. (1993). Microsatellite mapping of the gene causing weaver disease in cattle will allow the study of an associated quantitative trait locus. In: *Proceedings of the National Academy of Science of the USA.* 90 : 1058 – 1062.

Goldstein, D.B.; Linares, A.R.; Cavalli-Sforza, L.L. and Feldman, M.W. (1995a). An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139 : 463 – 471.

Goldstein, D.B.; Linares, A.R.; Cavalli-Sforza, L.L. and Feldman, M.W. (1995b). Genetic absolute dating based on microsatellites and the origin of modern humans. In: *Proceedings of the National Academy of Science of the USA* 92 : 6723 – 6727.

Goossens, B.; Waits, L.P. and Taberlet, P. (1998). Plucked hair samples as a source of DNA: reliability of dinucleotide microsatellite genotyping. *Molecular Ecology* 7: 1237 – 1241.

Gottelli, D.; Sillero-zubiri, C.; Applebaum, G.D.; Roy, M.S.; Girman, D.J.; Garcia-moreno, J.; Ostrander, E.A., and Wayne, R.K. (1994). Molecular genetics of the most endangered canid – the Ethiopian wolf *Canis simensis*. *Molecular Ecology* 3: 301 – 312.

Guo, S.W. and Thompson, E.A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48: 361 – 372.

Gwakisa, P.S.; Kemp. S.J. and Teale, A.J. (1994). Characterization of Zebu cattle breeds in Tanzania using random amplified polymorphic DNA markers. *Animal Genetics* 25: 89 - 94.

Haberfeld, A.; Cahaner, A.; Yoffe, O.; Plotsky, Y. and Hillel, J. (1991). DNA fingerprints of farm animals generated by microsatellite and minisatellite DNA probes. *Animal Genetics* 22: 299 - 305.

Hammond, K. (1993). Why conserve animal genetic resources? *Diversity* 9 (3): 30 – 35.

Hammond, K. (1994). Conservation of domestic Animal diversity: Global overview. In: *Proceedings of the 5th World Congress of Genetics Applied to Livestock Production.* 7 – 12 August, 1994 Guelph, Canada. Vol 21: 423 - 430.

Hanotte, O.; Burke, T.; Armour, J.A.L. and Jeffreys, A.J. (1991). Hypervariable minisatellite DNA sequences in the Indian Peafowl *Pavo cristatus*. *Genomics* 9 : 587 – 597.

Hartl, D.L. and Clark, A.G. (1997). *Principles of population genetics*, 3rd ed. Sinauer Associates. Sunderland, Massachusetts. pp 542.

Hodges, J. (1990). The global organisation of animal genetic resources. In: *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production.* Vol. XIV : 466 – 472.

Hunter, M.L. Jr. (1996). *Fundamentals of conservation biology*. Blackwell science, Inc., Cambridge, Massachusetts, U.S.A. pp 482.

Jeffreys, A.J. and Morton, D.B. (1987). DNA fingerprints of dogs and cats. *Animal Genetics* 18: 1 - 15.

Jeffreys, A.J.; Royle, N.J.; Wilson, V. and Wong, Z. (1988). Spontaneous mutation rates to new length alleles at tandem-repetitive loci in human DNA. *Nature* 332 : 278 – 281.

Jeffreys, A.J.; Tamaki, K.; MacLeod, A.; Monckton, D.G. and Neil, D.L. (1994). Complex gene conversion events in germline mutation at human minisatellites. *Nature Genetics* 6 : 136 – 145.

Jeffreys, A.J.; Wilson, V. and Thein, S.L. (1985). Hypervariable 'minisatellite' regions in human DNA. *Nature* 314: 67 - 73.

Joubert, D.M. (1973). Goats in the animal agriculture of Southern Africa. *Zeitschrift fur Tierzüchtung und Züchtungsbiologie* 90(3) : 245 – 262.

Kantanen, J. (1999). Genetic diversity of domestic cattle (*Bos taurus*) in North Europe. *PhD thesis*. University of Joensuu, Finland.

Karp, A.; Isaac, P.G. and Ingram, D.S. (1998). *Molecular tools for screening biodiversity. Plants and Animals*. Chapman & Hall. London, Weinheim, New York, Tokyo, Melbourne, Madras. pp 498.

Kashi, Y.; Tikochinsky, Y.; Genislav, E.; Iraqi, F.; Nave, A; Beckmann, J.S.; Gruenbaum, Y. and Soller, M. (1990). Large restriction fragments containing poly-TG are highly polymorphic in a variety of vertebrates. *Nucleic Acids Research* 18: 1129 - 1132.

Kemp, S.J. (1992). The potential contribution of biotechnology in breed characterisation. In: *African animal genetic resources: Their characterisation, conservation and utilisation. Proceedings of the research planning workshop* held at ILCA, Addis Ababa, Ethiopia 19 - 21 February, 1992. ILCA(International Livestock Centre for Africa) Addis Ababa, Ethiopia. pp 23 - 28.

Kikkawa, Y.; Yonekawa, H.; Suzuki, H. and Amano, T. (1997). Analysis of genetic diversity of domestic water buffaloes and anoas based on variations in the mitochondrial gene for cytochrome b. *Animal Genetics* 28 : 195 – 201.

Kuhnlein, U.; Zadworny, D.; Dawe, Y.; Fairfull, R.W. and Gavora, J.S. (1989). DNA fingerprinting: A tool for determining genetic distances between strains of poultry. *Theoretical and Applied Genetics* 77: 669 - 672.

Lacy, R.C. (1987). Loss of genetic variation from managed populations: Interacting effects of drift, mutation, immigration, selection and population subdivision. *Conservation Biology* 1: 143 – 158.

Lande, R. and Barrowclaugh, G.F. (1987). Effective population size, genetic variation, and their use in population management. In: *Viable populations for conservation* (Edited by M.E. Soulé). Cambridge University Press. pp 87 – 123.

Leberg, P.L. (1992). Effects of a population bottleneck on genetic variation as measured by allozyme electrophoresis. *Evolution* 46: 477 – 494.

Loftus, R.T.; MacHugh, D.E.; Ngere, L.O.; Balain, D.S.; Badi, A.M.; Bradley, D.G.; and Cunningham, E.P. (1994). Mitochondrial genetic variation in European, African and Indian cattle populations. *Animal Genetics* 25: 265 - 271.

Lomker, R. and Simon, D.L. (1994). Costs of and inbreeding in conservation strategies for endangered breeds of cattle. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock production*, Vol. XIV: 434 – 442.

Luikart, G.; Gielly, L.; Excoffier, L.; Vigne, J.; Bouvet, J; and Taberlet, P. (2001). Multiple maternal origins and weak phylogeographic structure in domestic goats. In : *Proceedings of the National Academy of Science of the USA* 98 : 5927 – 5932.

Lynch, M. (1988). Estimation of relatedness by DNA fingerprinting. *Molecular Biology and Evolution* 5: 584 – 599.

MacHugh, D.E.; Loftus, R.T.; Bradley, D.G.; Sharp, P.M. and Cunningham, E.P. (1994). Microsatellite DNA variation within and among European cattle breeds. In: *Proceedings of the Royal Society of London Series* B 256: 25 - 31.

MacHugh, D.E.; Loftus, R.T.; Cunningham, P. and Bradley, D.G. (1998). Genetic structure of seven European cattle breeds assessed using 20 microsatellite markers. *Animal Genetics* 29: 333 – 340.

MacHugh, D.E.; Shriver, M.D.; Loftus, R.T.; Cunningham, P. and Bradley, D.G. (1997). Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics* 146: 1071 – 1086.

Manly, B.F.J. (1986). *Multivariate Statistical Methods*: A Primer. Chapman and Hall. London.

Marshall, F. (2000). The origins and spread of domestic animals in East Africa. In: *The origins and development of African Livestock: archaeology, genetics, linguistics and ethnography* (Edited by Blench, R.M. and MacDonald, K.C). UCL Press. London, New York. pp 191 – 221.

Mason, I.L. (1951). The classification of West African livestock. Farnham royal, Slough, Bucks, UK. *Technical communication,Commonwealth Bureau of Animal Breeding and genetics* no. 7. pp 39.

Mason, I.L. (1981a). Wild goats and their domestication. In: *Goat production* (Edited by Gall C.). Academic press. London, New York, Toronto, San Francisco, Sydney. pp 35 – 55.

Mason, I.L. (1981b). Breeds. In: *Goat production* (Edited by Gall C.). Academic press. London, New York, Toronto, San Francisco, Sydney. pp 57- 110.

Mason, I.L. (1984). *Evolution of domesticated animals.* Longman. London and New York. pp 452.

Mason, I.L. and Maule, J.P. (1960). *The indigenous livestock of Eastern and Southern Africa.* Commonwealth Agricultural Bureaux, farnham Royal, Bucks, England. pp 179.

Meghen, C.; MacHugh, D.E. and Bradley, D.G. (1994). Genetic characterization and West African cattle. *World Animal Review* 78(1): 59-66.

Minch, E. (2000). MICROSAT, the microsatellite distance program. http://lotka.stanford.edu/microsat/microsat.htm

Minelli, A. (1993). *Biological systematics. The state of the art.* Capman & Hall. London, Glasgow, New York, Tokyo, Melbourne, Madras. pp 387.

Moore, H.D.M. (1985). Storage of gametes and embryos. In : *Advances in animal conservation. The proceedings of a symposium held at the Zoological Society of London* on 31 May – 1 June, 1984. No. 54 : 137 – 148.

Moore, S.S.; Sargeant, L.L.; King, T.J.; Mattick, J.S.; Georges, M. and Hetzel, D.J.S. (1991). The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics* 10 : 654 – 660.

Mopper, S.; Mitton, J.B.; Whitham, T.G.; Cobb, N.S. and Christensen, K.M. (1991). Genetic differentiation and heterozygosity in pinyon pine associated with resistance to herbivory and environmental stress. *Evolution* 45 : 989 – 999.

Mullis, K.B. and Faloona, F.A. (1987). Specific synthesis of DNA *in vitro* via a polymerase catalysed chain reaction. *Methods in Enzymology* 155: 335 - 350.

Nakumara, Y.; Lepper, M. and O'Connell, P. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* 235: 1616 – 1622.

Nei, M. (1972). Genetic distance between populations. *American Naturalist* 106 : 283 – 292.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. In: *Proceedings of the National Academy of Sciences of the U.S.A.* 70 : 3321 – 3323.

Nei, M. (1987). *Molecular Evolutionary Genetics.* Columbia University press, New York. pp 506.

Nei, M. (1996). Phylogenetic analysis in molecular evolutionary genetics. *Annual Review of Genetics* 30 : 371 – 403.

Nei, M. and Takezaki, N. (1994). Estimation of genetic distances and phylogenetic trees from DNA analysis. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production* , 7 – 12 August, 1994 Guelph, Canada 21 : 405 – 412.

Nei, M.; Tajima, F. and Tateno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution* 19 : 153 – 170.

Nelson, K. and Soulé, M. (1987). Genetic conservation of exploited fishes. In: *Population genetics and fishery management* (Edited by Nils Ryman and Fred Utter). Washington press. Seattle and London. pp 345 – 368.

Nguyen, T.C. (1990). Genetic systems of red cell blood groups in goats. *Animal Genetics* 21 : 233 – 245.

Notter, D.R.; Mariante, A. da S. and Sheng, Z. (1994). Modern approaches to active conservation of domestic animal diversity. In: *Proceedings of the 5th World Congress of Genetics Applied to Livestock Production.* 7 – 12 August, 1994 Guelph, Canada. Vol 21: 509 - 516.

Okomo, M.A.; Rege, J.E.O.; Teale, A. and Hanotte, O. (1998). Genetic characterisation of indigenous East African cattle breeds using microsatellite DNA markers. In: *Proceedings of the 6th World Congress of Genetics Applied to Livestock Production.* 11 – 16 January, 1998 Armidale, New Zealand. Vol 27: 243 - 246.

Ota, T. (1993). DISPAN: Genetic Distance and Phylogenetic Analysis. Pennsylvania State University, University Park PA, USA.

Paetkau, D.; Calvert, W.; Stirling, I. and Strobeck, C. (1995). Microsatellite analysis of population structure in polar bears. *Molecular Ecology* 4 : 347 – 354.

Page, R.D.M. (1996). Treeview: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 12 : 357 – 358.

Page, R.D.M. and Holmes, E.C. (1998). *Molecular evolution: A phylogenetic approach.* Blackwell science. University press, Cambridge. pp 346.

Payne, W.J.A. (1990). *An introduction to animal husbandry in the tropics.* Fourth edition. Longman Scientific & Technical, John Wiley & Sons, Inc., New York. pp 881.

Pemberton, J.M.; Albon, S.D.; Guinness, F.E.; Clutton-Brock, T.H. and Berry, R.J. (1988). Genetic variation and juvenile survival in red deer. *Evolution* 42 : 921 – 934.

Pepin, L. and Nguyen, T.C. (1994). Blood groups and protein polymorphisms in five goat breeds (*Capra hircus*). *Animal Genetics* 25 : 333 – 336.

Peters, K.J. (1987). Evaluation of goat populations in tropical and sub-tropical environments. *ILCA Bulletin* No. 28:14 - 21.

Primack, R.B. (1993). *Essentials of conservation biology.* Sinauer Associates, Inc., publishers. Sunderland, Massachusetts, U.S.A. pp 564.

Powell, W.; Morgante, M.; Andre, C.; Hanafey, M.; Vogel, J.; Tingey, S. and Rafalski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* 2 : 225 – 238.

Queller, D.C.; Strassmann, J.E. and Hughes, C.R. (1993). Microsatellites and kinship. *Trends in Ecology and Evolution* 8 : 285 – 288.

Ralls, K.; Ballou, J.D. and Templeton, A. (1988). Estimates of lethal equivalents and the costs of inbreeding in mammals. *Conservation Biology* 2: 185 – 193.

Raymond, M. and Rousset, F. (1995a). GENEPOP (Version 3.1): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86 : 248 – 249.

Raymond, M. and Rousset, F. (1995b). An exact test for population differentiation. *Evolution* 49 : 1280 – 1283.

Rege, J.E.O. (1992). Indigenous African small ruminants: a case for characterisation and improvement. In: *Proccedings of the second Biennial Conference of the African Small*

Ruminant Research Network, 7 – 11 December, 1992, AICC, Arusha, Tanzania. pp 205 – 211.

Rege, J.E.O.; Yapi-Gnaoré, C.V. and Tawah, C.L. (1996). The indigenous domestic ruminant genetic resources of Africa. In: *Proceedings of the second All Africa Conference on Animal Agriculture*, 1 – 4, April 1996, Pretoria, South Africa. pp. 57 – 75.

Reynolds, J.; Weir, B.S. and Cockerman, C.C. (1983). Estimation of the coancestry coefficient: bias for a short-term genetic distance. *Genetics* 105 : 767 – 779.

Richards, R.I. and Sutherland, G.R. (1994). Simple repeat DNA is not replicated simply. *Nature Genetics* 6: 114 – 116.

Rogers, J.S. (1972). Measures of genetic similarity and genetic distance. In: *Studies in Genetics VII.* University of Texas Publication 7213, Austin, TX. pp 145 – 153.

Roy, M.S.; Geffen, E.; Smith, D.; Ostrander, E.A. and Wayne, R.K. (1994). Patterns of differentiation and hybridisation in North American Wolflike canids, revealed by analysis of microsatellite loci. *Molecular Biology and Evolution* 11 : 553 – 570.

Rzhetsky, A. and Nei, M. (1992). Statistical properties of the ordinary least-squares, generalized least-square and minimum-evolution methods of phylogenetic inference. *Journal of Molecular Evolution* 35 : 367 – 375.

Saitbekova, N.; Gaillard, C.; Obexer-Ruff, G. and Dolf, G. (1999). Genetic diversity in Swiss goat breeds based on microsatellite analysis. *Animal Genetics* 30 : 36 – 41.

Saitou, N. and Imanishi, T. (1989). Relative efficiencies of the Fitch-Margoliash, Maximum-Parsimony, Maximum-Likelihood, Minimum-Evolution and Neighbour-joining methods of phylogenetic tree construction in obtaining the correct tree. *Molecular Biology and Evolution* 6 : 514 – 525.

Saitou, N. and Nei, M. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4 : 406 – 425.

Sambrook, J.; Fritsch, E.F. and Maniatis, T. (1989). *Molecular cloning – A laboratory manual. Vol. 2*, 2nd edition. Cold Spring Harbour, Laboratory Press, USA.

Schlötterer, C. and Tauz, D. (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research* 20: 211 – 215.

Selander, R.K. (1983). Evolutionary consequences of inbreeding. In: *Genetics and Conservation: A reference for managing wild animal and plant populations* (Edited by C.M. Schonewald-Cox, S.M. Chambers, B. MacBryde and L. Thomas). Benjamin/Cummings, menlopark, CA. pp 201 – 215.

Shriver, M.D.; Jin, L.; Boerwinkle, E.; Deka, R. and Ferrell, R.E. (1995). A novel measure of genetic distance for highly polymorphic tandem repeat loci. *Molecular Biology and Evolution* 12: 914 – 920.

Simm, G. (1998). *Genetic improvement of cattle and sheep*. Farming press. London. pp 433.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457 – 469.

Smith, A.B. (2000). The origins of the domesticated animals of southern Africa. In: *The origins and development of African Livestock: archaeology, genetics, linguistics and ethnography* (Edited by Blench, R.M. and MacDonald, K.C.). UCL press. London, New York. pp 222 – 231.

Smith, C. (1984). Rates of genetic change in farm livestock. *Research and Development in Agriculture* 1: 79 – 85.

Sneath, P.H.A. and Sokal, R.R. (1973). *Numerical Taxonomy.* W.H. Freeman and Co. San Francisco, California, USA. pp 573.

Soulé, M.E. (1980). Thresholds for survival: maintaining fitness and evolutionary potential. In: *Conservation biology* (Edited by M.E. Soulé and B.A. Wilcox). Sinauer, Sunderland, Massachusetts, USA. pp 151 – 168.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* 98: 503 - 517.

Swofford, D.L. and Olsen, G.J. (1990). Phylogeny reconstruction. In : *Molecular systematics* (Edited by D.M. Hills and C. Moritz). Sinauer, Sunderland, Massachusetts. pp 411 – 501.

Swofford, D.L.; Olsen, G.J.; Waddell, P.J. and Hills, D.M. (1996). Phylogenetic inference. In : *Molecular systematics* 2nd ed (Edited by D.M. Hills and C. Moritz). Sinauer, Sunderland, Massachusetts. pp 407 – 514.

Taberlet, P.; Camarra, J.J. and Griffin, S. (1997). Non-invasive genetic tracking of the endargered Pyrenean brown bear population. *Molecular Ecology* 6 : 869 – 876.

Takada, T.; Kikkawa, Y.; Yonekawa, H.; Kawakami, S. and Amano, T. (1997). Bezoar (*Capra aegagrus*) is a matriarchal candidate for ancestor of domestic goat (Capra hircus): evidence from the mitochondrial DNA diversity. *Biochemical Genetics* 35: 315 – 326.

Takezaki, N. and Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* 144 : 389 – 399.

Tateno, Y.; Nei, M. and Tajima, F. (1982). Accuracy of estimated phylogenetic trees from molecular data. I. Distantly related species. *Journal of Molecular Evolution* 18 : 387 - 404.

Tauz, D. (1989). Hypervariability of simple sequences as a general source of polymorphic DNA markers. *Nucleic Acids Research* 17: 6463 – 6471.

Taylor, A.C.; Sherwin, W.B. and Wayne, R.K. (1994). Genetic variation of microsatellite loci in a bottlenecked species: the hairy-nosed wombat (*Lasiorhinus krefftii*). *Molecular Ecology* 3 : 277 – 290.

Templeton, A.R. (1986). Coadaptation and outbreeding depression. In: *Conservation Biology. The science of scarcity and diversity* (Edited by M.E., Soule). Sinauer. Sunderland, Massachusetts. pp 105 – 116.

Valdes, A.M.; Slatkin, M. and Freimer, N.B. (1993). Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics* 133 : 737 – 749.

Vaiman, D.; Schibler, L. and Bourgeois, F. (1996). A genetic linkage map of the male goat genome. *Genetics* 144 : 279 – 283.

Vos, P.; Hogers, R.; Bleeker, M.; Reijans, M.; van de Lee, T.; Hornes, M.; Frijters, A.; Pot, J.; Peleman, J.; Kuiper, M. and Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23 : 4407 – 4414.

Wall, W.J.; Williamson, R.; Petrou, M.; Papaioannou, D. and Parkin, B.H. (1993). Variation of short tandem repeats within and between populations. *Human Molecular Genetics* 2(7) : 1023 - 1029.

Wayne, R.K.; Gilbert, D. and Lehman, N. (1991). Conservation genetics of the endangered Isle Royale gray wolf. *Conservation Biology* 5: 41 – 51.

Weber, J.L. (1990). Informativeness of human (dC-dA)n (dG-dT)n polymorphisms. *Genomics* 7 : 524 – 530.

Weber, J.L. and May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using polymerase chain reaction. *American Journal of Human Genetics* 44:388 - 396.

Weir, B.S. (1990). *Genetic Data Analysis – Methods for discrete population genetic data*. Sinauer Associates, Inc. publishers. Sunderland, Massachusetts. pp 377.

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

Weissenbach, J.; Gyapay, G.; Dib, C.; Vignal, A.; Morisette, J.; Millaseau, P.; Vaysseix, G. and Lathrop, M. (1992). A second generation linkage map of the human genome. *Nature* 359 : 794 – 801.

Welsh, J. and McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18 : 7213 – 7218.

Williams, J.G.; Kubelik, A.R.; Livak, K.J.C.; Rafalski, J.A. and Tingey, M.S.V. (1990). DNA polymorphysms amplified by arbitrary primers are useful as genetic makers. *Nucleic Acids Research* 18 (22): 6531 - 6535.

Wimmers, K.; Ponsuksili, S.; Hardge, T.; Valle-Zarate, A.; Mathur, P.K. and Horst, P. (2000). Genetic distinctness of African, Asian and South American local chickens. *Animal Genetics* 31: 159 – 165.

Winrock International (1983). *Sheep and goats in developing countries. Their present and potential roles*. The World Bank, Washington, D.C. U.S.A. pp 116.

Yang, L.; Zhao, S.H.; Li, K.; Peng, Z.Z. and Montgomery, G.W. (1999). Determination of genetic relationships among five indigenous Chinese goat breeds with six microsatellite markers. *Animal Genetics* 30 : 452 – 455.

Zeder, M.A. and Hesse, B. (2000). The initial domestication of goats (*Capra hircus*) in the Zagros Mountains 10,000 years ago. *Science* 287 : 2254 – 2257.

APPENDICES

Appendix 1: Number of alleles per breed for each locus in sub-Saharan African goats

Breed	Locus																		
	ILSTS 44	MAF 209	INRA 5	MAF 35	INRA 63	OarFCB 304	SRCR-SP3	BM 1818	OarAE 129	SRCR-SP7	ILSTS 87	ILSTS 17	BMS 1494	INRA 132	BMC 1222	BMS 357	MAF 65	ILSTS 11	ILSTS 5
Maasai	6	6	5	3	7	7	5	8	7	4	6	6	5	3	6	6	8	6	7
Kigezi	6	6	5	3	3	7	6	8	9	4	5	7	9	3	4	8	8	6	7
Mubende	5	5	7	4	7	4	6	6	8	6	12	7	10	4	7	9	9	10	8
Ndebele	6	4	6	4	7	8	6	8	8	6	8	7	8	7	8	9	8	6	6
WAD	4	5	6	3	5	8	7	11	10	5	3	7	7	5	6	5	6	8	5
NWH	4	4	5	2	5	9	5	5	8	4	7	7	7	3	5	8	11	9	2
Pafuri	4	3	6	5	6	8	4	9	5	3	5	5	6	3	8	4	9	11	5
Maure	6	3	4	3	5	10	6	10	7	5	6	8	9	3	4	7	8	7	4
Arsi-Bale	5	3	4	2	5	11	5	9	7	4	4	7	8	3	6	8	12	6	4
Djallonke	4	4	5	2	4	10	5	7	4	5	5	6	8	3	4	7	8	6	3
Grisons Striped	2	2	4	3	4	7	6	10	8	5	4	4	10	2	5	7	10	7	6
Toggenburg	1	1	3	3	3	7	3	6	4	3	3	4	6	2	4	6	7	7	4
Mongolian Cashmere	4	3	3	3	5	7	6	9	9	5	8	8	5	3	6	14	10	11	5
Bandipur	12	4	6	3	5	10	6	8	9	4	11	7	8	7	11	12	8	7	4
Arab Emirates	6	3	4	2	3	5	6	6	9	4	8	6	5	2	7	11	7	8	3

Appendix 2: Tests for deviation from HWE in sub-Saharan African goat breeds

Breed	Locus																		
	ILSTS 44	INRA 5	MAF 209	MAF 35	INRA 63	OarFCB 304	SRCRSP 3	BM 1818	OarAE 129	SRCRSP 7	ILSTS 87	ILSTS 17	BMS 1494	INRA 132	BMC 1222	BMS 357	MA F65	ILSTS 11	ILSTS 5
Maasai	*	ns	*	ns	***	*	**	ns	ns	***	ns	***	***	***	ns	***	***	ns	*
Kigezi	***	***	*	**	***	***	***	**	***	**	**	***	***	ns	***	***	*	***	***
Mubende	ns	***	***	***	ns	*	***	***	***	***	***	***	***	*	***	***	***	***	***
Ndebele	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	*	*
WAD	***	***	ns	***	***	***	***	***	***	***	***	***	***	**	***	***	ns	***	**
NWH	ns	***	ns	ns	**	***	ns	ns	***	***	ns	***	ns	ns	***	***	**	***	ns
Pafuri	***	**	ns	***	*	***	ns	***	ns	***	*	***	ns	ns	**	***	ns	***	ns
Maure	*	*	**	***	ns	***	***	***	ns	ns	ns	***	ns	ns	***	***	***	**	ns
Arsi-Bale	ns	ns	***	ns	ns	***	ns	*	**	**	ns	***	ns	ns	ns	***	***	ns	*
Djallonke	*	*	***	*	ns	***	ns	***	***	***	*	***	***	ns	***	***	*	ns	*
Arab	*	*	***	ns	*	**	***	*	ns	*	**	***	*	ns	**	**	ns	*	ns
Emirates Grisons	-	ns	*	***	*	ns	ns	***	***	**	***	ns	*	ns	*	ns	ns	ns	ns
Striped Toggenburg	-	-	ns	ns	**	ns	ns	ns	ns	ns	ns	***	ns	ns	ns	ns	**	***	***
Mongolian Cashmere	*	ns	ns	ns	ns	ns	ns	**	ns	***	ns	***	*	ns	***	**	ns	***	ns
Bandipur	ns	ns	**	**	*	***	ns	ns	*	ns	ns	***	ns	***	***	**	ns	***	ns

ns – not significant P ≥ 0.05
* - Significant P ≤ 0.05
** - Significant P ≤ 0.01
*** - Significant P ≤ 0.001

Appendix 3: Estimates of f for each locus and breed in sub-Saharan African goats

Breed	ILSTS 44	MAF 209	INRA 5	MAF 35	INRA 63	OARFCB 304	SRCRSP 3	BM 1818	OARAE 129	SRCRS P7	ILSTS 87	ILSTS 17	BMS 1494	INRA 132	BMC 1222	BMS 357	MAF 65	ILSTS 11	ILSTS 5
Masai	0.229	0.223	0.163	-0.028	0.217	0.323	0.343	-0.073	0.173	0.445	-0.190	0.540	0.263	0.698	0.237	0.738	0.504	0.006	0.250
Kigezi	0.331	0.096	0.402	0.428	0.646	0.434	0.381	0.277	0.334	0.530	0.370	0.558	0.314	0.343	0.471	0.746	0.273	0.562	0.663
Mubende	0.065	0.389	0.420	0.888	0.071	0.318	0.357	0.596	0.251	0.626	0.584	0.563	0.322	0.375	0.680	0.624	0.492	0.504	0.687
Emirates	0.415	0.810	0.310	0.216	0.318	0.446	0.441	0.207	0.234	0.397	0.229	0.857	-0.008	-0.016	0.397	0.419	0.109	0.281	-0.144
Ndebele	0.580	0.628	0.524	0.572	0.478	0.549	0.574	0.639	0.406	0.486	0.432	0.843	0.389	0.338	0.391	0.885	0.547	0.290	-0.045
WAD	0.839	0.139	0.263	1.000	0.339	0.873	0.353	0.336	0.311	0.392	0.534	0.620	0.059	0.409	0.696	0.839	0.203	-0.027	0.514
NWH	0.179	0.161	0.436	0.141	0.292	0.436	0.124	0.148	0.316	0.697	0.046	0.807	-0.054	-0.064	0.548	0.700	0.306	0.302	0.284
Pafuri	0.834	0.007	0.276	0.450	0.037	0.638	0.192	0.590	-0.042	0.719	0.318	0.644	-0.108	0.202	0.276	0.774	0.096	0.275	-0.094
Maure	0.417	0.548	-0.068	0.755	-0.119	0.786	0.434	0.250	0.158	-0.007	-0.120	0.747	-0.117	0.125	0.811	0.716	0.355	0.119	-0.029
Asibale	-0.069	0.492	-0.124	0.064	0.104	0.480	0.148	0.125	0.353	0.436	-0.071	0.545	-0.024	0.072	0.131	0.765	0.447	-0.008	0.284
Djallonke	0.435	0.490	-0.118	1.000	0.181	0.497	-0.066	0.314	0.588	0.515	-0.313	0.555	-0.176	-0.022	0.664	0.601	0.154	0.164	0.244
Grisons	NA	1.000	-0.054	-0.673	0.372	0.167	0.190	0.372	0.506	0.286	-0.381	0.300	-0.045	-0.043	0.215	-0.160	0.104	0.055	0.065
Toggenburg	NA	NA	-0.101	-0.053	0.428	-0.169	0.473	-0.011	0.055	0.272	0.138	0.628	-0.027	0.077	0.282	-0.132	0.393	-0.212	0.704
Mongolian	0.519	-0.110	0.066	0.099	0.134	0.118	0.017	0.025	-0.028	0.598	0.067	0.257	-0.051	-0.118	0.552	0.152	-0.012	0.327	0.115
Bandipur	-0.028	0.482	0.192	0.404	0.190	0.728	0.051	-0.125	0.123	0.318	0.093	0.454	-0.130	0.321	0.460	0.509	-0.133	0.563	0.122

NA – not available.

Appendix 4: Tests for genetic differentiation between pairs of breeds of sub-Saharan African goats

Population pair	ILSTS 44	MAF 209	INRA 5	MAF 35	INRA 63	OarFCB 304	SRCR-SP3	BM 1818	OarAE 129	SRCR-SP7	ILSTS 87	ILSTS 17	BMS 1494	INRA 132	BMC 1222	BMS3 57	MAF 65	ILSTS11	ILSTS 5
Maasai - Kigezi	**	ns	***	**	***	***	**	***	***	***	***	***	**	ns	ns	***	***	***	ns
Maasai - Mubende	ns	***	***	***	***	***	***	**	**	***	***	**	***	**	*	***	***	***	***
Maasai - Emirates	***	***	***	ns	***	**	***	***	***	*	***	***	***	***	***	***	***	***	**
Maasai - Ndebele	***	**	**	***	*	***	***	***	***	***	***	*	***	***	***	***	***	***	***
Maasai - WAD	***	**	*	***	*	***	**	***	***	***	ns	***	***	ns	***	***	***	***	ns
Maasai - NWH	ns	***	**	ns	**	***	ns	**	***	*	ns	***	***	*	***	***	***	***	***
Maasai - Pafuri	***	***	***	*	**	**	***	***	***	***	*	**	***	*	***	***	***	ns	***
Maasai - Maure	***	***	ns	***	*	***	***	*	***	***	**	*	***	ns	***	***	***	ns	**
Maasai - Arsi-Bale	*	***	***	*	***	***	ns	*	***	**	***	***	**	ns	***	***	***	*	**
Maasai - Djallonke	***	*	***	ns	***	***	***	***	***	***	ns	***	***	ns	*	***	***	***	**
Maasai - Grisons	***	***	ns	***	***	***	**	***	***	***	***	***	***	**	***	***	***	***	***
Maasai - Toggenburg	***	***	ns	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Maasai - Mongolian	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Maasai - Bandipur	***	***	***	***	ns	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Kigezi - Mubende	**	***	*	***	ns	***	***	ns	***	***	***	***	***	ns	***	**	***	**	*
Kigezi - Emirates	***	***	ns	*	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Kigezi - Ndebele	**	ns	ns	ns	***	***	***	*	***	**	***	***	***	***	***	***	***	***	**
Kigezi - WAD	***	ns	***	***	***	***	***	***	***	***	***	***	***	**	***	***	***	***	**
Kigezi - NWH	***	***	ns	*	***	**	***	***	***	**	***	***	***	ns	***	***	***	***	***
Kigezi - Pafuri	***	***	ns	**	***	***	***	*	***	***	***	***	***	ns	**	***	***	*	*
Kigezi - Maure	**	***	***	***	***	***	***	ns	***	***	***	***	***	*	***	*	***	***	***
Kigezi - Arsi-Bale	**	***	*	*	*	**	***	***	***	***	***	***	***	ns	***	***	***	***	***
Kigezi - Djallonke	***	ns	***	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Kigezi - Grisons	***	***	***	***	*	***	***	***	***	***	***	***	ns	ns	***	***	***	***	***
Kigezi - Toggenburg	***	***	***	***	***	***	***	***	***	***	***	***	*	**	***	***	***	***	***
Kigezi - Mongolian	***	***	**	***	***	***	***	***	***	***	***	***	**	***	***	***	***	***	**
Kigezi - Bandipur	***	***	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Mubende - Emirates	***	***	*	***	***	**	ns	***	***	**	***	***	***	***	***	***	***	***	**
Mubende - Ndebele	***	***	ns	**	***	***	***	***	***	***	*	**	***	***	***	***	***	**	***
Mubende - WAD	***	***	***	***	***	***	ns	***	***	*	***	***	***	***	***	***	***	***	**
Mubende - NWH	*	***	**	***	ns	***	**	***	***	**	***	***	***	ns	***	***	***	***	***
Mubende - Pafuri	***	***	***	*	**	**	***	**	***	*	***	***	***	*	**	***	***	ns	*

Population pair	ILST S44	MAF 209	INRA 5	MAF 35	INRA 63	OarFC B304	SRC RSP3	BM18 18	OarA E129	SRC RSP7	ILST S87	ILST S17	BMS 1494	INRA 132	BMC 1222	BMS 357	MAF 65	ILST S11	ILST S5
Mubende - Maure	ns	***	***	***	***	***	**	*	***	***	***	***	***	***	***	ns	***	***	***
Mubende - Arsi-Bale	ns	***	**	***	ns	***	***	***	***	*	***	***	***	ns	***	***	***	***	***
Mubende - Djallonke	***	***	***	**	***	***	ns	***	***	*	***	***	***	***	***	***	***	***	***
Mubende - Grisons	***	***	***	***	ns	***	***	***	***	***	***	***	***	**	***	***	***	***	***
Mubende - Toggenburg	***	***	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Mubende - Mongolian	***	***	***	***	***	***	*	***	***	***	***	***	***	***	***	***	***	ns	**
Mubende - Bandipur	***	***	**	***	***	***	***	***	***	***	***	***	***	**	***	***	***	***	**
Emirates - Ndebele	***	***	ns	*	***	***	***	***	***	***	***	ns	***	***	***	***	***	***	***
Emirates - WAD	***	***	***	***	***	***	*	***	***	**	***	***	***	***	***	***	***	***	*
Emirates - NWH	***	**	**	ns	*	***	*	***	***	ns	***	**	***	***	***	***	***	***	***
Emirates - Pafuri	***	***	*	ns	ns	ns	***	***	***	***	***	***	**	***	***	***	***	***	***
Emirates - Maure	***	***	***	***	**	***	***	***	***	*	***	ns	***	***	***	***	***	***	***
Emirates - Arsi-Bale	***	**	**	ns	*	***	***	**	***	**	***	***	***	***	***	***	***	***	***
Emirates - Djallonke	***	***	***	ns	ns	***	*	***	***	***	***	**	***	***	***	***	***	***	ns
Emirates - Grisons	***	***	***	***	ns	***	***	***	***	ns	***	***	***	***	***	***	***	***	***
Emirates - Toggenburg	***	***	***	***	***	***	***	***	***	***	***	**	***	***	***	***	***	***	***
Emirates - Mongolian	***	ns	***	***	***	***	**	***	***	***	***	***	***	***	***	***	***	***	***
Emirates - Bandipur	***	***	***	***	***	***	***	***	***	***	***	***	**	***	***	***	***	***	*
Ndebele - WAD	***	ns	***	***	***	***	***	***	***	**	***	***	***	***	***	***	***	***	***
Ndebele - NWH	***	*	ns	**	***	***	***	***	***	***	***	ns	***	***	***	***	***	***	***
Ndebele - Pafuri	*	**	***	ns	***	ns	ns	***	***	***	***	*	***	***	**	***	***	**	**
Ndebele - Maure	***	***	***	***	***	***	***	**	***	**	***	ns	***	***	***	***	***	***	***
Ndebele - Arsi-Bale	***	*	*	*	***	***	***	***	***	***	***	ns	***	***	***	***	***	***	***
Ndebele - Djallonke	***	ns	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Ndebele - Grisons	***	***	***	***	***	***	**	***	***	***	***	***	***	***	***	***	***	***	***
Ndebele - Toggenburg	**	***	**	***	***	***	ns	***	***	***	***	***	***	***	***	***	*	***	**
Ndebele - Mongolian	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	*	*
Ndebele - Bandipur	***	***	*	***	*	***	***	***	***	***	***	*	***	***	***	***	***	***	***
WAD - NWH	***	ns	***	***	***	***	***	***	***	**	***	***	***	***	**	***	***	***	*
WAD - Pafuri	***	*	***	***	**	***	***	***	***	ns	***	***	***	***	***	***	***	***	***
WAD - Maure	**	***	ns	***	ns	***	ns	***	ns	ns	***	**	***	ns	*	***	ns	***	ns
WAD - Arsi-Bale	***	*	***	***	***	***	*	***	***	ns	***	***	***	**	**	***	***	***	*
WAD - Djallonke	ns	ns	**	***	***	***	*	***	***	***	ns	***	***	ns	*	***	**	***	*
WAD – Grisons	***	***	**	***	***	***	***	***	***	**	***	***	***	***	***	***	*	***	**

Population pair	ILST S44	MAF 209	INRA 5	MAF 35	INRA 63	OarFC B304	SRC RSP3	BM18 18	OarA E129	SRC RSP7	ILST S87	ILST S17	BMS 1494	INRA 132	BMC 1222	BMS 357	MAF 65	ILST S11	ILST S5
WAD - Toggenburg	***	***	**	***	***	***	***	***	***	**	***	***	**	**	***	***	***	***	***
WAD - Mongolian	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
WAD - Bandipur	***	***	***	***	*	***	***	***	***	***	***	ns	**	***	***	***	***	***	ns
NWH - Pafuri	***	*	**	ns	ns	***	***	***	***	**	**	***	ns	ns	***	***	***	***	***
NWH - Maure	***	***	***	***	ns	***	***	***	***	***	**	***	***	**	ns	***	***	***	ns
NWH - Arsi-Bale	**	ns	ns	ns	ns	***	ns	ns	**	ns	***	ns	ns	ns	ns	***	***	***	*
NWH - Djallonke	***	*	***	ns	***	***	**	***	***	***	**	***	***	***	*	***	***	***	***
NWH - Grisons	***	***	***	***	ns	***	***	***	***	***	***	***	***	ns	***	***	*	***	**
NWH - Toggenburg	***	***	*	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
NWH - Mongolian	***	***	**	***	***	***	***	***	***	***	***	***	ns	***	***	***	*	***	***
NWH - Bandipur	***	***	ns	***	**	***	***	***	***	***	***	***	**	***	***	***	***	***	***
Pafuri - Maure	***	ns	***	***	ns	***	***	***	***	*	***	**	***	**	***	***	***	ns	***
Pafuri - Arsi-Bale	**	ns	*	ns	ns	***	***	***	***	**	***	***	ns	ns	***	***	***	***	***
Pafuri - Djallonke	***	***	***	*	*	***	***	***	***	**	**	***	***	***	***	***	***	***	***
Pafuri - Grisons	***	***	***	***	ns	***	**	***	***	***	***	ns	***	ns	***	***	***	***	***
Pafuri - Toggenburg	***	**	***	***	***	***	ns	***	***	***	***	***	***	**	***	***	***	***	***
Pafuri - Mongolian	***	*	ns	***	***	***	*	***	***	***	***	***	*	***	***	***	***	*	ns
Pafuri - Bandipur	***	**	**	***	**	***	***	***	***	***	***	***	ns	***	***	***	***	***	*
Maure - Arsi-Bale	ns	***	***	***	*	***	***	**	***	***	***	***	***	*	**	***	***	***	ns
Maure - Djallonke	***	***	***	***	**	***	ns	***	***	***	***	***	***	ns	ns	***	ns	***	***
Maure - Grisons	***	ns	**	***	*	***	***	***	***	*	***	***	***	***	***	***	*	***	*
Maure - Toggenburg	***	ns	ns	***	***	***	***	***	***	***	***	***	**	**	***	***	***	***	***
Maure - Mongolian	***	**	***	***	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Maure - Bandipur	***	ns	***	*	ns	***	***	***	***	***	***	ns	***	***	***	***	***	***	*
Arsi-Bale - Djallonke	***	**	***	*	**	***	***	***	***	***	***	***	***	***	**	***	***	***	***
Arsi-Bale - Grisons	***	***	***	***	ns	***	***	***	***	***	***	***	***	ns	***	***	***	***	ns
Arsi-Bale - Toggenburg	***	***	***	***	***	***	***	***	***	***	***	***	***	**	***	***	*	***	***
Arsi-Bale - Mongolian	***	**	**	***	***	***	***	***	**	***	***	***	ns	***	***	***	**	***	***
Arsi-Bale - Bandipur	***	***	ns	***	***	***	***	***	***	***	***	***	ns	***	***	***	***	***	***
Djallonke - Grisons	***	***	***	***	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Djallonke - Mongolian	***	***	***	***	***	***	***	***	***	***	*	***	***	***	***	***	***	***	***
Djallonke - Mongolian	***	***	***	***	***	***	**	***	***	***	***	***	***	***	***	***	***	***	***
Djallonke - Bandipur	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	**
Grisons - Toggenburg	ns	ns	ns	***	***	***	ns	**	***	**	ns	***	*	***	**	**	***	*	***

Population pair	ILSTS 44	MAF 209	INRA 5	MAF 35	INRA 63	OarFCB 304	SRCRSP 3	BM 1818	OarAE 129	SRCRSP 7	ILSTS 87	ILSTS 17	BMS 1494	INRA 132	BMC 1222	BMS 357	MAF 65	ILSTS 11	ILSTS 5
Grisons - Mongolian	**	***	***	ns	***	***	***	***	***	***	***	***	*	***	***	***	**	**	***
Grisons - Bandipur	***	ns	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Toggenburg - Mongolian	ns	***	***	***	**	***	**	***	***	***	***	***	***	***	**	***	***	***	***
Toggenburg - Bandipur	**	ns	***	***	***	***	*	***	***	***	***	***	***	***	***	***	***	***	***
Mongolian - Bandipur	***	***	***	***	***	***	***	***	***	ns	**	***	**	***	**	***	***	***	**

Appendix 5: Number of alleles per population for each locus in eastern African goat populations

Population	Locus																		
	INRA 63	ILSTS8 7	SRCR-SP3	SRCR-SP7	OarAE 129	BMS 357	INRA 5	BMS 1494	BMC 1222	INRA 132	ILSTS17	BM 1818	MAF 209	OarFCB 304	MAF 35	ILSTS4 4	ILSTS5	MAF 65	ILSTS 11
KSEA	4	6	5	4	6	9	5	4	4	3	8	9	3	11	3	4	5	9	5
Boran	4	3	5	3	8	6	4	3	4	3	7	10	6	13	2	5	3	11	5
Galla	5	5	4	3	7	5	5	3	4	3	6	8	3	12	2	5	7	10	8
NEH	5	5	5	5	9	10	4	4	6	4	7	6	5	10	2	6	5	10	6
Afar	3	8	5	4	10	11	5	5	5	4	7	9	3	11	3	6	6	11	8
Tanzanian Coastal	5	8	5	4	9	7	5	4	7	5	6	9	5	8	3	6	8	9	6
Ugogo	5	6	4	3	9	10	6	4	5	6	7	9	3	8	2	5	5	9	7
Ujiji	5	9	4	4	9	10	6	4	8	4	9	7	3	6	4	4	5	9	7
Maasai	5	8	4	3	9	9	5	4	5	5	6	8	6	7	2	6	7	8	6
Sukuma	5	6	4	3	10	10	6	4	6	4	6	12	3	9	2	5	5	11	8
Newala	5	6	4	4	6	7	4	4	5	2	5	8	4	13	5	5	6	8	8
Mbeya	5	7	6	4	7	7	5	4	5	6	8	7	3	12	5	4	8	9	6
Landim	4	8	5	4	10	7	5	6	7	4	4	8	3	13	2	7	4	8	9

Appendix 6: Tests for deviation from HWE in eastern African goat populations

Breed	Locus																		
	ILSTS 44	INRA 5	MAF 209	MAF 35	INRA 63	OarFCB 304	SRCRSP 3	BM 1818	OarAE 129	SRCRSP 7	ILSTS 87	ILSTS 17	BMS 1494	INRA 132	BMC 1222	BMS 357	MAF 65	ILSTS 11	ILSTS 5
KSEA	ns	ns	**	ns	ns	*	ns	*	ns	*	ns	**	ns	***	ns	***	***	ns	ns
Boran	***	ns	**	ns	***	ns	***	ns	***	ns	***	ns	**	***	***	**	ns	***	ns
Galla	ns	ns	ns	ns	***	*	ns	ns	ns	ns	*	ns	ns	ns	ns	**	ns	ns	***
NEH	**	ns	ns	ns	*	ns	**	*	ns	ns	***	*	**	*	*	***	ns	**	ns
Afar	ns	ns	ns	*	ns	ns	ns	**	ns	**	***	**	*	***	ns	*	*	ns	***
Tanzanian Coastal	*	*	ns	ns	ns	*	ns	***	*	ns	***	***	**	*	*	***	ns	*	***
Ugogo	ns	*	ns	ns	**	ns	*	*	***	**	***	***	ns	***	*	***	***	***	***
Ujiji	ns	*	ns	**	**	*	*	ns	ns	*	***	***	ns	**	**	***	ns	*	ns
Maasai	*	ns	**	ns	*	ns	**	ns	ns	*	**	***	*	ns	ns	***	***	ns	*
Sukuma	**	ns	ns	ns	ns	ns	ns	***	ns	*	***	*	*	ns	**	***	**	***	ns
Newala	***	ns	ns	*	**	ns	ns	*	***	*	***	**	*	ns	***	**	ns	**	ns
Mbeya	ns	***	ns	***	**	ns	**	ns	***	***	***	***	ns	***	***	***	ns	***	*
Landim	**	ns	ns	ns	*	ns	ns	**	***	****	***	ns	*	ns	**	***	ns	*	ns

Appendix 7: Estimates of *f* for each locus per breed for eastern African goat populations

Population	INRA 63	ILSTS 87	SRCR-SP3	SRCR-SP7	OARAE 129	BMS 357	INRA 5	BMS 1494	BMC 1222	INRA 132	ILSTS 17	BM 1818	MAF 209	OARFC B304	MAF 35	ILSTS 44	ILSTS 5	MAF 65	ILSTS 11
Tswana	0.232	-0.100	-0.258	0.117	0.192	0.347	0.033	0.307	0.127	0.510	0.456	-0.023	0.258	0.195	0.441	0.579	-0.038	0.017	0.243
Venda	0.535	0.296	0.261	0.048	0.013	0.796	-0.004	0.105	0.280	0.004	0.444	-0.153	0.344	0.090	-0.037	0.580	-0.238	0.097	0.233
KSEA	0.023	-0.224	-0.186	-0.020	0.066	0.377	0.093	0.230	0.084	-0.674	0.438	0.116	0.147	0.165	-0.113	0.138	-0.057	0.542	0.205
Toggenburg	0.210	-0.023	-0.245	-0.044	0.040	0.672	-0.225	0.037	0.079	-0.241	0.645	0.150	NA	0.078	-0.022	1.000	0.623	0.255	-0.025
Landim	0.097	-0.107	0.158	0.284	0.130	0.457	-0.111	0.060	0.500	-0.293	-0.197	0.356	0.379	0.120	-0.014	0.395	0.120	0.167	0.222
Boran	0.201	-0.777	0.140	0.143	0.219	0.085	-0.179	0.317	0.419	-0.320	0.215	0.058	0.131	0.029	-0.101	0.404	0.290	0.159	0.316
Galla	0.534	-0.351	-0.163	0.100	0.101	0.029	-0.008	-0.377	0.033	-0.192	0.026	-0.036	0.097	0.316	0.151	-0.013	0.757	0.106	0.137
NEH	0.163	-0.034	-0.079	-0.003	0.079	0.334	-0.128	0.313	0.170	0.106	0.246	0.217	-0.047	-0.153	0.082	0.168	0.126	-0.009	0.111
Afar	-0.277	0.072	0.022	-0.021	0.068	0.166	0.032	0.103	-0.013	-0.438	0.358	0.300	-0.078	0.153	0.246	0.127	0.439	0.030	0.148
Tanzanian Coastal	0.072	0.364	0.246	0.219	0.022	0.270	0.230	0.357	0.179	0.156	0.638	0.349	0.136	0.224	0.090	0.256	0.632	0.015	0.061
Ugogo	0.124	0.360	-0.123	0.447	0.211	0.273	0.095	0.115	0.165	-0.182	0.474	0.100	0.158	0.069	-0.033	0.150	0.661	0.197	0.455
Ujiji	0.113	0.701	0.134	0.359	-0.092	0.236	0.048	0.000	0.264	0.034	0.491	0.172	-0.103	0.208	0.235	0.054	0.009	-0.097	0.108
Maasai	0.038	0.314	0.236	0.318	0.001	0.524	0.094	-0.084	0.066	0.035	0.465	0.071	0.148	-0.002	-0.011	0.248	0.222	0.504	0.006
Sukuma	0.036	0.507	-0.018	0.335	0.009	0.243	0.071	0.194	0.286	0.033	0.394	0.266	-0.045	-0.096	-0.033	0.229	-0.036	0.211	0.195
Newala	0.259	-0.029	-0.049	0.064	0.397	0.102	0.175	0.051	0.379	-0.032	0.384	0.155	0.203	0.098	0.107	0.776	-0.105	-0.015	0.146
Mbeya	0.161	0.038	0.124	0.410	-0.024	0.037	0.282	0.012	0.421	0.253	0.634	0.109	-0.104	-0.049	0.407	0.150	-0.024	0.044	0.332
RedSokoto	-0.061	-0.311	-0.294	0.237	0.053	0.006	0.128	0.063	0.183	0.049	-0.010	-0.012	0.123	0.005	-0.039	0.692	-0.048	0.133	0.452

NA – not available

Appendix 8: Tests for genetic differentiation between pairs of populations of eastern African goats

Population pair		INRA 63	ILSTS 87	SRCR SP3	SRCR SP7	OarAE 129	BMS 357	INRA 5	BMS 1494	BMC 1222	INRA 132	ILSTS 17	BM 1818	MAF 209	OarFCB 304	MAF 35	ILSTS 44	ILSTS 5	MAF 65	ILSTS 11
KSEA	- Landim	***	***	ns	***	**	***	ns	*	***	**	*	*	***	**	ns	***	***	***	ns
KSEA	- Boran	***	**	ns	ns	***	***	ns	ns	ns	ns	ns	ns	***	*	ns	*	*	**	***
KSEA	- Galla	**	ns	ns	ns	***	**	ns	ns	*	**	ns	ns	**	ns	ns	***	*	*	*
KSEA	- NEH	ns	**	ns	ns	***	***	ns	***	**	ns	***	***	***	***	ns	*	**	***	**
KSEA	- Afar	***	***	*	*	**	***	***	***	ns	ns	***	**	***	***	ns	ns	ns	***	***
KSEA	- Ta-Co	***	***	***	***	**	***	ns	*	***	*	***	***	ns	***	***	ns	*	***	ns
KSEA	- Ugogo	***	***	*	ns	ns	***	***	ns	ns	*	*	**	*	***	***	ns	*	***	***
KSEA	- Ujiji	***	***	ns	ns	ns	***	ns	*	***	***	ns	**	ns	***	***	ns	**	***	***
KSEA	- Maasai	***	***	ns	*	*	***	ns	ns	ns	ns	**	***	ns	***	***	ns	ns	***	ns
KSEA	- Sukuma	***	***	ns	ns	***	***	***	ns	ns	ns	**	***	**	***	***	ns	ns	ns	***
KSEA	- Newala	ns	***	***	***	***	***	ns	*	***	ns	***	***	*	***	***	*	***	***	*
KSEA	- Mbeya	***	**	ns	*	ns	***	***	ns	ns	ns	***	***	ns	***	***	**	ns	**	ns
Landim	- Boran	ns	***	ns	***	***	***	*	***	***	*	ns	ns	***	**	ns	***	***	***	***
Landim	- Galla	ns	**	ns	**	***	***	ns	ns	***	***	ns	ns	*	*	**	ns	***	***	**
Landim	- NEH	***	***	*	**	*	***	*	***	***	*	***	***	***	***	*	***	***	***	***
Landim	- Afar	**	ns	ns	**	**	***	***	***	***	ns	***	***	**	***	*	***	***	***	***
Landim	- Ta-Co	*	***	*	ns	***	**	ns	**	***	*	***	***	***	***	***	***	***	***	***
Landim	- Ugogo	***	*	*	***	*	***	**	ns	***	**	*	***	***	***	***	***	***	***	***
Landim	- Ujiji	***	***	**	***	***	***	*	ns	***	***	***	***	***	***	***	***	***	***	***
Landim	- Maasai	**	**	*	***	**	***	ns	*	***	ns	*	***	***	***	***	***	***	***	*
Landim	- Sukuma	*	***	ns	***	***	***	**	*	***	ns	ns	ns	***	***	***	***	***	***	***
Landim	- Newala	***	**	***	***	***	***	ns	***	***	**	ns	***	***	***	***	***	***	**	***
Landim	- Mbeya	***	*	ns	ns	**	**	***	*	***	*	***	ns	***	***	***	***	***	***	***

Ta-Co – Tanzanian Coastal

Population pair		INRA 63	ILSTS 87	SRCR SP3	SRCR SP7	OarAE 129	BMS 357	INRA 5	BMS 1494	BMC 1222	INRA 132	ILSTS 17	BM 1818	MAF 209	OarFCB 304	MAF 35	ILSTS 44	ILSTS 5	MAF 65	ILSTS 11
Boran	- Galla	**	*	ns	ns	ns	***	ns	**	ns	**	ns	ns	ns	ns	ns	ns	*	*	*
Boran	- NEH	***	***	ns	ns	ns	***	ns	***	***	ns	ns	***	*	**	ns	***	***	***	***
Boran	- Afar	***	***	ns	*	**	***	*	***	*	ns	***	***	ns	*	ns	**	***	****	***
Boran	- Ta-Coa	*	***	***	***	***	***	*	***	****	**	***	***	*	**	***	***	*	***	***
Boran	- Ugogo	***	***	ns	*	***	***	***	**	ns	*	**	*	ns	ns	***	***	**	***	***
Boran	- Ujiji	**	***	ns	*	***	*	ns	***	***	***	ns	*	***	***	***	*	***	***	***
Boran	- Maasai	***	***	ns	**	***	***	*	ns	ns	ns	***	***	***	***	***	ns	***	***	***
Boran	- Sukuma	**	***	ns	*	*	***	***	ns	ns	ns	*	*	*	*	***	ns	**	ns	***
Boran	- Newala	***	***	***	***	ns	***	*	***	***	*	***	*	**	***	***	*	***	***	***
Boran	- Mbeya	***	***	ns	***	***	***	***	**	ns	ns	***	*	*	ns	***	***	***	*	***
Galla	- NEH	ns	*	ns	ns	*	ns	ns	*	*	*	ns	***	ns	**	ns	***	***	***	*
Galla	- Afar	*	*	ns	ns	***	ns	ns	ns	**	**	*	**	ns	**	ns	***	ns	ns	ns
Galla	- Ta-Co	***	***	***	***	***	***	ns	ns	ns	***	ns	***	ns	***	***	***	ns	ns	*
Galla	- Ugogo	***	***	ns	ns	***	***	ns	ns	ns	ns	*	ns	ns	***	***	***	ns	**	***
Galla	- Ujiji	**	***	ns	ns	***	*	ns	ns	**	***	ns	*	**	***	***	***	***	***	***
Galla	- Maasai	*	**	ns	*	***	*	*	*	**	*	*	***	*	***	***	*	ns	***	**
Galla	- Sukuma	**	***	ns	ns	*	**	*	ns	**	*	ns	ns	ns	***	***	*	*	***	**
Galla	- Newala	***	***	***	***	***	***	*	ns	***	***	**	**	*	***	***	**	***	***	***
Galla	- Mbeya	*	ns	ns	ns	***	***	**	ns	ns	ns	*	*	*	**	***	***	**	***	**
NEH	- Afar	**	***	ns	*	ns	*	ns	***	ns	ns	**	***	ns	ns	ns	ns	**	***	**
NEH	- Ta-Co	***	***	***	***	***	***	ns	**	***	*	***	***	**	***	***	*	***	***	***
NEH	- Ugogo	***	***	ns	ns	ns	***	**	***	***	*	***	***	*	***	***	ns	***	***	***
NEH	- Ujiji	***	***	ns	ns	***	***	ns	***	ns	***	***	***	***	***	***	ns	***	***	***
NEH	- Maasai	***	***	ns	*	*	**	*	***	**	ns	***	**	***	***	***	*	***	***	**
NEH	- Sukuma	***	***	ns	ns	ns	***	*	***	***	ns	***	***	***	***	***	***	***	***	***
NEH	- Newala	***	***	***	***	*	***	*	***	***	ns	***	***	***	***	***	***	***	***	***
NEH	- Mbeya	***	***	*	ns	ns	***		0 ***	***	ns	***	***	***	***	***	***	***	***	***
Afar	- Ta-Co	***	***	***	***	*	***	ns	**	***	*	***	***	***	***	***	ns	***	*	***
Afar	- Ugogo	***	*	ns	***	ns	***	ns	**	ns	*	***	***	ns	***	***	ns	**	***	***

Population pair	INRA 63	ILSTS 87	SRCR SP3	SRCR SP7	OarAE 129	BMS 357	INRA 5	BMS 1494	BMC 1222	INRA 132	ILSTS 17	BM 1818	MAF 209	OarFCB 304	MAF 35	ILSTS 44	ILSTS 5	MAF 65	ILSTS 11
Afar - Ujiji	***	***	*	***	*	**	ns	**	ns	***	***	***	***	***	***	ns	***	***	***
Afar - Maasai	***	**	ns	***	ns	**	**	**	ns	ns	***	***	***	***	***	ns	ns	***	***
Afar - Sukuma	***	***	*	**	ns	***	**	*	ns	ns	***	***	**	***	***	**	***	***	***
Afar - Newala	***	**	***	***	***	***	***	*	**	*	***	***	***	***	***	***	***	***	***
Afar - Mbeya	***	ns	*	*	*	***	**	*	**	ns	***	***	**	***	***	**	***	***	***
Ta-Co - Ugogo	ns	**	***	***	*	***	*	ns	***	ns	ns	***	ns	*	ns	ns	**	ns	***
Ta-Co - Ujiji	***	ns	***	***	***	***	ns	***	***	**	***	***	ns	ns	ns	*	***	***	***
Ta-Co - Maasai	*	***	***	***	***	***	**	***	***	ns	*	***	*	**	ns	ns	***	***	ns
Ta-Co - Sukuma	ns	***	***	***	***	*	ns	*	***	*	***	***	***	*	ns	**	*	***	***
Ta-Co - Newala	***	***	***	***	***	***	**	ns	***	***	*	***	ns	***	ns	***	***	***	***
Ta-Co - Mbeya	*	***	***	ns	*	***	***	ns	***	ns	*	***	ns	ns	ns	ns	***	***	***
Ugogo - Ujiji	***	***	*	ns	ns	**	ns	*	***	*	***	*	**	*	*	ns	***	***	***
Ugogo - Maasai	ns	***	ns	ns	ns	***	***	ns	ns	ns	*	ns	*	**	ns	*	ns	***	***
Ugogo - Sukuma	**	***	*	ns	ns	ns	ns	ns	*	ns	*	**	ns	ns	ns	***	ns	***	***
Ugogo - Newala	***	***	***	***	***	***	***	ns	***	***	ns	***	*	***	ns	***	***	***	***
Ugogo - Mbeya	ns	*	ns	ns	ns	***	ns	ns	ns	ns	*	***	ns	*	*	***	***	***	***
Ujiji - Maasai	**	***	ns	ns	ns	ns	*	**	***	***	***	***	*	**	***	ns	ns	***	***
Ujiji - Sukuma	***	***	*	ns	ns	*	*	**	***	***	***	***	ns	*	*	*	ns	***	***
Ujiji - Newala	***	***	***	***	***	***	*	***	***	***	***	***	***	***	ns	***	***	***	***
Ujiji - Mbeya	*	***	***	*	ns	***	***	**	***	***	***	*	**	**	ns	***	**	***	***
Maasai - Sukuma	*	***	ns	*	**	**	***	ns	**	ns	*	***	**	**	ns	ns	ns	***	***
Maasai - Newala	***	***	***	***	***	***	ns	**	***	**	ns	***	**	**	ns	ns	***	***	***
Maasai - Mbeya	ns	**	ns	**	ns	***	***	*	*	ns	***	***	ns	**	**	***	ns	***	**
Sukuma - Newala	***	**	***	***	***	***	***	ns	***	**	ns	***	***	**	ns	*	***	***	***
Sukuma - Mbeya	**	***	ns	ns	**	***	***	ns	*	ns	***	***	**	ns	*	***	*	*	***
Newala - Mbeya	***	***	***	***	***	***	***	ns	***	*	***	***	ns	***	ns	***	**	***	ns