MORPHOLOGICAL AND GENETIC CHARACTERISTICS OF *BOS INDICUS* INDIGENOUS CATTLE OF ERITREA

GOITOM SOLOMON GHEBREGERGISH

DOCTOR OF PHILOSOPHY

Animal Breeding and Genetics

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

2019

Morphological and Genetic Characteristics of *Bos indicus* Indigenous Cattle of Eritrea

Goitom Solomon Ghebregergish

A Thesis Submitted in Fulfilment for the Degree of Doctor of Philosophy in Animal Breeding and Genetics in the Jomo Kenyatta University of Agriculture and Technology

2019

DECLARATION

This thesis is my original work and has not been presented for a degree in any other University

Signature……. Date……

Goitom Solomon Ghebregergish

This thesis has been submitted for examination with our approval as University supervisors.

Dr Mathew G. Gicheha, PhD
JKUAT, Kenya

Signature……. Date……

Dr Francis K. Njonge, PhD
JKUAT, Kenya

Signature……. Date……

Dr Kiplangat Ngeno, PhD
Egerton, Kenya

DEDICATION

I dedicate this work to my late father Solomon Ghebregergish, my late sister Sara Solomon and late workmate Dr Teklehimanot.

ACKNOWLEDGEMENT

Firstly, I would like to appreciate Dr Mathew Gicheha, Dr Francis Njonge and Dr Kiplangat Ngeno who supervised this work for their continued support from the conceptualisation through data collection, analysis and interpretation of the results up until the successful compilation of this thesis. I salute my supervisors. To Professor Nyende my sincere gratitude for all the assistance he granted me from the start to the end of my PhD journey. My heart is filled with appreciation.

I would also want to recognize the Department of Animal Sciences at Jomo Kenyatta University of Agriculture and Technology for giving me a chance to undertake my studies in the department. And also for the University for granted me partial scholarship towards my tuition fees. Thank also goes to members of the staff of JKUAT for making my stay at the University good, and more specifically members of staff of the Department of Animal Sciences, you were my brothers and sisters away from home.

Mr. Semere Amlesom, the Dean Hamelmalo Agricultural College (HAC), played a very great role for accepting my request for leave to undertake my PhD studies and for the financial support granted to me by HAC, I will always remain grateful. Professor Zemenfes Tsige is highly appreciating his assistance as well as the National Commission for Higher Board (NCHE) of Eritrea in organizing and funding my PhD studies. I further recognize the support provided by the Ministry of Agriculture, Zonal Administrations.

Special recognition goes to the Japan International Cooperation Agency (JICA) for funding the research component of my studies. Special thanks go to Mr Suruzaki, JICA representative in Eritrea. Thank also goes to MoA, MoH and Zonal administration for helping me during data collection, sample preservation and packaging.

Lastly but not least, I would like to thank my mother Mrs. Dahab Solomon, my lovely wife Mrs. Nazreit Debesay and my beloved daughter Sabrina Goitom for their continued encouragement, prayers and having to contend with my absence during the study. I salute my sister Abrehet Solomon for continuous support and encouragement during the start to the end. God bless you abundantly.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

Abbreviation	Description
A_F	Allelic frequency
AFLP	Amplified Fragment Length Polymorphism
AEZ	Agro-Ecological Zone
AMOVA	Analysis of Molecular Variance
AnGR	Animal Genetic Resources
ANOVA	Analysis of Variance
ASAL	Arid and semi-arid land
BECA	Biosciences in Eastern and Central Africa
BAM	Binary Alignment Map
BGI	Beijing Genomic Institute
BLAST	Basic Local Alignment Search Tool
Bp	Base Pair
BWA	Borrow Wheel Aligner
CV	Coefficient of Variation
DAGRIS	Domestic Animal Genetic Resources Information System
DNA	Deoxyribonucleic Acid
EDTA	Ethylene Diamine Tetra-acetic Acid
F_{IT}	Departure from HWE in the entire population
F_{IS}	Departure from HWE within sub-population
F_{ST}	Departure from HWE among sub-population
EHH	Extended Haplotype Homozygosity
FAO	Food and Agriculture Organization
FAOSTAT	FAO Statistics
FASTA	FAST Alignment
GBS	Genotyping by Sequencing
GDP	Gross Domestic Product
GIS	Geographic Information System
GO	Gene Ontology

Abbreviation	Description
GWAS	Genome-Wide Association Studies
HAC	Hamelmalo Agricultural College
H_O	Observed Heterozygosity

Abbreviation	Description
H_E	Expected Heterozygosity
IBD	Identical by Distance
ILRI	International Livestock Research Institute
InDel	Insertion/deletion
JICA	Japan International Cooperation Agency
JKUAT	Jomo Kenyatta University of Agriculture and Technology
Kb	Kilo base
LD	Linkage Disequilibrium
MAF	Minor Allele Frequency
ML	Maximum Likelihood
MoA	Ministry of Agriculture
MP	Maximum Parsimony
NARI	National Agricultural Research Institute
N_f	Number of female animals
NCBI	National Centre for Biotechnology Information
NGS	Next Generation Sequencing
NJ	Neighbour-Joining
NLDP	National Livestock Development Project
N_f	Number of male animals
OR	Odds Ratio
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
Pe	Probability of incorrectly called
PLINK	Population-Based Linkage Analyses
QTN	Quantitative Trait Nucleotides
RAPD	Random Amplification of Polymorphic DNA

RFLP	Restrict Fragment Length Polymorphism
RNA	Ribosomal Nucleic Acid
SAHN	Single-linkage Agglomerative Hierarchical and Non-overlapping
SE	Standard Error
SAM	Sequence Alignment Map or Spatial Analysis Method
SNP	Single Nucleotide Polymorphism
SPSS	Statistical Program for Social Sciences
SSR	Simple Sequence Repeats
TASSEL	Trait Analysis by Association Evolution and Linkage
UPGMA	Un-weighted Pair Group Method with Arithmetic mean
VCF	Variant Call Format
VEP75	Variant Effect Prediction 75 data base
ZF_{ST}	Standardized F_{ST}

LIST OF SYMBOLS

Symbol	Description
F	Inbreeding coefficient
Ne	Effective population size
Qt	Quintal = 100 kg
β_0	Intercept
β	Linear regression coefficient
μ	Population mean
$\bar{x}$	Sample mean
%	Percentage
Mg	Microgram
μl	Microliter
σ	Standard deviation
χ^2	Chi-square

ABSTRACT

Indigenous cattle production plays important socio-economic roles at households and national levels. An improvement would therefore be desirable in enhancing household incomes, food and nutrition security. This can only happen in structured cattle breeds genetic improvements whose initial phase is determination of the genetic resources inventory. This study was therefore aimed at establishing the morphological and genetic differences that exist within and between indigenous cattle populations. The data was on 243 animals found in 27 sampling populations found in the three agro-ecological zones (AEZs) spread throughout Eritrea. Data for morphological classification of cattle populations was collected from 12 body characteristics, and analysis done by cluster and discriminant analyses. Analysis for genetic diversity study was used to compare variabilities among/within cattle populations. Structure analysis was done to infer the current number of different cattle populations. Analysis of Molecular Variance (AMOVA) was used to test the significance variability between and within cattle populations. Analysis of production data indicated that the average milk yield from Barka breed was 3.48 litres per day compared to Arado breeds at 1.98 litres. Irrespective of the farming system, most farmers selected bulls based on body size criterion as their first choice, and cows based on milk yield performance. Morphologically, indigenous cattle resources of Eritrea clustered into two groups (breeds; Barka and Arado). Generally, the mean morphometric measurements of Arado breed were lower in all body measurements. In molecular analysis, a total of 16,388 polymorphic autosomal SNPs were produced following the filtering and used for diversity, structure and signature analysis. Average allele frequency (A_F) of 0.157 was found for all cattle populations. The expected heterozygosity (H_E) per population ranged between 0.192 to 0.343. The cluster analysis was carried out and resulted three distinct groups. The signature analysis produced nine candidate genes which were significantly annotated for Gene Ontology (GO) terms. The genomic regions under positive selection harboured genes of facial morphology, skeletal and muscle system development, mineral homeostasis and immune systems. This association could have effect on characterization of ICPs and need to be considered in setting breeding program.

Keywords: AEZs, Characterization, Diversity, Landscape, Morphology, Signature

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Background information

In Eritrea, agriculture has been identified as the top priority given that about 80% of country population lives in the rural areas where subsistence agriculture is practiced. There has been a notable decline in the contribution realised from agriculture sector to the Gross Domestic Product (GDP) from as high as 34.7% in 2009 to less than 13.3% in the year 2011 (Mehler *et al.* 2014). This decline and/or slower than potential growth in the agricultural sector needs to be investigated and the necessary steps taken to increase productivity. According to the Government of the State of Eritrea in the National Livestock Development Project (NLDP, 2007) most of the country's animal and crop production occurs in the smallholder production systems with an estimated 1.9 million head of cattle, 2.1 million sheep, 4.6 million goats, 2.5 million birds and 0.1 million camels. Further statistics show that the numbers have been on the decrease although it is very slowly (FAO, 2006).

According the report of government of Eritrea in 2012, about 57.1% of its landmass of is arid and semi-arid land (ASAL) which can only be used for livestock production through grazing. Even where crop production is practiced mixed livestock-crop production system is common as animals are able to utilise crop residue as feed in addition to grazing. Most of the livestock production in Eritrea is characterised by different farming systems but is mainly dominated by smallholder pastoral, agro-pastoral and crop-livestock systems with much of the production being practiced to satisfy nutritional needs with little emphasis in commercial production (MoA, 2012). The production is mainly practiced in the ASALs and the main breeds kept are indigenous in the country. The animals are therefore adapted to the nutritional and environmental constraints experienced in such regions. The Indigenous Cattle Populations (ICPs) are distributed widely throughout the diverse geographical and Agro Ecological Zones (AEZs). These geographically isolated ICPs are subjected to local climatic conditions which has impacted the populations

differently resulting into detectable and unique characteristics that can help them to survive and reproduce in the harsh environments. A good example is the cattle populations around the Red Sea coast which have anatomical and physiological characteristics that equip them with the ability to produce and reproduce in hot (above 40°C) and salty environments (Figure 1.1; C). This is one of the many examples of adaptive characteristics acquired by the ICPs. Despite the importance

Figure 1.1. Indigenous cattle of Eritrea (Arado; A, Barka; B and

attached to the cattle production sector, little or no efforts have been made to carry out morphological and genetic inventory (characterization) of the animals so as to develop sustainable conservation and utilisation programmes.

Breed characterization based on morphological descriptions has importance in regard to production and reproduction performances. According to FAO (2012) phenotypic characterization is defined as the process of identifying distinct breeds and describing their external and production characteristics in a given environment and management system. Furthermore, phenotypic characterization is the description of breeds in terms of external characteristics (coat colour, ear shape, horn shape and others), linear body measurements (height at wither, heart girth, body length, ear length and others), production traits (body weight, milk yield and others) and reproductive traits (age at first calving, calving rate and others) (FAO, 2012). According to Notter (1999), there exists phenotypic diversity in livestock populations especially among ICPs. These phenotypic differences resulted from genetic diversity and the production environmental differences. Phenotypic characterization is a cheap tool for breed classification that can be used for rapid selection of animals in the field (Dossa *et al.* 2007). However, it is highly influenced by environmental effects, and sometimes by strong genetic and environmental interactions. Genetic diversity can be expressed based on the phenotypic level of production traits, exterior appearances,

reproduction traits and health traits. Similarly, genetic diversity has been expressed as the variety of alleles and reflected in morphological, physiological and behavioural differences between individuals and populations (Frankham *et al.* 1999). Therefore, it is important the morphological classifications to be supported by molecular characterization techniques (Gizaw *et al.* 2011).

1.2 Statement of the problem

The horn of Africa region, where Eritrea is located, is among the poorest parts of the world according to the FAO (2004) report. The report further noted that people in this area relay on agriculture in order to ensure/enhance food security. The main agricultural activities in the region are crop and livestock productions. The latter mainly occurs in ASALs, and is characterized by low animal productivities. Cattle production sector has the potential to expand (Goitom *et al.* 2016), however, only marginal growth has been experienced over time. According to Breuil *et al.* (2014), the livestock sector contribution to the Gross Domestic Product (GDP) in Eritrea was paltry 7-8 percent since 2011.

Eritrea lies in the Sahelian arid region where rainfall is scarce, but relatively suitable for livestock compared to crop production. Cattle relatively good animals in the Eritrea where they can utilize the available resources converting them into meat and milk products. Their importance to farming society are many such as provision of food for the family, draft power for the cultivation of land, cash income and other services. Despite having such huge resource potential and export opportunities, cattle production in Eritrea remains unexploited. This is despite the improved living standards and increased incomes to households which has translated to an unmet increase in meat, milk and other animal derived products and services.

Various projects have been initiated by the Government of the State of Eritrea with support from development partners. However, most of them have favoured interventions in animal health management with little emphasis on the potential use of existing genetic resources in mitigating the identified production constraints despite the existence of evidence from different parts of the world that genetic diversity has been used to address some of the factors limiting optimal production

from developing countries. The possible explanation for the failure to use the ICPs genetic resources is the lack of proper (using scientific approaches) information on the production, marketing and breeding systems in which the cattle are reared. The first step in sustainable utilisation and subsequent improvement in livestock production is the characterization of the country's animal genetic resources. Species characterization should involve all aspects relating to the identification and documentation of qualitative and quantitative attributes of breeds (FAO, 2011; Solomon *et al.* 2011). Mason and Maul (1960) described cattle populations in Eritrea, however, the approach used was non-scientific, where herders used traditional descriptors which are based on their respective ethnic groups. Furthermore, the resulting morphological and performance differences were not backed up with molecular analysis. The current research was carried with an aim of characterising the production systems in Eritrea including the morphological and genetic characteristics of the ICPs kept. The information obtained could be used in design of sustainable production and breeds improvement programmes.

1.3 Justifications of the study

Cattle production plays important socio-economic roles both at household and national levels in Eritrea. The grazed cattle sector dominates the production occurring mainly in pastoral and agro-pastoral systems. This is as a result of much of the country being too arid for sustainable crop production to occur. Indigenous cattle breeds are almost exclusively reared in the two production systems. Goitom *et al.* (2016) emphasised the importance of the indigenous cattle production sector in satisfying social and economic needs of significant population in the rural areas of Eritrea. Further, the rural livelihoods of Eritrean farmers are directly or indirectly dependent on cattle they raise. This has resulted to an enhancement of the role played by the livestock sector in the wellbeing of the people of Eritrea. More needs to be done as the productivity per animal remains one of the lowest in the region (Goitom *et al.* 2016). Phenotypic improvement has been proven to be sustainable approach to increased animal production as the genetic gain is locked in the animal population and is passed on from generation to generation. Furthermore, there is significant

phenotypic differences based on performance traits in ICPs of Eritrea which can be used in improving populations production and profitability.

Designing effective genetic improvement programme requires adequate knowledge of the available genetic resources since genetic diversity is the most important factor in livestock improvement plans. Based on the diversity of cattle, communities develop different sets of cultural and social values by which they judge, appraise and decide on breeding animal (Zechner *et al.* 2001). The genetic diversity information of the ICPs in Eritrea is lacking making genetic improvement of animal populations becomes a daunting task. This study was carried out to determine the morphological and genetic differences in Eritrean ICPs. This contributes to the breeding programmes of cattle populations of Eritrea.

In light of the climatic change being experienced, it is anticipated the changes in animal farm systems in response to the subsequent increase in climatic variability. There is need for Eritrea to determine the performance of the breeds that exist in the continent with a view of conserving those that can withstand higher ambient temperatures as well as fluctuating pastures supply and quality. The information obtained from this research would contribute to genetic improvement of the cattle populations as well as aid in matching the available genetic resources to production systems where they can be optimally utilised.

1.4 Objectives

1.4.1 General objective

The general objective of this study was to morphologically and genetically characterize indigenous cattle of Eritrea. This information is crucial in design of efficient and sustainable indigenous cattle breeding programs.

1.4.2 Specific objectives

1. To describe the indigenous cattle production systems in Eritrea,
2. To morphologically characterise the ICPs in Eritrea,
3. To genetically characterise the ICPs in Eritrea,

4. To determine the relationships between and among the cattle populations genetic resources and ecological landscapes of Eritrea.

1.4.3 Hypotheses

1. There are distinct production, breeding and marketing strategies between and among cattle producers in different agro-ecological zones in Eritrea
2. Indigenous cattle populations in different agro-ecological zones in Eritrea are morphologically similar
3. Indigenous cattle populations in different agro-ecological zones in Eritrea are genetically similar
4. Ecological landscapes of Eritrea have no impact on the relationships between and among the cattle populations genetic resources

1.5 Outline of the thesis

1. General introduction,
2. Literature review,
3. Studies,
    i. Characterization of cattle production systems and status of Eritrean indigenous cattle resources,
    ii. Morphological characterization of indigenous cattle breeds in Eritrea,
    iii. Genome-wide genetic diversity, population structure and admixture analysis in Eritrean indigenous cattle,
    iv. Landscape genomics and signature analysis to understand the genetic adaptation of Eritrean ICPs.

General conclusions and recommendations.

CHAPTER TWO

LITERATURE REVIEW

2.1 Cattle

2.1.1 Origin and classification of cattle

The wild Auroch (*Bos primigenius*) (Payne, 1970) which is considered an ancestor of all modern cattle species were originally domesticated in the nascent civilization era in the near East at about 8000 and 9000 years ago (Epstein & Mason, 1984). This was followed by development (selection) of the *Bos primigenius* to the more modern *Bos nomadicus* (Chen *et al.* 2010). This resulted from the migration of the cattle from the origin of domestication to the South-eastern areas (Indian sub-continent) which were relatively drier. There was continued migration of the new type of cattle to much of the world in the last 8000 years ago (Payne, 1970).

The first cattle in Africa came from the Near East through North of western Africa and they were *Bos taurus* type. The other route was through North-eastern Africa around Nile-basin where possibly interbreeding with local wild variants occurred (Epstein and Mason, 1984). There has been continuous movement of cattle through all ages of domestication (Dackson, 2008).

The *Bos-indicus* (Zebu) has been introduced from Indian sub-continent to Eastern Africa in the recent migration period (Dackson, 2008). Findings from Epstein and Mason (1984) described that the migration of *Bos indicus* cattle to Africa from Arabia and Asia was by Semitic tribes. The long-horned Zebu cattle crossed with humpless longhorn cattle and produced the cervico-thoracic humped cattle currently identified as Sanga (Hanotte *et al.* 2002). The breed was later spread in the greater parts of Africa (Southern Africa) by the migrating herders.

The concept of cattle classification (based on breed) started in the 19th century with categorisation focusing on human-oriented characterization (Rege *et al.* 1999). According to Dackson (2008) the breeds served as differently to human needs, and selection was targeted on individuals that expressed a trait of interest. Over a long time, breeds have evolved to specialise to specific functions such as

meat/milk/draught power production. Phenotypic and genetic variation has resulted from the cattle breeds being selected to serve different human needs as well habituating different environments. The emergent breeds of cattle, therefore, serve as a source of genetic variation, which forms the base for selection, and a wide range of breeds that have evolved in various environments represent unique sets of genetic diversity (Dackson, 2008). Generally, it has been estimated that since domestication, over 6,379 documented breeds from 30 species of livestock have been developed globally (FAO, 2000).

A survey carried out by Rege (1999) revealed that sub-Saharan Africa is the home of a total of 145 cattle breeds/strains. These breeds/strains comprise of 2 taurine longhorns, 15 taurine shorthorns, 75 Zebu (*Bos indicus*), 30 Sanga, 8 Zanga (Zebu-Sanga crosses), and 9 breeds derived from interbreeding and 6 composite breeds. Furthermore, nowadays a total of 990 cattle breeds have been reported throughout the world with 897 being classified as local or indigenous breeds (Scherf, 2000). Sub-Saharan African countries accounts for over 200 million heads of livestock, 32% of which are cattle (Herlocker, 1999) and 95% of the cattle are considered indigenous (Rege, 1994).

Eritrean cattle are broadly classified into *Bos indicus* Zebu beside the classification carried out by Mason and Maule (1960) based on their body sizes. According to Rege and Tawah, (1999), the East African Zebu breeds in which the Eritrean cattle fall are classified into sub-groups namely the large East Africa Zebu, and Zenga cattle (see Table 2.1).

Table 2.1. East Africa indigenous cattle breeds with their genetic characteristics (DAGRIS, 2007)

Group	Breed name	Main characteristics
1. Large East African Zebu	1) Barka	Active disposition
	2) Karamajong zebu	Adapted to a very dry climate
	3) Kenyan Boran	Walking and mothering ability, and large sex dimorphism
	4) Orma Boran	Tolerant to trypanosomiasis
	5) Turkana	Survive on scarce pasture and water, and walking ability
2. Small East African Zebu	1) Angoni	Adapted browsing, variable coat colour and size horns
	2) Arsi	Difficult to milk, extremely active and aggressive
	3) Jem-Jem	Well adapted to the wet and cold climate
	4) Mongolla	Tolerant to trypanosomiasis and well fleshed
	5) Nuba Zebu	Dwarf and tolerant to trypanosomiasis
	6) Ogaden	Good dairy and beef characteristics
	7) Ugogo Grey	Adapted to browsing during the dry season
3. Zenga	1) Alur	Trypanotolerant
	2) Arado	Docile, good work animal and low milk yield
	3) Bovines of Tete	Trypanotolerant
	4) Fogera	Docile temperament
	5) Borgou	Distinct sexual dimorphism

According to Rege and Tawah (1999) the Eritrean Barka falls in the large East African Zebu group which comprises of 13 breeds. The breeds are also exclusively found in drier parts of Sudan, Eritrea, Ethiopia, Somalia, Kenya, Tanzania and Uganda. Arado breed is the other major breed found in Eritrea and belongs to the Zenga cattle. The breed originated from crossbreeding the Zebu and Sanga breeds (Rege and Tawah, 1999). The Zenga population accounts for 8 breeds namely the

Arado, Fogera and Horro of Ethiopia, Jiddu of Southern Somalia, Alur cattle of Congo, Nganda of Uganda, Sukuma of Tanzania, and Bovines of Tete Mozambique.

2.1.2 Cattle in Eritrea

Eritrean history indicates that livestock rearing is an age old practice. Cattle farming is specifically carried out for provision of food and other rural household services such as draught power supply on cultivated lands, pulling carts, store for wealth status among other functions (FAO, 2006). According to the year 2013 Ministry of Agriculture (MoA, 2013) of the state of Eritrea annual report, the livestock populations were estimated at 2.2 million heads of cattle, 2.1 million sheep, 2.5 million poultry and 0.1 million camels. Approximately 95% of the cattle population is indigenous to Eritrea (Rege, 1994). Table 2.2 presents the different livestock species population trends between 2001 and 2005. Cattle population remained constant as did that of the goats, camels and sheep. The meat and milk production followed the same trend.

Table 2.2. Livestock numbers (million heads), meat and milk production of Eritrea in 2001-2005

Species	Years					
	2001	**2002**	**2003**	**2004**	**2005**	**Mean**
Cattle	1.95	1.90	1.93	1.93	1.95	1.93
Sheep	2.15	2.00	2.10	2.10	2.10	2.09
Goat	1.70	1.70	1.70	1.70	1.70	1.70
Goat meat production (10^3 Mt)	5.80	5.80	5.80	5.80	5.80	5.80
Total milk production (10^3 Mt)	69.5	56.7	56.7	56.7	56.7	59.26

Source (FAO, 2006)

2.1.3 Importance of cattle in Eritrean society

Eritrea is located in the horn of Africa which is considered among the poorest regions of the world (FAO, 2004). The area is dry and crop production is limited by inadequate and erratic rainfalls. This is despite the need to have tremendous

agricultural production growth to reduce poverty and increase food and nutrition security. Indigenous cattle resources in Eritrea satisfy social and economic needs of significant population in the rural areas of Eritrea (Goitom *et al.* 2016). They are mainly reared in traditional systems where feeding, breeding and general management practices have evolved overtime to match with respective agro-ecological zones resources and knowledge base. The animals are exclusively grazed except during crop harvest time when they are supplemented with the crop residues in the crop-livestock production systems. Much of the grazing occurs in the naturally available grasslands in arid and semi-arid areas. The populations are also characterised by low productivity and profitability mainly provide means for subsistence to rural households (MoA, 2012). Table 2.3 presents the land use distribution in Eritrea and it is evident that browsing and grazing land accounts for more than half of the landmass in the country at 57.16%. This is an indication of the importance of the grazed livestock sector in the country's economy.

Table 2.3. Land use distribution of Eritrea (Government of Eritrea, 2012)

Land use	Percentage
Cultivated land rain-fed	3.42
Irrigated land	0.18
Disturbed forest	0.43
Forest plantations	0.08
Woodland and scrubland	5.52
Browsing and grazing land	57.16
Barren land	33.21
Potential (possible) irrigated land	4.92
Potential rain-fed land	8.61
Total	**100%**

2.2 Cattle production systems

2.2.1 Types of production systems

According to Hans (1982), there are five distinguishable production systems in agriculture and they are: pastoral range-livestock production systems, crop-livestock

production systems in the lowland, crop-livestock production systems in the highland, ranching systems, and landless livestock production systems. FAO (1996) classified the systems based on the level of investment, and levels of technology and complexity into either traditional or modern with much of Eritrea systems falling into the traditional category which is characterised by low input and corresponding low output. The modern systems of production system require huge capital investment accompanied by highly skilled labour whereas the traditional type relies on the family labour and is characterised by extensive use of farmland resources. FAO (2002) further classified the production systems into grassland-based cattle production systems (pastoralism and ranching), mixed rain-fed systems, mixed irrigated systems and landless systems. All these variants are found in Eritrea with differing proportions. Kremen *et al*. (2012) classified farming systems as either being diversified or specialized. The diversified farming system analysis considers whole-agricultural system including the agro-ecological principles and social issues. Table 2.4 shows the classification of traditional livestock production systems in Sub-Saharan Africa.

Table 2.4. Classification of traditional livestock production systems in sub-Saharan Africa based on the priority of livestock species

Grassland-Based		Mixed rain-fed		
		Semi-arid	Humid/Sub-humid	Highland (HL)
Species	Cattle Sheep Goat	1. Cattle 2. Goat 3. Sheep	1. Cattle 2. Goats 3. Sheep	1. Cattle 2. Sheep 3. Goats
Breeds	Indigenous	Indigenous	Indigenous & Exotic *	Indigenous & Exotic ***
Output	Milk Meat	1. Milk 2.Draught power	1. Meat 2.Milk & Draught power	1. Draught power 2. Meat & Milk

Use of exotic breeds: *** = very important, ** = moderately important, * = some importance

Source (FAO, 2002)

With the anticipated increase in animal protein demand resulting from the increase in human population and better living standards there is a need to orient the production systems towards specialised production whose focus is an increase in productivity on per unit production resource (e.g. per cow basis). This is besides the impact of climate change on crop and animal production which has made agricultural production unpredictable to satisfy the growing demand (Cribb, 2010; Childers *et al.* 2011).

Cattle production systems are rooted in ecological conditions. Laura (2014) described agro-ecology as the application of ecological concepts and principles to the design and management of sustainable agro-ecosystem. Eritrea is characterised by three major Agro-Ecological Zones (AEZs) with diversity within each of the zones. Much of the land is considered arid or semi-arid which is characterised by extreme seasonal variations in temperature, humidity and sunshine intensity. This results into diversity in agro-ecological conditions ranging from relatively cold conditions in the high altitude (the highest being 3018 metre above sea level) to very hot conditions in areas lying very low in Eritrea. The three AEZs are the highland (HL), western lowland (WLL), and eastern lowland (ELL) or eastern coastal plain zones. Within each of these there are sub-agro-ecological zones which are detailed in MoA (2009) report as follows:

1. The HL zone is found on average altitude of 1,500 metres above sea level, with average annual rainfall ranging between 400 and 700 mm per year. This zone includes three sub-zones namely; highland plateau, Midlands, and East and West escarpments.
2. The WLL is low in elevation and mostly flatland with semi-arid climatic conditions. It has two sub-zones namely the South-WLL and the dry North-WLL.
3. The ELL is a sea coastal plain zone with an elevation 600 meters above sea level to below zero sea level. The climate is desert-like, where annual rainfall is less than 200 mm.

2.3 Cattle characterization

2.3.1 Phenotypic characterization

Cattle characterization is defined as the process of identifying distinct cattle breed by describing their characteristics and their production environments (FAO, 2011). Two approaches were mentioned in the characterization of cattle breeds (FAO, 2011).

1. Exploratory method is used to explore the presence of diverse breeds in the focused area.
2. Whereas the confirmatory approach is carried out in a situation where breed basic information such as distribution and identification is available. In such conditions, only need to authenticate the identified breed.

Irrespective of the characterization approach used in identification of the concerned breeds, data on physical features, appearance, productive and reproductive attributes of the breeds, images of typical adult males and females, origin and environment, known functional and genetic relationships, responses of the breed to environmental stressors, and relevant indigenous knowledge should be collected (FAO, 2011). Phenotypic (morphological) and genetic (molecular) characterization are the two major approaches used in identification of differences that exist between and within species populations. Phenotypic characterization denotes the procedure of identification of distinct populations and description of their external features and

performances within a specific production environment (Gamaniel and Gwaza, 2017). Production environment includes the natural environment (climate, terrain, vegetation, are among others) and management practices. The phenotypic characterization describes the individuals morphometric and physical characteristics. Delgado *et al* (2001) noted that having knowledge on the morphological variations between and within animal populations as being the first step in characterization of local cattle. According to FAO (2011) there are two descriptors of the variation and they are the qualitative and quantitative characteristics. The qualitative classification is based on external physical form, shape, colour and appearance of animals. These qualitative traits are affected by few genes unlike quantitative traits which are governed by many genes (polygenes). Quantitative traits are measurable and are directly correlated to production and reproduction functions.

2.3.2 Genetic characterization

Phenotypic characterization provides a rough estimate of the genetic make-up of a given population (Meghen *et al.* 1994) while the genetic characterization reveals the genetic basis of phenotypes of an animal, their patterns of inheritance, within-breed genetic structure and levels of variability, and relationships between breeds/populations (FAO, 2015). Genetic characterization of breeds allows the evaluation of genetic variability which is a fundamental element in working outbreeding strategies and genetic conservation plans.

DNA polymorphisms have become a tool for the assessment of genetic diversity among cattle breeds. The ability to detect polymorphisms at the DNA level (marker) has led to new approaches in the genetic analysis of farm animals (Güven *et al.* 2010). A marker is an easily detectable gene or sequence of DNA that shows polymorphisms and serves as a source of information in identifying genes responsible for a trait. This is true if the type of polymorphisms is found in the coding region (exon). Therefore, molecular markers have revolutionized the ability to characterize genetic variation and rationalize genetic selection (Barcaccia *et al.* 2013). Furthermore, understanding of the pattern of genetic variability among breeds

through genetic characterization helps in the development of more rational breeding programs. There are a variety of markers as discussed below.

biochemical markers which includes blood types and isozymes (Lirón, 2002). These markers represent biochemical traits that could be analysed by protein electrophoresis. The application of these markers is limited because proteins and isozymes are not genetic materials. They are products of gene expression, so they are affected by environmental factors (Drinkwater and Hazel, 1991).

A genetic marker is a gene or DNA sequence with a known location on a chromosome and associated with a particular gene or trait (Firas *et al.* 2015). A genetic marker may be a short DNA sequence, such as a sequence surrounding a single base-pair change (Single Nucleotide Polymorphism (SNP)), or long like microsatellites. Genetic markers have been comprehensively exploited to assess genetic variability as they contribute information on every region of the genome. For a successful implementation of Quantitative Trait Loci (QTL) within selection programs, identification of specific marker is needed (Ron and Weller, 2007).

A study by Parker *et al.* (1998) has been shown that markers revealing polymorphism at the DNA level play an important role in cattle relationship studies. A breeder wants to select cattle with superior genetic potential as parents for the next generation, and for this purpose genetic markers are used to support these conventional breeding strategies. There are a number of genetic markers that have been identified. They are Restriction Fragment Length Polymorphism (RFLP), Random Amplification of Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), Microsatellites also referred to as Simple Sequence Repeated (SSR) or Variable Number of Tandem Repeats (VNTRs) andSingle Nucleotide Polymorphism (SNP). Deoxyribonucleic acid sequencing is the process of determining the complete DNA sequence of an organism's genome at a time. The genomes of several specimens are sequenced to discover large numbers of markers (SNPs) for exploring within-species diversity, performing Genome-Wide Association Studies (GWAS) (Metzker, 2010) and genomic prediction. The evolution of DNA sequencing has come a long way since in the 1970s and summarized below into three sequencing generations (first, second and third generations)

Next generation sequencing involves sequencing of either whole genome, specific genomic region, whole-exome or RNA. Such sequencing generates a lot of data and for scientists to make sense out of the data, several steps and numerous bioinformatics methods are used. The general steps in variant calling includes base calling, quality control, trimming reads, mapping reads to the reference genome, recalibration of quality scores, and variant identification. All these steps rely on different bioinformatics tools and methods.

2.4 Genetic diversity

Diversity among living organisms is a result of variations in DNA sequences (mutation) and environmental effects (Rege, 1999). The diversity of cattle breeds started due to their coexistence with humans. Humans started selected the modern cattle from wild ancestors to play certain functions such as provision of meat, milk, power leading to the application of selective pressure and subsequently distinct breeds (Laercio, 2013). Cattle genetic diversity is important in meeting the current production needs in various environments and allow sustainable genetic improvement (Notter, 1999). Breeds of cattle, therefore, act as sources of genetic variation which forms the base for selection. In farm animals and particularly in cattle, artificial selection pressure has increased very rapidly since the advent of quantitative genetic methods and reproductive technologies (mainly artificial insemination) (Powell *et al.* 2003). These techniques aggravate the loss of genetic diversity by selecting animals with good production traits and culling others even though they have importance in non-production traits. Traditional farmers are characterized by maintaining diverse livestock species using their traditional knowledge (Miguel *et al.* 2012).

Diversity could be described based on either the phenotype or genetic make-up of an animal. Phenotypic diversity can be measured by calculating the variances in the performance of the phenotype. It has more advantages over genotypic diversity due to fewer algorithms used in computing phenotypic diversity which usually use the morphological, production and reproduction values of the individuals. This implies that it is easier to carry out and takes much less time compared to the genotypic

diversity approach. However, genetic diversity is more important in breed characterization because it provides information on the genetic variations within and between populations. Genetic markers such as SNPs and microsatellites are commonly used for assessment of genetic variability in cattle studies. Distance and character-based methods, Hardy-Weinberg Equilibrium (HWE) and F-Statistics, Principal Components Analyses (PCA) and cluster analysis are the approaches used in evaluation of populations genetic diversity.

Phylogenetic analysis is the procedures used to reconstruct the evolutionary relationships among a group of related molecules or organisms (Khan *et al.* 2014). Phylogenetic inference is designed to enable users to differentiate orthologs (related by speciation events) from paralogs (gene duplication) (William, 2003). A phylogenetic tree is a mathematical structure which is used to represent the evolutionary history of a group of sequences or organisms (Roderic, 2003). A tree consists of nodes (ancestors), edges (connection), leaves (terminal taxa) of sequences or organisms (Roderic, 2003).

2.5. Landscape genomics and adaptation

Indigenous cattle living in different AEZs are subjected to different natural and human selection pressures. This has exemplified the genetic make-up of these cattle. Natural selection imposed by the environment especially in the arid area is believed to have resulted in animals with different adaptive genes than their counterparts found in the wet environment. Therefore, cattle populations raised under diverse production systems across variable environmental areas are based on the demand of nature or/and human beings.

2.5.1. Landscape genomics

Manel *et al* (2003) explained landscape genetics is the union of population genetics and landscape ecology. Landscape genomics, further described as the correlation between genomic data and environmental parameters usable in inferencing the genomic basis of local adaptation (Joost *et al.* 2007). Separate studies on the genetic make-up of cattle and its environmental effects could not show the adaptability of a

specific cattle population in a specific environment. Therefore, landscape genetics has revealed the causes of adaptability. Manel *et al* (2010) also stated the use of landscape genomics to uncover the environmental drivers of local adaptation and the underlying candidate gene/genes networks. Similarly, Joost *et al* (2007) stated that landscape genomics has a goal in identifying loci having adaptive significance in the genome.

Recently, NGS technique and high-density SNPs chips allow characterization and detection of adaptive loci (Cock *et al.* 2013). Through landscape genomics, source of functional diversity and uniqueness of indigenous cattle populations can be delineated in terms of natural selection imposed by the extended geographic distribution across agro-ecological zones, production systems and artificial selection. New methods of detection of signatures in cattle are based on the application of spatial analysis (Joost et al. 2007). Spatial Analysis Method (SAM) requires a geo-referenced data set to describe the sampling location. Similarly, the genomic extent of local adaptation can also differ among populations depending on the degree of genetic isolation (Feder and Nosil, 2010). Stratification of populations is done by the distance among the populations. Too long distance among cattle populations resulted divergence of populations due to Isolation-By-Distance (IBD) (Joost *et al.* 2016). Besides the reduction in gene flow due to IBD, divergence can also be hastened by the geographical barrier (Barton and Bengtsson, 1986; Gavrilets and Cruzan, 1998) in a process often referred to as isolation by adaptation (Hendry, 2004; Nosil *et al.* 2009). Geographical coordinates are often used as a substitution for the climatic and production selection pressures, which are modelled to investigate the genetic adaptation (Mdladla, 2016).

2.5.2. Tracing selection footprints in the cattle genome

Detection of selection signatures within the genome of organisms is a key point, since it allows a greater understanding of what proportion of a genome or which genes have been/are being shaped by past and ongoing natural selection (Joost *et al.* 2007). Similarly, Perez *et al* (2014) has explained on the determination of signatures in specific regions of a genome that have been preferentially increased in frequency

and fixed in a population due to natural or artificial selection. In general, regions of the genome that are under selection are likely to be of functional importance, and inferences regarding selection may provide important information (Nielsen, 2005).

A number of statistical methods are used to scan the genome for signals of selection due to domestication and adaptation. The two commonly used methods used to scan the genome for signals of selection due to domestication and adaptation are the single-locus (F_{ST}) and haplotype-based (EHH; extended haplotype homozygosity) (Kim *et al.* 2016). The single-locus method uses average allele frequency at each SNP locus and the fixation index F_{ST} (Weir and Cockerham, 1984; Rubin *et al.* 2010). Identification of a regions affected by selection is based on the level of F_{ST}. Genomic region that shows elevated levels of F_{ST} between populations are considered as selection signatures. The second approach is based on EHH using integrated haplotype score to identify the selection of signatures in the cattle population. The selected regions could have strong selective advantage increases quickly in frequency until reaching fixation is known as hard sweep. In contrast, when the selected allele is less pronounced and the increase frequency is at the beginning phase is known as soft sweep (Przeworski *et al.* 2005). On the other hand selective sweep has balancing effect favors maintenance of polymorphism (Oleksyk *et al.* 2010).

CHAPTER THREE

CHARACTERIZATION OF PRODUCTION SYSTEMS AND TRADITIONAL HUSBANDRY PRACTICES OF INDIGENOUS CATTLE RESOURCES

3.1 Introduction

3.1.1 Background

Agriculture is the most important sector of the Eritrean rural economy and accounted for 8% of the gross domestic product (GDP) in 2005 (MoA, 2013). However, much of the sector relies on rain-fed agriculture which has been disrupted by climate change. The sector also depends on traditional methods of production which are limited by need to produce more from each individual animal. The climatic condition of Eritrea is conducive for livestock production with approximately 49% of the total land area being suitable for livestock grazing compared to 17% which is usable for crop production (FAO, 1994).

Most of the livestock production occurs in the arid and semi-arid regions (ASALs) of Eritrea. Most areas of the country are found in the ASALs where farmers practice pastoral and agro-pastoral cattle production system in the lowlands. Nevertheless, in the highlands (HL) which are characterised by low temperatures and relatively high rainfall mixed crop-livestock production system has been practiced. Cattle are mainly used as draft power in these systems. Based on the report of National Livestock Development Project of Eritrea (NLDP, 2007), most of the country's cattle production is found in the smallholder production systems. The animals are mainly kept in traditional systems where feeding, breeding and general management practices have evolved to match with respective agro-ecological zones (AEZs) resources and knowledge base.

Indigenous cattle resources satisfy social and economic needs of significant population in the rural areas of Eritrea (Goitom *et al.* 2016). The rural livelihoods of Eritrean farmers are directly or indirectly dependent on cattle they raise. To enhance

the role played by the livestock sector in the wellbeing of the people of Eritrea, the Eritrean government is in the process of implementing livestock improvement programs to optimise on their productivity thus harnessing food and nutrition security (MoA, 2013). The information presented this study would contribute to genetic improvement of the cattle populations as well as aid in matching the available genetic resources to production systems where they can be optimally utilised. The findings presented in the thesis would form the basis of design and implementation of indigenous cattle breeds improvement programme.

There are three identifiable cattle production systems in Eritrea. These are pastoralism, agro-pastoralism and mixed crop-livestock system. The production approaches and animals reared in the three systems vary significantly, however, in all systems, traditional knowledge and practices are used in livestock production (MoA, 2013) with an exception of the mixed crop-livestock system where to a lesser extent modern livestock production practices have been adopted. The pastoralist and agro-pastoralist cattle production systems are mostly practiced in Western lowland (WLL) and Eastern coast (ELL) regions of the country. These areas are characterized by low annual rainfall with about nine-month dry season. Figure 3.1 (Keru sampling area) shows herders sheltering under a make shift structure. The area is characterised by scarce grazing resources and very high ambient temperatures.

Figure 3.1 Herders (pastoralists) sheltering from the high temperatures as the cattle drink water from a temporary pan in Keru sampling area. Source: (own photo)

Cattle movements are determined by the availability of water and feed resources. The implication being that such system requires animals that are hardy in order to walk long distances in the tough environment characterised by scarce feed and water resources (Saidu and Omedo, 2010). Breeding practices in such constrained production environments dictate that the selection criteria be on adaptive traits that counteract the extreme climatic conditions, feed scarcity as well as water shortages.

The mixed crop-livestock is the most predominant production system in the highlands (HL). The animals are mainly utilised in ploughing land (indigenous breeds) and milk production (exotic). The farmers use the crop residue and crop by-products are used in supplementing the grazing animals. The system is characterised by varying levels of intensification depending on the economic and knowledge of the producer. Farms run by well to do and educated producers (have access to

information and capital) are characterised by high input/high output resources. Animal droppings collected from the animal yard is used as a manure in crop farm (MoA, 2013).

The existing literature indicates that there is scarce information on livestock production systems in Eritrea. The little available information is mainly obtained from development reports which do not strictly adhere to scientific approaches in the systems analysis. This chapter is dedicated to the description of the indigenous cattle production systems in Eritrea through a review traditional breeding practices with special emphasis on selection preferences based on cattle production system and AEZs, evaluation of production and reproduction performances of indigenous cattle populations and the assessment of the status of cattle population dynamics in relation to inbreeding status. The aim of this study was to characterize the cattle husbandry practices of Eritrea and determining the status of indigenous cattle resources of Eritrea in relation to population dynamics by determining inbreeding status,

3.2 Materials and methods

3.2.1 Study area

Figure 3.2 shows the study area covered in this study. The research was conducted in three AEZs of Eritrea which lies between latitudes 12^0 42' N to 18^0 2' N and longitudes 36^0 30' E to 43^0 20' E. The three AEZs are namely the Western lowland (WLL), the Highland (HL) and Eastern lowland (ELL). Data were collected from all the regions of the country (Figure 3.2).

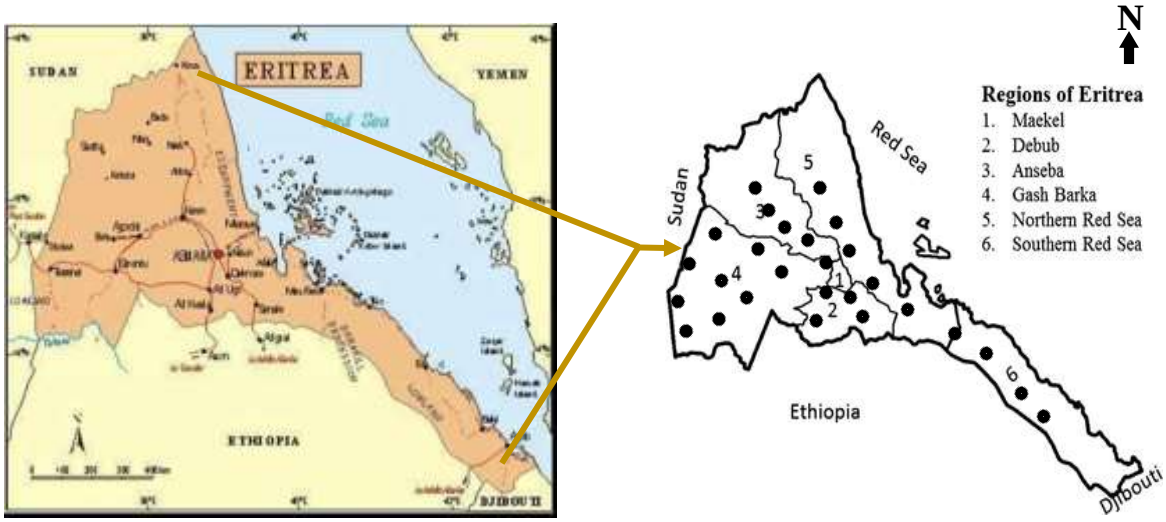


Figure 3.2. Map of Eritrea and sampling sites (Source: World Atlas, modified)

3.2.2 Data collection

A structured questionnaire was used in data collection (Appendix 1). A total of 243 producers were interviewed and relevant data recorded for analysis based on reasonable standard error (Arsham, 2005) and guidelines prescribed in FAO (2015). The questions were specifically set to collect data on the production, breeding and marketing practices utilised by different producers. The questionnaire was developed following the guidelines presented in FAO (2012) report. The questionnaire was piloted by engaging zonal administration officials, zonal ministry of agriculture officers (experts) and representative local farmers. The questionnaire was adjusted where need be identified following the piloting. The questionnaire was then used to collect data from all respondents for the respective analysis. The data obtained corresponded to 15 demography information questions, 26 breeding questions, 27 husbandry questions, 13 production/reproduction indicators (utility characteristics), 1 farm constrain question. The analysis reliability and validity was supported through use of records obtained from focus groups through discussions and informal interviews to describe traditional cattle production. Existing literature on production and breeding systems in other countries neighbouring Eritrea were also analysed to support the data obtained. Table 3.1 shows the number of samples in the three AEZs and six administration zones (Anseba, Debub, Gash-Barka, Maekel, Northern Red Sea and Southern Red Sea).

Table 3.1. Distribution of samples of cattle population and number observations (animals) based on AEZs

AEZ	Sub-AEZ	Cattle populations	Number of Animals/producers
1. HL	Southern highland	4	36
	Northern Highland	5	45
2. WLL	South western lowland	5	45
	North western lowland	4	36
3. ELL	South ELL	4	36
	North ELL	5	45
Total		**27**	**243**

HL= Highland, WLL=Western Lowland and ELL=Eastern Lowland

3.2.3 Data analysis

3.2.3.1 Descriptive statistics

Data on the status of the cattle and traditional husbandry practices were analysed to determine descriptive statistics (mean ($\bar{x}$), standard deviation (σ) and percentage (%)) for demography, utility characteristics (production and reproduction performances), breeding, feeding, importance, and constraints were analysed using SPSS software, (2006).

3.2.3.2 Indexing

For trait preference and farm constrains data, ranking index was developed. The rank ranged between 1 and 3 in accordance to the importance with 3 ranking lowest (Kosgey, 2004). A total of 8 traits were presented to each participant (Appendix 2). The traits included body size, milk/meat yield, libido, pedigree, adaptability to environmental stress, growth, temperament, and age at first maturity. The farmer was asked to rank them for bull or cow selection. This implies that a trait would be ranked high in selecting a cow and low in selecting a bull. Sex limited traits were considered in comparing the selection criteria for the two sexes. However, they were ranked respective to the sexes. Index was calculated according Kosgey, 2004; Index

= Sum [3 for rank 1, 2 for rank 2, and 1 for rank 3] of individual reason divided by summed for all reasons.

3.2.3.3 Univariate analysis

Univariate mixed model was used to analyse quantitative data taking AEZs as fixed effect and cattle populations as random effect in the analysis using Genstat software, (2014). The statistical model that used for quantitative measurements was:

$$\gamma_{ijk} = \mu + \alpha_i + \beta_j + (\alpha \times \beta)_{ij} + \varepsilon_{ijk}$$

Where: γ_{ijk} = K^{th} observation on quantitative measurements; μ= the overall mean; α_i = the fixed effect of agro-ecological zone (i = HL, WLL and ELL (Sea coasts)); β_j = random effect of a population nested on agro-ecological zone (j = 27 cattle populations); $(\alpha \times \beta)_{ij}$ = the interaction effect of agro-ecological zone with animal population; ε_{ijk} = the effect of random error.

3.2.3.4 Logistic regression

Logistic regression analysis was used to determine the probability of selecting a cow/bull following different selection criteria and agro-ecological zones while Kruskal-Wallis test (Sheskin, 2011) was used in calculation of the odds ratio. The ratio was used to determine the probability that a farmer would prefer either of the eight traits considered in this study. Agro-ecological zones were fitted to the model as independent variable. The logit model used for analysis was:

$$\text{In}\left[\frac{P}{1-P}\right] = \beta_0 + \beta_1 X_1 + e$$

where; P is the probability of a trait preferred, (1–P) is the probability that a trait is not preferred, $\left[\frac{P}{1-P}\right]$ is the odds ratio, X_1 is the agro-ecological zones, β_0 is the intercept; β_1 is the linear regression coefficient and e is the random residual error.

3.3 Effective population number

The Ne of the parental population was calculated based on the formula by Falconer and Mackay (1996).

$$Ne = \frac{(4N_m N_f)}{(N_m + N_f)}$$

where; N_m and N_f are actual number of male and female animals respectively

The Ne in this study was determined for Barka breed due to the unavailability of data for the Arado breed. Furthermore, Barka breed was found to be more dominantly reared in Eritrea. The inbreeding coefficient was calculated as:

$$F_t = 1 - \left[1 - \frac{1}{2Ne}\right]^t$$

where F_t is the inbreeding coefficient within time (t) interval.

3.3.1 Demographic information

Table 3.2 shows the analysis of the demographic information obtained from the study areas. Most respondents practiced traditional animal husbandry while the head of most of the households interviewed were men at 97.3%. Approximately 72.7% of the farmers (head of households) were aged 60 years and above. Approximately 45.6% of the producers were also illiterate implying that they could neither read nor write. The level of illiteracy varied between and amongst agro-ecological zones. Approximately 69.2% of the literate farmers were found in the highlands compared to only 36.4% in ELL.

Table 3.2. Household information in percentage in the three AEZs

	Level of education (%)			**Family size (%)**		
AEZ	**Literate (1-12)**	**Illiterate**	**Read & Write**	**Small (1-4)**	**Medium (5-9)**	**Large (>9)**
HL (N=72)	45.7	30.9	23.5	9.1	75.3	15.6
WLL (N=64)	39.4	42.4	18.2	17.9	74.6	7.5
ELL (N=76)	27.3	63.6	09.1	22.7	64.0	13.3
Mean	**37.5**	**45.6**	**16.9**	**16.6**	**71.3**	**12.1**
	Age (%)		**Sex (%)**			
AEZ	**Young & medium (<60 years)**	**Old (>60 years)**	**Male**	**Female**		

HL (N=72)	16.7	83.3	97.5	2.5
WLL (N=64)	29.7	70.3	97.0	3.0
ELL (N=76)	35.5	64.5	97.5	2.5
Mean	**27.3**	**72.7**	**97.3**	**2.7**

Table 3.3 presents the family incomes, farm activities, farm sizes, number of cattle, relative importance of livestock and the household incomes in the three AEZs of Eritrea. Agriculture accounted for mean 90.5% of the family income with trade and other activities accounting for 8.7% and 0.8% respectively. Those respondents mostly involved in agriculture have farmlands that are less than or equal 4 hectares (<1 ha; 15.7% and 1-4 ha; 75.2%) and above 4 hectares (9.1%). The number of animals reared by households differed amongst the AEZs with the highest being less than 5 heads of cattle in the highlands at 65.4% and the lowest in WLL. The livestock importance index was highest for cattle at 0.51 and lowest for the poultry at 0.04. Whereas, family income from crop production ranked first (0.43) proceeding by cattle production.

Table 3.3. The percentage family income, farm activities, farmland (hectare), cattle number, livestock importance and household income in the three AEZs

	Family income (%)			**Farm activity (%)**		
AEZ	**Trade**	**Agriculture**	**Others**	**Crop**	**Livestock**	**Both**
HL (N=81)	3.70	96.3	0.00	25.90	2.50	71.60
WLL (N=67)	5.00	95.0	0.00	0.00	10.00	90.00
ELL (N=81)	17.30	80.2	2.40	40.70	40.70	18.50
Mean	**8.70**	**90.50**	**0.80**	**22.20**	**17.70**	**60.00**
	Farmland (%)			**Cattle number (%)**		
AEZ	**<1 ha**	**1-4 ha**	**>4 ha**	**< 5 heads**	**5-10 heads**	**>10 heads**
HL (N=81)	24.60	71.50	3.70	65.43	32.10	2.47
WLL (N=67)	0.00	81.40	18.90	21.25	27.50	51.25

ELL (N=81)	22.20	72.80	4.90	59.26	35.80	4.94
Mean	**15.70**	**75.20**	**9.10**	**48.65**	**31.80**	**19.55**

	Livestock importance		**Household income**		
AEZ	**%**	**Index**	**Production systems**	**%**	**Index**
Cattle	74.90	0.51	Cattle	43.80	0.42
Sheep	15.06	0.28	Crop	47.93	0.43
Goat	9.62	0.17	Sheep & goat	07.85	0.15
Poultry	0.42	0.04	Horticulture	0.41	0.01

3.4 Results

3.4.1 Cattle husbandry practices

Table 3.4 shows the cattle husbandry practices carried out by different cattle keeping communities in the three AEZs. Different practices were carried out for different reasons with culling (cows) in relation to old age (46.5%) being the most common practice amongst the livestock keeping communities in the study area. Castration was mainly done to avoid unwanted breeding with a proportion of farmers (27.6%) doing it for fattening function. Most (63%) of the cattle production occurred in communally owned and grazed land. Approximately 51.4% farmers supplemented their animals with crop by-products and grounded cereals with 41.2% mostly doing it during the dry seasons.

Table 3.4. Proportion of major indigenous cattle husbandry practices

Practices	**Option**	**%**	**Practices**	**Option**	**%**
Castration	Yes	60.1	Type of grazing	Communal	61.3
- Purpose	Fattening	27.6	Shelter	No	51.4
Cow culling	Yes	95.1	Dry season grazing	Free grazing	42.8
- Reason	Old age	46.5	Wet season grazing	Herded	30.5
Bull culling	Yes	86.4	Provide supplement	Yes	51.4
- Reason	Old age	49.8	- When	Dry season	41.2

Shelter was not provided to animals in pastoral systems except makeshift structures meant for nights (51.4%). Some farmers practising mixed crop-livestock and agro-pastoral systems however, provided shelter to their animals to protect them from adverse weather as well as predation.

3.4.1.1 Traditional knowledge in husbandry

Traditional knowledge and alternative management practices varied between and within the AEZs. Table 3.5 presents most commonly shared traditional practices between and among different indigenous cattle producers in Eritrea.

Table 3.5. Shared traditional cattle rearing practices in Eritrea

Practice	Description
1. Milk let down stimulation	
• Suckling	Allowing the calf to suckle dam during milking (face to face milking)
• *Arem*	In all cases where the practice is common, the died/slaughtered calf's skin was sprayed with milk and the cow allowed to lick and this skin would always be shown to the cow during milking. Unstimulated milk let down was limited in many indigenous cattle breeds mainly reared in pastoral systems
• *Mbkuae*	Air is blown into the vulva of a cow whose calf has died or slaughtered. The air would be retained in the vulva which would result to milk let down
2. Milking schedule	
• *Dario*	Skipping milking for a day mainly when there is drought. This ensures the cows are not subjected to nutrients deficit in already stressful time of the year.
• Drying/control suckling	Use of thorns when drying off and use of net around the calf's mouth for restrict suckling
3. Feeding and watering	
• *Timyo*	Failure to provide the cattle with water for a whole day during dry seasons or during long distances to get water. *Sitio* is a term used during the days when the cattle are provided with water. Timyo mainly occured in the lowlands
• *Mdgual*	In this practice grazing animals are forced to rest under tree shades (Figure 3.3 A)
• *Aleda*	Involves night grazing to avoid the high temperatures during the day
4. Housing	
• *Dembe*	A form of shelter where the animals are surrounded using thorny bushes either within or outside the homestead (Figure 3.3 B)

Figure 3.3. Medagul (A) in Shambuko site and dembe shelter (B) Goluj site (own photo)

Traditional cattle naming

Table 3.6 presents the indigenous cattle naming system used in different agro-ecological zones in Eritrea. They are mainly based on the coat colour, presence/absence of the horns besides the positioning, and body size. The names differ between sexes for the same naming criterion. The differences in the naming system is explainable by the differences in culture and language amongst the cattle rearing communities.

Table 3.6. Some of local cattle names based on external appearances in highland and western lowland in Eritrea

Names based on	Characteristics	Local name		Common AEZs[1]
		Cow	Bull/Ox	
Coat colour	Pure black	*Drit*	*Duruy*	WL
	Light black	*Gobayt*	*Gobay*	HL
	White	*Halibet*	*Halib*	WL
		-	*Shawush*	HL
	Red	*Tala*	*Hamer*	WL
		-	*Lemin*	HL
	Red with black eyes	*Kuhil*	-	HL
	Red with white face	*Berha*	*Ashaal*	HL
	Patchy black and white	*Adelway*	-	WL
		-	*Berih*	HL
	Any colour with white spots on neck	*Koayt*	*Koay*	WL

	Creamy	*Sotay*	*Sotay*	WL
		-	*Senay*	HL
	Grey	*Chelayt*	*Megal*	HL
		-	*Chelay*	
	Grey around neck and head	*Tekulay/Fahray*	-	HL
	Red with small white spots	*Urube*	*Urub*	HL
	Red with white strips	*Ashkeray/Sulum*	-	HL
Horn position	Polled	*Dombay*	*Dombay*	HL, WL
	Curved upward	*Guaguday*	*Guaguday*	HL, WL
	Long curved upward	*Felak*	*Felak*	HL, WL
	Short horn	-	*Gombel*	HL
	Lateral horn	-	*Game*	HL
	One horn upward other down ward	*Kalshay*	*Kalshay*	WL
Body size	Long body	*Berik*	*Barkay*	WL
	Small body (round)	*Kebab*	-	HL

[1]WL (Western lowland); HL (Highland)

3.4.2 Utility characteristics

3.4.2.1 Milk production and reproduction performances

The results for milk production and reproduction performance of different cattle types studied are presented in Table 3.7. The milk yield was 1.98 ±0.22 and 3.48 ±0.27 litres for Arado and Barka breeds respectively. The respective lactation length (LL) in months for Barka and Arado breeds were 4.23 ±0.54 and 5.64 ±0.16, while age at first mating (AFM) in years 3.57 ±0.22 and 3.04 ±0.09, and calving interval (CI) in years was 1.77 ±0.07 and 1.45 ±0.12.

Table 3.7. Mean milk production and reproduction performances of cattle breeds (P <0.05)

Cattle type	Milk Production (lt)	Lactation length (month)	Age at first mating (year)	Calving interval (year)
Arado	1.98 ±0.22	4.23 ±0.54	3.57 ±0.22	1.77 ±0.07
Barka	3.48 ±0.27	5.64 ±0.16	3.04 ±0.09	1.45 ±0.12

Much (62.0%) of the calving occurred in the months of June to September with 97.5% of the cattle giving birth to one calf per calving. Approximately 33.7% of the producers interviewed did not practice weaning with the calf being left to run with the mother until such a time she could no longer allow the calf to suckle. About 65.0% of the farmers reported that they were milking the cows twice (morning and evening) with the rest milking their cows once in the morning while allowing the calf to run with the cow (mother) throughout the day while being separated for night. Approximately 93.4% of the interviewed households reported consuming milk while 36.0% of those who had excess sold it locally. Locally processing of the raw milk into butter for sale and household use was reported by 67.1% of the respondents.

3.4.3 Trait preference

3.4.3.1 Ranking preferences

Trait preference differed between bull and cow selection are shown in Table 3.8. Bulls were mainly selected based on body size (52.26%; index=0.35) which ranked first and pedigree (12.35%; index=0.16) which ranked second. Milk yield ranked highest at 49.8% with an index of 0.49 in cow selection followed by body size which ranked second at 34.2%; and an index of 0.45. Generally, the preferred traits corresponded to milk and/or meat production.

Table 3.8. Ranking and proportion of trait preferences for bulls and cows in

traditional husbandry practices

Bull selection			Cow selection		
Selection criteria	**%**	**Index**	**Selection criteria**	**%**	**Index**
Body size	52.26	0.35	Milk yield	49.79	0.49
Temperament	1.23	0.06	Body size	34.16	0.45
Yield of milk/meat	12.35	0.13	Mothering ability	5.35	0.12
Libido	0.41	0.02	Calving frequency	2.06	0.10
Pedigree	16.05	0.16	Udder size	2.47	0.06
Adaptability	1.23	0.03	Colour	0.41	0.05
Growth	1.65	0.08	Growth/survival	1.23	0.07
Do not select	14.40	0.15	• Do not select	1.65	0.01
• Age at first maturity	0.41	0.03	• Age at first maturity	2.88	0.03

Its noteworthy that in pure pastoral systems no selection was practiced but rather the animals are left to mate indiscriminately although castration is carried out to prevent high levels of inbreeding though the bigger proportion of bulls are those that are poor in growth performance.

3.4.3.2Likelihood preferences

Table 3.9 shows the odds ratio, traits preference and related statistical analysis. Most odds ratio analysis showed that there is a high probability for trait preference change with AEZ. This implies that each AEZ has its own selection criteria for bulls and cows. Although different factors affected selection priority, AEZs had a more (P <0.05) profound effect on farmers' first and second trait preference of cows, and similarly different effects (P <0.05) of AEZs on second trait preference for bulls. However, preference for third criterion on cows, and the first and third criterions on bulls had no differences (P <0.05) among AEZs. Milk yield selection criterion had a probability of 50% and cow size (37%) which is an indirect selection for milk yield, being selected as top most preferred traits in cows as did body size in bulls and milk yield (52 and 16% respectively).

Table 3.9. Odds ratio, wald score, coefficient and preferred probability of traits for

selection cows and bulls

Cow/bull	Priority	Selection criteria	Coefficient (±SE)	Wald (score)	Odds ratio	Probability (%)
Cow	1st	Milk yield	-0.28±0.19	6.24*	0.99	50
	2nd	Size	-0.40±0.16	5.15*	0.58	37
	3rd	Calf growth	-0.38±0.17	2.18NS	0.31	24
Bull	1st	Size	1.10±0.26	3.54NS	1.10	52
	2nd	Milk yield	0.30±0.16	17.63**	0.19	16
	3rd	Maturity	-0.05±0.22	0.05NS	0.18	15

[1] NS; not significant, * significant at 5% and ** significant at 1%

3.4.4 Constraints in traditional cattle production

The major constraints in traditional cattle husbandry were evaluated as whole for all cattle populations found in the three AEZs. The ranking of the constraints by the farmers are presented in Table 3.10. Out of the constraints identified as limiting the expansion of herds by producers, feed and water shortages were ranked the highest with an index values 0.41 and 0.22 respectively. Diseases were ranked third constraint in most of the areas studied however ranked first in Habero and Shaha sites where the two areas are known for frequent Rabies disease outbreaks. Predators were also identified as important constraints threatening cattle production in many areas.

Table 3.10. Proportion and ranking of constraints in the three AEZs (HL, WLL and ELL) in Eritrea

Constraints	AEZs			Grand mean	Index
	HL (%)	WLL (%)	ELL (%)		
Feed shortage	65.43	54.32	29.63	49.79	0.41
Drought (water shortage)	8.64	12.35	66.67	29.22	0.22
Disease	17.28	23.46	2.47	14.4	0.21
Labour shortage	2.47	3.70	0.00	2.06	0.05
Lack of extension	0.00	3.70	0.00	1.23	0.02
Lack of input	3.70	2.47	1.23	2.47	0.06
Predator	2.47	0.00	0.00	0.82	0.08

3.5 Discussion

3.4.1. Demographic and importance of cattle

Evaluation of the status of cattle production system is the fundamental step in cattle improvement program. Besides the livestock species, it is imperative that social and cultural practices are identified as social aspects such as the level of education of the producers, dominant gender in care/rearing of a specific species, age structure of the animal owners among others influence the adoption of practices aimed at increasing animal production and profitability. Analysis of the data obtained from this study indicated that farmers in the HL agro-ecological zone were younger and more educated compared to the WLL and ELL zones. In most of the latter two zones, nomadic pastoralism is mainly practiced in the cattle production implying that the producers are constantly in movement hence this hinder education acquisition. Women participation in cattle production was low which is similar what was found by Endashaw (2012) in Ethiopia which has similar cattle production practices as Eritrea. This differs from what was obtained in Tanzania, where women participation in cattle (Nguni breed) was at 21% (Obert, 2013). Most of the producers in the Tanzania case were young and educated which could explain the difference in the

gender participation in the production. Mwai *et al* (2015) observed that age and level of education have a significant impact on livestock production practices.

3.4.2. Cattle husbandry practices

The male to female sex ratio differed between cattle populations with pastoralists in WLL eliminating male calves at birth to manage inbreeding and increase cattle population. This led to only few males within the cattle populations. The reverse occurred in the HL zones where the birth of a male calf was preferred as it meant more source of farm power. Then oxen are primarily used for ploughing and threshing crops in the mixed crop-livestock production systems in the highlands. The average household herd size was found low than the 13.2 heads reported by Mekonnen (2012) in similar production environments in Ethiopia. This can be explained by the more frequent drought in Eritrea than in Ethiopia. Furthermore, the country is young having attained her independence in 1991. In Eritrea a large number of cattle are owned by pastoralists and agro-pastoralists. Their main purpose of keeping the cattle is production of milk for home consumption and sale of live animals to generate household income. The amounts of milk harvested are low due to constraints related to breeding, feed and general management since the pastoralists are averse to new livestock husbandry practices and technologies which can lead to increases in herd productivity and profitability. They mainly rely on their traditional knowledge and skills which has evolved over the long times they have reared livestock.

Findings from this study confirms the similarity amongst many traditional livestock keepers in Africa. For instance, in almost all the communities where livestock is extensively grazed use of traditional knowledge is widespread as does practices such as castration, dehorning, weaning among others (Chali, 2014). The objective of carrying out a certain practice may differ at times. For instance, most indigenous cattle producers in Tanzania castrated bulls to enhance draft power provision as well as reduce temperament of the males for easy handling (Msanga, 2012). This differ from the findings of this study where majority of the farmers castrated their animals

to control indiscriminate mating and enhancement of fattening. The latter with the reasons given in Chali (2014) for the Arsi cattle of Ethiopia.

It was evident from the results obtained in this study that culling was majorly based on age. This concurs with the findings presented in Mekonnen (2012) for the Horo cattle populations in Ethiopia. This is besides the observation that much of the pastoralism was characterised by communal land ownership which is similar to what is reported by Ftiwi and Tamir (2015), and Mekonnen (2012) for the Begait and Horo cattle populations respectively. Milk production varied with AEZs as well as the breed and its use in the farm system. The Arado breed which was mainly reared in the HL and ELL produced an average of 1.98 litres of milk per day. This is more than the Horo and Arsi breeds (1.5 litres/cow/day) in Ethiopia (Mekonnen *et al.* 2012; Chali, 2014) which are subjected to similar production environment as the Arado in Eritrea. There is more emphasis on the survival and adaptation traits when selecting the breed as it performs in harsh environment specially, feed compared to Barka which produced an average of 3.5 litres of milk daily for these mainly kept in crop-livestock production systems where selection is geared towards higher production.

Reproductive performance of indigenous cattle of Eritrea was generally good taking the average age at first mating as indicator of reproductive performance. Average age at first mating of cows of this study Arado and Barka cattle breeds is in agreement with the result of 3 - 4 years in Ethiopia on Arado and Arsi cattle populations (Mekonnen, 2013; Chali, 2014) but it is less than the research done by Kanai *et al* (2013) which is 4.7 years.

3.4.3. Trait preference and cattle naming

The findings of this study indicates that traditional cattle herders of different ethnic societies have different traditional knowledge in selection of cattle for breeding purpose. However, two clear objectives are detectable in the data obtained and are milk and meat production. Traits that influence either milk and/or meat production are preferred in selection. Its noteworthy that much of the production occurs in areas characterised by poor supply of quality nutrients and high temperatures. This affects

producers' choice of breeding stock as indicated by the high frequency with which producers mentioned the ability to produce and reproduce in the hostile environment (survival traits). Since there is no performance recording, selection is purely based on the phenotype preference. Using traditional knowledge, animals were mainly selected by the role being played (milk, meat or provision of draught). Farmers have very cautious not to lose the adaptive characteristics that made the breeds productive and reproduce in different challenging conditions. Pastoralists generally preferred the birth of female calves probably to ensure that their flock sizes increased overtime while crop-livestock production system preferred the birth of male calves as they are useful in provision of draft power in their crop farms. This does not concur with the findings presented in Chebo *et al* (2014) who indicated that lowland producers in southern Ethiopia preferred male calves for traction while the HL producers concentrated on survival traits. Most farmers in the current study selected bulls based on body size as their first choice and pedigree as their second choice. Cows were selected based on their milk yield performance and body size respectively. Selection based on calf growth is also another indirect measurement of milk yield. The findings from this study are in agreement with previous study by Ftiwi and Tamir (2015) and Endachaw (2012) who noted that consideration should be given towards the highly ranked traits in design of a breeding program for indigenous cattle breeds in the tropics.

There were differences in traditional cattle naming systems among producers in different AEZs. However, in many instances the naming was based on coat colour and horn type in different communities. Traditional cattle farming society have major constraints, feed shortage that ranked first, water shortage ranked as second and disease as third. However, in some areas predators were also ranked first just as vector for Rabies disease and killing animals. Similar result also found by Ftiwi and Tamir (2015) that feed shortage was considered as the most important problem which ranked first.

Most of the breeds used in the production system should be well adapted to the prevailing farming conditions. Cattle production practiced by the farmers have evolved over a long time aimed to match with the production system. Smaller animals are preferred in drier areas when compared to animals that are large framed

probably to match the feed supply or to adapt to the need to cover long grazing distances. In such environments as is the case in Eritrea, survival characteristics tend to be preferred to production traits (Goitom *et al.* 2016). It is important to analyse the low input dryland production systems and relate the information to the selection practices in establishment of indigenous livestock resources improvement program. Failure to match genetic resources to the production system leads to sub-optimal utilisation and conservation of the resources. However, it is important to consider local farmers breeding preferences in design of improvement plans irrespective whether the traits they consider have economic value. Failure to consider producers interest in breeding programme has been identified as the single most important factor affecting the success of such programmes (Kosgey *et al.* 2003).

3.6 Conclusions

This study confirmed the existence of three indigenous cattle production systems in Eritrea namely the crop-livestock, agro-pastoral and pastoral systems. The pastoralists were mainly found in lowlands while crop-livestock producers were in highlands. The result further indicates that much of Eritrea landmass is characterised by aridity thus favouring pastoralism. The Barka breed averaged high milk production and is mainly kept in the crop-livestock and agro-pastoral production systems where it can be bred with exotic breeds to increase its milk production. The Arado breed whose milk production is lower than Barka breed was mainly found in crop-livestock production system. Irrespective of the farming system, most farmers selected bulls based on body size criterion as their first choice, and cows based on milk yield performance. Traditional knowledge exists for cattle naming, feeding, housing and even selection. The two main selection objectives were milk and meat production, however, power provision and survivability are key to the livestock keepers. Diversity in traditional cattle management systems imply that there is opportunity for shared knowledge in indigenous cattle production.

CHAPTER FOUR

MORPHOLOGICAL CHARACTERIZATION OF INDIGENOUS CATTLE BREEDS IN ERITREA

4.1. Introduction

4.1.1. Background

Optimal utilisation of livestock biodiversity is crucial in ensuring sustainable food production, enhanced food security, higher household incomes and development of animal based agro-industry sector in Africa (Kugonza *et al.* 2011). Phenotypic characterization is important in design and implementation of efficient breed use and conservation programmes (Pilling *et al.* 2010; Kugonza *et al.* 2011). The morphological differences within and between cattle populations are attributable to the genetic makeup of individuals and environmental conditions in which they exist (Rege, 1999). This implies that performance of different breeds will vary in different environments (Pilling *et al.* 2010). The differences can be captured through morphological measurements (phenotypic characterization) which are then related to suitability of the selected individuals to perform in the specific environment. Phenotypic characterization is easy and feasible to implement and can be used for rapid selection of animals in the field (Dossa *et al.* 2007) making it suitable for resources constrained developing countries. It is however important to support the morphological classifications by molecular characterization techniques (Gizaw *et al.* 2011) as it is highly influenced by environmental effects as well as the strong genetic and environmental interactions.

In Eritrea, cattle production plays important socio-economic roles both at household and national levels (Goitom *et al.* 2016) with the grazed cattle sector dominating the livestock production industry. Much of the country's landmass is classified as ASAL and is therefore more suited to grazed animal production. Even in areas where crop production is practiced, mixed livestock-crop production system is common as animals are reared to utilise the crop residues to supplement grazing. According to

the Government of the State of Eritrea National Livestock development Project Report (NLDP, 2007), most of the country's animal production occur in the smallholder production systems which are mostly dominated by pastoral and agro-pastoral systems. The main livestock breeds kept are indigenous to the country and are therefore well adapted to the nutritional and environmental constraints experienced in the country's ASALs conditions. The indigenous cattle reared are characterised by low productivity forcing the government and other livestock sector stakeholders to invest in programmes to improve the performance (NLDP, 2007).

Designing effective genetic improvement programme requires adequate knowledge of the available genetic resources. Genetic diversity is the most important resource in livestock plans. Besides, communities have different sets of cultural and social values by which to judge, appraise and decide on breeding animal based on physical appearance (Zechner *et al.* 2001). The genetic diversity information of the ICPs in Eritrea is lacking making genetic improvement of animal populations a daunting task. This is despite the fact that cattle production is a very important component of the Eritrean economic growth and one which has the potential to increase food security in the country and the region in general.

In the sub-Saharan Africa, about 180 breeds of cattle have been recognized; among which 150 are indigenous to the continent (Rege, 1999; Rege *et al.* 2003). In the review of the African cattle breeds, Rege, (1999) grouped Eritrean cattle population into North Sudan Zebu comprising mainly of the Barka breed and Zenga comprising of the Arado breed. The Barka inhabit in the WLL AEZ of Eritrea, with a small population being found in the neighbouring country of Ethiopia. The breed is primarily kept for milk and meat production. The breed has potential to be improved for use in the ASAL areas of Africa. The Arado breed which is smaller in body size as compared to Barka is suited in mountainous regions and has been traditionally used in provision draught power (Genzebu *et al.* 2016) which is important in crop lands. Margaret (2002) indicated that Arado cattle is an intermediate between Zebu and Sanga breeds and noted the importance of the breed as a source of meat and milk.

In light of the climatic change being experienced and the anticipated changes in animal farm systems in response to the subsequent increase in climatic variability,

there is need for Africa to determine the performance of the breeds that exist in the continent with a view of conserving those that can withstand higher ambient temperatures as well as fluctuating pastures supply and quality. The Barka and Arado breeds are good candidates for use in mitigating the effects of climate change as they have been selected to produce and reproduce in the ASALs of Eritrea and parts of Ethiopia characterised by high ambient temperatures and low quality grazed pastures. In order to sustainably utilise and conserve these breeds, there is a need to determine its phenotypic and genetic diversity. Phenotypic categorisation is the first step in characterization of genetic resources (Delgado *et al.* 2001) for use in design of sustainable improvement programmes. The aim of this study was to characterize morphologically Eritrean cattle populations and to determine their relationship between and within these cattle populations. Eventually, the results of this study will be helpful in providing information for the improvement program and effective utilization of indigenous cattle resources.

4.2. Materials and methods

4.2.1. Study area

The study was conducted in Eritrea which located in East of Africa. Data was obtained from all the regions representing all the main Agro Ecological Zones (AEZs) namely western lowland (WLL), the highland (HL) and eastern lowland (ELL) region of Eritrea. A total of 243 unrelated cattle were randomly sampled for qualitative and quantitative phenotypic characters data necessary for identification and description of distinct populations, breeds or eco-types. The three ecological zones are found in the six administrative units (Zone) namely Anseba, Debub, Gash-Barka, Maekel, Northern Red Sea and Southern Red Sea that make up Eritrea. Each Zone contributed differently to the total samples collected depending on the size of the cattle population and agro ecological conditions. The lowest number of samples were collected from Southern Red Sea zone which had the least cattle population due to its desert climate conditions (see detailed information in Table 3.1). Samples of cattle populations were collected from a total of 27 sites spread out in the six Zone and which were Goluj, Awgaro, Tekreret, Tekombia, Cambo-10, Keru, Barentu,

Akordat, Shambuko, Serejeka, Galanefhi, Mai-alba, Bada, Sheha, Enghel-Eila, Menkanile, Emberemi, Shieb-Seleba, Afelba, Rekumbedin, Habero, Hamelmalo, Shieb-Mensheb, Shieb-Gedged, Foro, Cheguaro and Dongolo.

4.2.2. Data collection

A pilot survey was carried out prior to the actual characterization. This was to establish the distribution of cattle populations and composition in the study area as well to test the suitability of the questionnaire and other data collections tools to obtain all the information and samples necessary for the study. The data collection tools were adjusted following the pilot study for use in the main characterization. Morphological characterization in the study adopted the 13 body characteristics proposed by FAO (1986) with 7 morphometric measures, 5 physical traits and 1 coat colour characteristic as summarised in Table 4.1.

Table 4.1. Description of morphological characterization traits (FAO, 1986)

Measurement	Abbreviation	Description	
Morphometric			
Body length	BL	The distance between point of shoulder to the pin bone	
Heart girth (chest girth)	HG	Circumference of the body behind the base of the hump	
Height at withers	HW	The vertical distance from ground to the point of wither	
Horn length	HL	Base of the horn to tip of the horn	
Ear length	EL	Base of the ear to the pointed end	
Dewlap width	DW	The widest part of dewlap	
Tail length	TL	From the base of the tail to the pointed end of the tail	
Physical			
Horn type	HT	Horned, polled, loose	
Horn orientation	HO	Lateral, curved upward, curved to front	
Head profile	HP	Straight/flat, concave, markedly convex	
Ear formation	EF	Erect, semi-pendulous	
Hair characteristic	HC	Short and coarse, short and smooth	
Colour			
Coat colour		White, brown, red, black, mixed	CC

4.2.3. Data analysis

4.2.3.1. Descriptive statistics

Descriptive statistics were carried out to describe different morphological characteristics of cattle populations. Frequencies of each level of the qualitative data of the studied characteristics were computed using SPSS statistical software (SPSS,

2006). Chi-square analysis and student's *t*-test were used to determine the effect of the AEZ on morphological and physical traits among the studied populations.

4.2.3.2. Univariate analysis

Analysis of variance was used to assess the statistically significant differences between cattle populations and the corresponding AEZs. The mixed model of analysis of variance was used to analyse quantitative (morphological) data taking AEZs as fixed effect and animal populations as random effect. The statistical model that was used for the quantitative measurements was:

$$\gamma_{ijk} = \mu + \alpha_i + \beta_j + (\alpha \times \beta)_{ij} + \varepsilon_{ijk}$$

Where γ_{ijk} is an observation of a quantitative measurements on i^{th} agro-ecology zone and j^{th} cattle population; μ is the overall mean; α_i is i^{th} fixed effect of AEZ (i^{th} = one of the three ecological zones); β_j is random effect of j^{th} cattle population nested on AEZs (j^{th} = is one of nine populations); $(\alpha \times \beta)_{ij}$ is the interaction between the effect of AEZs and cattle populations; ε_{ijk} is random residual effect, where, N: = (0, 1).

4.2.3.3. Multivariate analysis

The data was first entered into Microsoft Excel for filtering and detection of inconsistencies before the analysis. The data was then exported to SPSS database system for the design of the structure of the variables for analysis. Principal Component Analysis (PCA) was carried out to visualize grouping and variation of the components. The first two principal components (vectors) were selected based on the rule of the highest proportion of variance (eigen value). This was important so as to detect the high variation by avoiding redundancy of characteristics. Rotation of principal components was done using VARIMAX to obtain high correlation within a single component besides to increasing difference between new principal components (Manly, 1994).

The classification was based on the 27 cattle populations studied and was aimed at forming distinct homogenous cattle populations. Cluster analysis was applied based on the model of nearest neighbour clustering using hierarchical single linkage

algorithm technique (Sneath and Sokal, 1973). Distance between two clusters was calculated based on the average distance among members within a single cluster to average distance of other cluster. Homogenous groups of cattle populations were formed based on their morphological characteristics resemblance. Grouping of similar animal populations were visualized by graphical hierarchical classification tree (dendrogram) and heuristics decisions were extracted to determine the number of clusters.

4.2.3.4. Discriminant analysis

Discriminant function analysis was carried out to certify whether the classification of cattle populations into clusters was correct. Furthermore, the analysis was used to determine which variable had more discriminating power in classification cattle populations. The percentage of correct allocation for each cattle group was calculated to determine how well populations were separated using the available variables. Stepwise discriminant function analysis was used to determine characteristics with the largest canonical coefficients implying a higher contribution of the discrimination between populations. The Mahalanobis square distance obtained through discriminant analysis was used to measure variability among cattle populations found in different AEZs. The Pearson correlation analysis was done between all morphometric and physical characteristics followed with a statistical significances test on the traits to determine the actual differences among the correlation values.

4.3. Results

4.3.1. Morphometric evaluations

The morphometric measurements of cattle populations in different AEZs in Eritrea are shown in Table 4.2. The WLL cattle (Barka) population recorded higher morphometric measurements, (BL; 126.8 ±0.7 cm), (HG; 156.4 ±0.7 cm), (HW; 125.3 ±0.6 cm) and (DW; 18.7 ±0.6 cm) than the HL cattle (Arado); (BL; 112.2 ±2.1 cm), (HG; 131.7 ±1.0 cm), (HW; 110.7 ±0.8 cm) and (DW; 11.4 ±1.1 cm) and the ELL cattle (Arebo), (BL; 111.3 ±2.2 cm), (HG; 137.6 ±0.8 cm), (HW; 108.1 ±0.8 cm) and (DW; 14.3 ±0.4 cm). Significant differences in morphometric measurements

are found among cattle populations found in different AEZs (P <0.001). Body length in Arado HL cattle was found to be shortest (107.3 cm) in Mai-alba cattle population and the tallest (130.3 cm) recorded in Keru cattle population in North of WLL (commonly known as dowhin eco-type). Height at wither was found to be short in HL (Arado) and in ELL (Arebo) with 110.7 cm and 108.1 cm respectively (Table 4.2). Arado cattle were significantly shorter than cattle populations in WLL, mainly the Barka cattle type (125.3 cm) (P <0.001).

Heart (Chest) girth is another important quantitative measurement in cattle description, and also important for the prediction body weight of an animal. Small value implies light animal, and the smallest (127.6 cm) observed in Hamelmalo cattle population (North part of HL AEZ). The highest chest girth value (158.6 cm) which is heavy animal is found in Augaro cattle population which is Barka breed specifically in Begait eco-type which is found in South of the WLL. However, mean values are 134.4 cm for Arado breed and 156.4 cm for Barka breed.

Dewlap is folded skin that extends its length from the base of head up to thoracic area, and maximum width is found near middle of the total length. Dewlap width has a direct indicator for a specific cattle type. Cattle population (Barka breed) found in WLL specifically, in South-WLL around Gash River had a significant (P <0.001) wide dewlap than other cattle populations (Table 4.2).

Table 4.2. Morphometric measurements (cm) of cattle populations in different AEZs

Population	**BL**	**HG**	**HW**	**DW**
HL (Arado)	112.2 ±2.1	131.7 ±1.0	110.7 ±0.8	11.4 ±1.1
WLL (Barka)	126.8 ±0.7	156.4 ±0.7	125.3 ±0.6	18.7 ±0.6
ELL (Arebo)	111.3 ±2.2	137.6 ±0.8	108.1 ±0.8	14.3 ±0.4
Mean	116.77 ±1.7	141.90 ±0.8	114.70 ±0.9	14.80 ±0.7
P-value	P <0.001			

4.3.2. Physical characteristics

Results for the coat colour, head profile, horn type and horn orientation are presented in Table 4.3. Red coat dominated the cattle populations found in the HL and ELL regions at 54.3% while white and grey coat coloured cattle types were least at 8.7%. The Barka breed was dominated by mixed coloured (dominantly black and white) cattle at 39.5% followed closely by individuals coloured black (gray) at 34.6%. White coloured individuals were least at 1.2%. Approximately 6.2% of the WLL cattle had brown coat colour while there was no in the HL and ELL cattle populations. Besides, only 2.5% of the WLL cattle had red coat compared to 54.3% in the HL and ELL regions.

Most (90.1%) of the Arado cattle had a straight/flat head profile. There was no single individual with a markedly convex head/face profile among the cattle populations in the HL and ELL regions but there was 1.2% in WLL. Straight/flat head accounted for 65.4% of the sampled cattle in the WLL with the concave shaped individuals representing a 33.3%. 92.6% of the HL and ELL cattle populations (Arado) were horned while only 2.5% were polled and similarly, the WLL cattle (Barka) had 77.8% and 6.2%. Most cattle had upward curved horns orientation at 74.1% for the HL and ELL regions and 84% for the WLL cattle (Table 4.3).

Table 4.3. Frequency and proportion of coat colour, head profile, horn type and horn orientation of cattle in different AEZs in Eritrea

	HL and ELL (East Coast) (Arado)		WLL (Barka)	
	Frequency	**%**	**Frequency**	**%**
Coat colour				
White	7	8.7	1	1.2
Brown	-	-	5	6.2
Red	44	54.3	2	2.5
Black	10	12.3	28	34.6
Grey	7	8.7	13	16.0
Mixed (black & white)	13	16.0	32	39.5
Head profile				
Straight/flat	73	90.1	53	65.4
Concave	8	9.9	27	33.3

Markedly convex	-	-	1	1.3
Horn type				
Horned	75	92.6	63	77.8
Polled	2	2.5	5	6.2
Loose	4	4.9	13	16.0
Horn orientation				
Lateral	5	6.2	-	-
Curved upward	60	74.1	68	84.0
Curved to front	16	19.7	13	16.0
P-value	***P* <0.01**		***P* <0.01**	

Table 4.4 presents frequency and proportion of ear formation, ear length, hair type and tail length of different cattle populations in different AEZs. A total of 95.1% of the Arado cattle types had erect ear while 72.8% for the Barka cattle type different ear type (P <0.05). No individual was sampled with long ears type amongst the HL and ELL cattle populations and similarly short ear in WLL. The most common hair type was short and smooth at 88.9% and 90.1% for the HL and ELL regions, and WLL cattle populations respectively. Similarly, most sampled cattle had long tail at 84% for the HL and ELL and 74.1% for the WLL cattle types.

Table 4.4. Frequency and proportion of ear formation, ear length, hair type and tail length of cattle populations in different AEZs

	HL and ELL (Arado)		WLL (Barka)	
	Frequency	**%**	**Frequency**	**%**
Ear formation				
Erect	77	95.1	59	72.8
Semi-pendulous	4	4.9	22	27.2
Ear length				
Long	-	-	38	46.9
Medium	78	96.3	43	53.1
Short	3	3.7	-	-
Hair characteristic				
Short and course	9	11.1	8	9.9
Short and smooth	72	88.9	73	90.1
Tail length				
Long	68	84.0	60	74.1
Medium	13	16.0	20	24.7
P-value	***P* <0.01**		***P* <0.01**	

Table 4.5 presents the Pearson's correlation coefficients of the physical and morphometric characteristics of Arado and Barka breeds. Significant (P <0.05) positive correlations were found among HW, HG and BL for the Barka breed but not for Arado breed. Dewlab width of Arado breed has significant (P <0.05) positive correlations with BL and HG. However, coat colour was inversely related with most morphometric characteristics especially with HG and DW.

Table 4.5. Pearson's correlation coefficients of the physical and morphometric characteristics of Arado and Barka breeds[1]

		Arado								
		BL	HG	HW	CC	HP	DW	TL	HT	EL
Barka	BL		.148	.140	.050	-.046	.279**	-.137	-.005	.108
	HG	.173		.164	-.104	.094	.312**	-.013	.130	-.103
	HW	.371**	.276*		-.078	-.079	.007	-.042	.064	.114
	CC	-.014	-.276*	-.026		.124	-.229**	-.099	.025	-.078
	HP	.086	-.128	.036	-.131		-.085	.027	.009	-.062
	DW	.043	.067	-.138	-.219	-.013		.211**	-.056	-.042
	TL	.060	-.080	.105	.033	.024	.024		-.135	-.111
	HT	-.038	-.023	-.043	.011	.102	-.037	.044		.079
	EL	.135	-.049	-.084	-.062	-.119	-.059	-.013	.136	

*Correlation is significant at the 0.05 level and ** correlation is significant at the 0.01 level (2-tailed); [1]See Table 4.1 for the description of the physical and morphometric characteristics of the Arado and Barka breeds.

4.3.3. Classification of indigenous cattle types

Table 4.6 shows the latent vectors and proportion of variation of the first two principal components obtained using the multivariate techniques (PCA and cluster analysis) based on different characteristics in the cattle classification. The first two principal components accounted for 77.7% of the total variance. The first principal component focused on the contrasting between high negative value of morphological measurements (BL and HG) with the low negative values of physical characteristics (CC, HP, HO, HT and HC). The second principal component indicated the contrast between the high positive value of HG (0.63) and the high negative value of BL (-0.77).

Table 4.6. Latent Vectors and proportion of variation of the first two principal components

	PC1	PC2
Proportion of variation	74.48	3.20
Characteristics[1]		
BL	-0.515	-0.767
HG	-0.704	0.633
CC	-0.032	-0.008
DW	-0.175	-0.097
EF	-0.013	-0.005
EL	-0.016	-0.005
HC	0.001	0.004
HP	-0.007	-0.002
HW	-0.454	-0.102
HO	-0.026	0.006
HT	-0.005	0.001
TL	-0.034	0.011

[1]See table 4 1 for the description of the characteristics

Classification of indigenous cattle types was done based on hierarchical single linkage algorithm technique and produced two clusters (Figure 4.1). The first cluster had cattle populations from Goluj, Awgaro, Tekreret, Tekombia, Cambo-10, Keru, Barentu, Akordat and Shambuko populations (all are Barka) while the second cluster had populations from Serejeka, Galanefhi, Mai-alba, Bada, Sheha, Enghel-Eila, Menkanile, Emberemi, Shieb-Seleba, Afelba, Rekumbedin, Habero, Hamelmalo, Shieb-Mensheb, Shieb-Gedged, Foro, Cheguaro and Dongolo (all are Arado). The Shambuko sub-population formed as sub-cluster within the first cluster while Dongolo formed a sub cluster in the second cluster. Figure 4.2 shows the photos of the two major clusters (breeds).

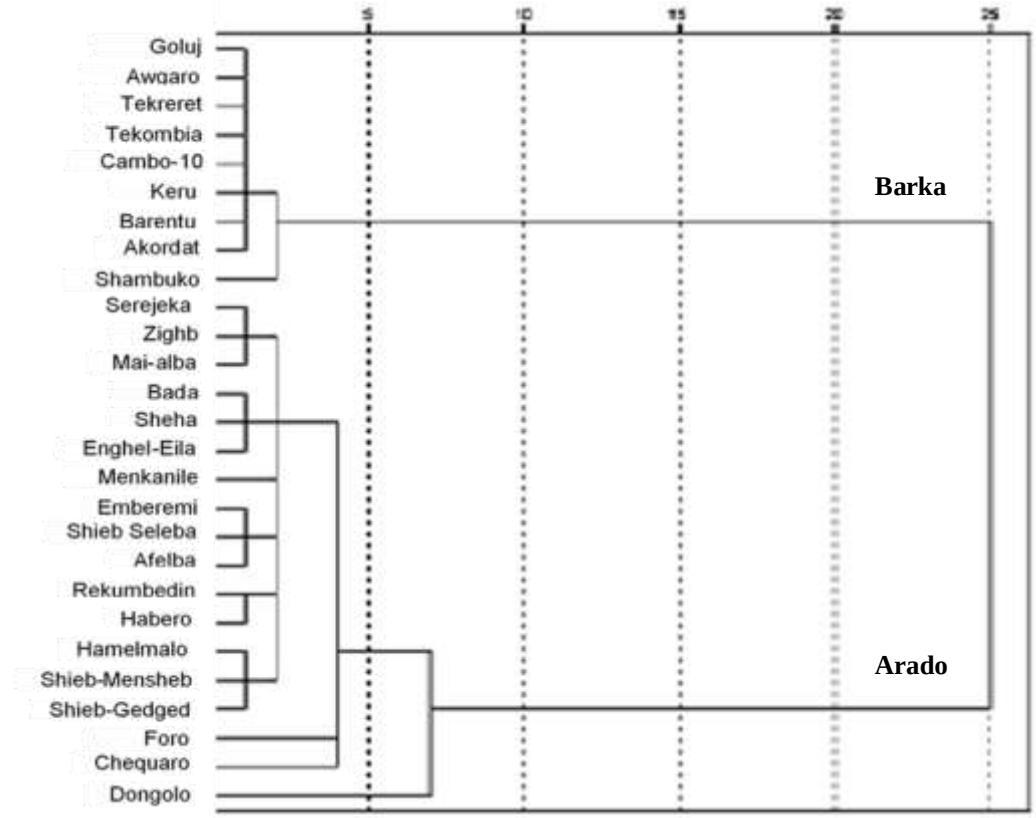


Figure 4.1. Dendrogram using single linkage in classification cattle populations

Figure 4.2. Photos of the two cattle groups identified in the analysis. Source: (own photo)

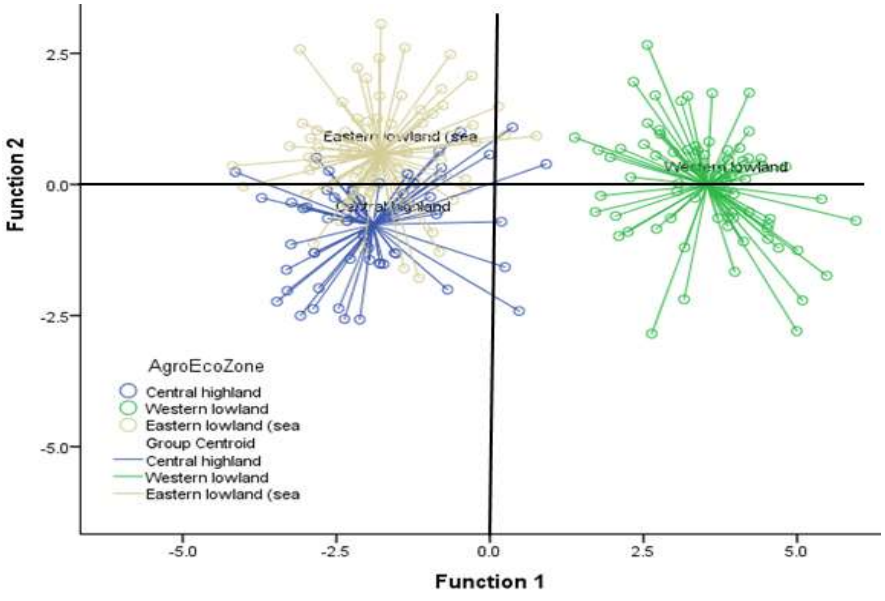
Results from the canonical discriminant analysis are presented in Table 4.7. The first function accounted for 95.7% of the total variance. Further results from stepwise discriminant analysis which were used to check the discriminating powers of the variables within a function showed that out of the 7 variables considered, only three had significant (P <0.001) discriminating power. These were HW, CC and TL had the best discriminating power in the first canonical function. Similarly, the TL and DW had good discrimination power when used in the classification populations by the second function.

Table 4.7. Standardized canonical discriminant function coefficients

Characteristics[1]	First function	Second function
HG	0.347	0.217
HW	0.511*	-0.545
CC	0.418*	-0.082
HP	0.228	-0.172
DW	0.271	0.563*
TL	0.464*	0.437*
EL	0.416	-0.360

[1]See Table 4.1 for the description of the characteristics; *Characteristics with discriminating power

Result of classification of cattle populations found in different AEZs based on the first two canonical discriminant functions are presented in Figure 4.3. The first canonical discriminant function separated the cattle populations into two cattle



breeds (Barka and Arado).

Figure 4.3. Result of classification of cattle populations found in different AEZs based on the first two canonical discriminant functions

Table 4.8 presents the results for the Mahalanobis distance between WLL cattle populations in one group and the HL and ELL cattle types in the second group.

Table 4.8. Mahalanobis distance and proportion of variations of the two functions for WLL and HL and ELL of cattle populations

Functions (Groups)	Wilks' Lambda	% of variance	P-value
1st function (WLL Vs HL and ELL)	0.101	95.7	*P* <.001
2nd function (HL Vs ELL)	0.769	4.3	*P* <.001

The squared distances between standardized classes means allows for pairwise comparisons of morphological characteristics between the different populations. The distances between HL and ELL cattle populations was the longest distance at 5.1 as presented in Table 4.9.

Table 4.9. The squared Mahalanobis distances between HL, WLL and ELL cattle populations

AEZ	HL	WL	EL
HL	**		
WLL	1.93	**	
ELL	5.10	4.72	**

Table 4.10 presents results of discriminant analysis based on the number of observations and percentages of correct classification in brackets of the cattle breeds. Analysis of correctness of the breeds/AEZs allocation were 100% for the Barka (WLL), 77% for the Arado (HL) and 77.8% for the Arado (ELL). Generally, 84.9% of the original grouping of the cattle populations into breeds/ecotypes were correctly allocated.

Table 4.10. Classification results of discriminant analysis by the number of observations and percent correctly classified (bracket) in different breeds and ecotypes

Agro-ecological zone	HL	WLL	ELL
Arado (HL)	47 (77.0%)	1(1.6%)	13 (21.3%)
Barka (WLL)	0 (0.0%)	74 (100.0%)	0 (0.0%)
Arebo (ELL)	18 (22.2%)	0 (0.0%)	63 (77.8%)

4.4. Discussion

Results from this study indicate that there exist differences in the physical and morphometric measurements among, between and within the indigenous cattle breeds reared in Eritrea. Yakubu *et al* (2009) in Nigeria and Mwacharo *et al* (2006) noted that the differences in morphometric measurements are usable in categorising between and within breed variation. The differences were used to morphologically characterise the cattle breed in Eritrea where major cattle types were identified by clustering system.

4.4.1 Physical and morphometric measurements

The Barka breed had higher values for heart girth and height at withers which correspond to its larger than Arado breed stature. The HG and HW have been used to determine the live-weight of cattle in Senegal (Tebug *et al.* 2016). Moreover, the HG has been identified as the most reliable body measurement parameter to predict body weight (Lesosky *et al.* 2012; Lukuyu *et al.* 2016). In areas where feed resource is not a constraint in Eritrea it would be advisable for producers to keep Barka breed to increase the meat productivity per unit area of grazing land. Furthermore, the breed can be crossed with the smaller Arado breed to take advantage of the higher live-weight. The lower values of the HG in Arado breeds can be explained by the low quantity/quality feeds found in the areas where the breed is found. Long height of the Barka breed is associated with its long length and huge body conformation. The range of the HW values in the Arado breeds in this study are comparable to those

obtained for Nandi and Mongalla indigenous cattle breeds in Nigeria (Abdulmojeed, 2010; 111.84 cm). Similarly, Barka breed has consistency with white Fulani (Kanai *et al.* 2013; 151cm) in Nigeria, Arado breed in Ethiopia (Dessalegn 2012; 138.0 cm) and Baggara breed in Sudan (Alsiddig *et al.* 2010; 132.35 cm). The production system from which these results were obtained are similar implying that the environment could played an important part in dictating the morphology of the breeds.

Most of the cattle populations in Eritrea had short and smooth hair type. The hair type was expected as an adaptation to high temperatures as much of the country is characterised by high ambient temperatures. Dry areas are also characterised by high number of biting insects which corresponds to the medium to long tail length observed in most of the cattle populations studied. The long tails are an adaptation to protect the cattle from the high external insect infestation.

Coat colour has been used to identify cattle types by farmers and has been associated with thermal regulation selection pressure (Finch *et al.* 1984) with dark coated animals absorbing more heat from solar radiation than light coated contemporaries (Seo *et al.* 2007; Desta *et al.* 2011). Results obtained from the study indicated that the Arado breed was mainly reddish compared to the black and white observed in Barka. This matches the characteristic ambient temperatures in the environments which each of the breeds dominate.

4.4.2. Morphological classifications

Analysis of the morphological characteristics data obtained in this study indicate that there are two major groups (Barka and Arado) accounting for the cattle populations in Eritrea which support the classification according to their historical origin. Where, the Barka mainly includes cattle types from the WLL while the Arado is majorly composed of cattle breed type found in the HL and ELL regions. The morphological characterization uses taxonomic units (Sneath and Sokal, 1973) in which the aggregation is based on the similarity relationships. However, there were sub-population within the main group (Barka cluster) as evidenced by the cattle population found in Shambuko area commonly called as Begait cattle. This is

possibly due to speciation from a common ancestor of Barka breed. Besides, Shambuko area has higher rainfall implying the cattle populations have access to better quality and sufficient feed resource. The Dongolo cattle population from the Arado cluster also formed a sub-population probably due to similar reason since the area they are mainly kept is characterised by having two rainy seasons resulting to higher supply of forage compared to the other areas in the HL and ELL regions where the breed is found.

Results from the PCA indicated that the first two principal components accounted for about 77.7% of the variation. Specifically, the PC1 (74.48%) indicated that there was no relationship between morphometric measures and physical attributes such as the coat colour. Furthermore, there existed high positive morphological measurements (HG, BL and HW) compared to low other measurements. However, the contrasting relationships of HG at 0.63 and BL at -0.77 in the PC2 showed that a high value for HG did not necessarily imply a high value for BL.

Findings from the discriminant analysis confirmed that there were no animals from the Barka breed that were clustered together with the Arado breed. It also resulted in perfect (100%; see Table 10) allocation of the Barka cattle to Barka breed which implies that the breed is clearly distinguishable based on its morphological characteristics. Moreover, the Arado breed found in the HL and the ELL regions had similar correct allocations at 77% and 77.8% respectively indicating that they have common morphological features.

4.5. Conclusions

There exists between and within breeds variation amongst cattle populations reared in Eritrea. Morphological and physical characteristics cluster all the cattle populations in two groups mainly the large bodied WLL Barka breed and the smaller statured Arado breed found in the HL and ELL regions. Based on the finding from this research, cattle genetic improvement programme based on within and/or between breed types would be feasible for the ICPs in Eritrea. This would be important in that the genetic diversity of Eritrean indigenous cattle would be maintained. The larger bodied Barka cattle could be used as the sire breed to improve meat production

proving relatively good managements. The results in this study further indicate that there has been little loss of genetic diversity in Eritrean ICPs. There is need to back up the findings of this study with a molecular characterization of the cattle breeds of Eritrea.

CHAPTER FIVE

GENETIC DIVERSITY POPULATION STRUCTURE AND ADMIXTURE ANALYSIS IN ERITREAN INDIGENOUS CATTLE

5.1. Introduction

5.1.1. Background

The concept of the breed was developed during the eighteenth century and is meant to describe the differences (morphological and genetic) between populations. A wide range of the breeds represents unique sets of genetic diversity (Dackson, 2008). The diversity of cattle breeds mainly resulted from their coexistence with humans as a result of artificial selection for different functions which has led to the high genetic diversity observed in different cattle breeds and/or populations (Laercio**,** 2013). Cattle may also have genetically diverged as a result of natural selection.

Indigenous cattle populations (ICPs) in Eritrea are widely distributed throughout diversified geographical and Agro-Ecological Zones (AEZs). Geographically isolated ICPs are subjected to local climatic conditions and each region may host some unique characteristics that can help them to survive and reproduce in the harsh environments. Cattle around Red Sea coast are one example that have ability to resist very hot climate (above 40°C) and tolerate high salty resources in the environment. However, there is no research work or records available to show genetic uniqueness of these cattle populations. Cluster analysis based on single linkage agglomerative hierarchical and non-overlapping (SAHN) technique using morphological features clustered Eritrean ICPs into two groups; Arado (Highland (HL) and Eastern Lowland (ELL) populations), and Barka (Western Lowland (WLL) populations) (Goitom *et al.* 2016). However, Parker *et al* (1998) recognized the need to confirm the phenotypic distinction through molecular characterization as the latter eliminates the effects of the selection pressure. Characterization based on molecular markers provide large unbiased estimates of similarities or differences of cattle populations. Markers have been comprehensively exploited to access genetic variability as they

contribute information on every region of the genome. Specifically, the use of single-nucleotide polymorphisms (SNP) markers is one of the most powerful means of studying genetic diversity and admixture (Putman and Carbone, 2014). High-throughput sequencing technologies have improved the power to discover and genotype SNPs. Genotyping-by-Sequencing (GBS) is the most cost-effective (De Donato *et al.* 2013) and powerful genomic method for genetic diversity analysis due to its ability to identify genomic variations throughout the genome (Baral *et al.* 2018). However, utilization of high-throughput sequencing technologies using high density markers (SNP) to study the genetic variations and population structure of Eritrean ICPs has not been done. Therefore, the current study was aimed at assessing the genetic variations, population structure and admixture of Eritrean indigenous cattle populations from three distinct and diverse AEZs. Besides, the study was meant to confirm the morphological cattle categories of Eritrea.

5.2. Materials and methods

5.2.1. Study area

The study was conducted in Eritrea which located in North eastern of Africa. It is bordered by the Red Sea to the East, Sudan on the West, Ethiopia to the South and Djibouti on its South-eastern side.

5.2.2. Data collection

The data used in this study was collected from different sub-regions (27) located within the three AEZs (WLL, HL and ELL) in the country (Table 5.1). Blood samples were collected from a total of 243 mature cows. The sampling criteria was based on the guidelines presented in FAO, (2015) report which indicated that at least 40 animals are enough to characterize a breed. This is in addition to the 60 km geographical distances between the sub-regions within the AEZs. The names of cattle populations were assigned based on the names of the sub-regions (refer to Table 3.1 and Table 5.1).

Table 5.1. The AEZs, sub-regions and their respective abbreviations

AEZs	Populations	Code
	Shambuko	SHE
	Cambo-10	CAM
	Barentu	BAR
	Awgaro	AW
WLL	Goluj	GOL
	Tekombia	TEKO
	Keru	KER
	Tekreret	TEK
	Akordet	CF
	Mai-Alba	MA
	Galanefhi	GAL
	Sheha	SHEH
	Afelba	AFE
HL	Senafe	SEN
	Serejeka	SER
	Habero	HAB
	Shieb-Seleba	SHES
	Hamelmalo	HAM
	Shieb-Mensheb	MENS
	Dongolo	DOGO
	Shieb-Gedged	GHE
	Emberemi	EMB
ELL	Foro	FOR
	Enghel-Eila	ENG
	Menkaneli	MENK
	Rekumbedin	REK
	Bada	BADA

5.2.3. Data analysis

5.2.3.1. Blood sampling and DNA extraction

The animal was restrained and 5 ml blood sample was collected using 10 ml vacutainer coated with EDTA from the jugular vein as shown in Figure 5.1. Preservation of the whole blood samples was performed in ice box in field and in deep freezer set at -20ºC. Proper packaging and shipment of the whole blood were done for DNA extraction in Kenya. The DNA was extracted following the standard protocol described by Sambrook (1998) which involves the use of proteinase K digestion and phenol-chloroform extraction procedures in Kenya.

Figure 5.1. Blood sampling from Jaguar vein using vacutainer in Bada (A) and Tekombia (B) sites. Source: (own photo)

Quantity and quality of the extracted genomic DNA were determined using gel electrophoresis to visualized quality of bands produced. The Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, USA) using optical density was also used to determine the amount and concentration of the DNA (Figure 5.2). This was done to ensure that the samples with high DNA concentration (>50 ng/µl) and quality (>1.8 - 2.0 $A_{260/280}$). For each sample, 30 ng/µl of DNA was diluted using double distilled water and stored at 4°C before being shipped to Beijing

Genomic Institute (BGI) laboratory for sequencing using Genotyping-by-Sequencing (GBS) approach. The GBS protocol follows the procedure described by Elshire *et al.* (2011). The procedure includes the digestion of genomic DNA with restriction enzymes, ligation of barcode adapter, PCR amplification and sequencing on a single lane of flow cells.

Figure 5.2. DNA extraction activity (A) and quality determination using Nanodrop (B). Source: (own photo)

5.2.3.2. Sequencing

Libraries were created for all samples with unique barcodes. A 96 multiplexed libraries (including controls) per lane were prepare by BGI for GBS using Illumina Hiseq 4000 PE 101 machine. After sequencing, demultiplexing of the barcodes was done by the sequencing company (BGI) and demultiplex barcoded reads were separated into individual files.

5.2.3.3. Data validation and filtering

After the demultiplexing, raw FastQ data were subjected to quality control checks. QualityTrim software (Robinson, 2015) was used to trim sequences with a minimum quality set at 20, the minimum length of 50 bp and maximum poor bases as well as N bases at 3. SAMtools v.1.3.1 (Li *et al.* 2009) was used to convert SAM files to BAM format before alignment. Further quality by removing duplicates, sorting, indexing and merging procedures were also done on the sequences. The resultant quality sequence reads were aligned against *Bos taurus* reference genome (UMD_3.1.1/BosTau8) using Borrow Wheel Aligner (BWA) (Li *et al.* 2009).

Variant calling was done by invoking SAMtools Mpile-up function with the called variants being stored in the Variant Call Format (VCF) files using BCFtools. Many filtering procedures to reduce false-positive SNPs calls were conducted before using for further analyses. Sex chromosomes, insertion and deletion (InDel) and SNPs with quality below 20 were filtered using BCFtools v.0.1.15 (Danecek *et al.* 2011). Additional quality control for SNPs was done by removing SNPs with calling rate below 98%, MAF above 5%, Linkage Disequilibrium (LD) pruning (50 5 0.2) and HWE (P >0.001) was carried out using PLINK v1.07 software (Purcell *et al.* 2007).

5.2.4. Diversity analysis

Subsequently, the remainder of the autosomal SNPs were used in making inferences to genetic diversity and relationships among the cattle populations. The PGDSpider Software version 2.1.1.3 (Lischer and Exoffier, 2012) was used to convert PLINK PED and MAP files to Arlequin software format for the population differentiation analysis. The TASSEL 5.2.43 Software (https://bitbucket.org) was used for determination of Nei's standard genetic distance (Nei, 1972) between pairs of cattle populations. Distance measure to assess the divergence among cattle populations was done using allele frequencies (F_{ST}). F_{ST} fixation indices were calculated using Arlequin Software version 3.5 (Excoffier *et al.* 2005).

Genetic parameters A_F, heterozygosity and inbreeding coefficient were calculated to compare the levels of heterogeneity between and within cattle populations using PLINK Software version 1.07 (Purcell *et al.* 2007).

5.2.5. Structure analysis

Cluster analysis using the unweighted pair group method with arithmetic mean (UPGMA) was used to determine population relationships (Sneath and Sokal, 1973). Visualization of population relationships was done by cluster analysis using the UPGMA. Patterns of population classification and genetic structures were assessed using Principal Component Analysis (PCA) in TASSEL 5.2.43 Software (https://bitbucket.org/) and admixture analysis using ADMIXTURE 1.2.3 Software (Alexander *et al.* 2009). In making inferences on the number of clusters (K), one to six clusters were assumed and optimum number of K was determined using the lowest cross-validation error to indicate the best representative clusters. The Analysis of Molecular Variance (AMOVA) was conducted to test the significance of variability between and within cattle populations where cattle populations were nested on AEZs.

5.3. Results

5.3.1. SNP identification and characterization

A total of 109.6 million reads were produced from the qualified 188 sampled animals. Numerous quality control checks were done to ensure the quality of SNPs. Insertion, deletions and SNPs on the sex chromosomes were filtered. The SNPs were further filtered for quality (>20), MAF (<0.05), HWE (*P* >0.001) and LD-based pruning (50 5 0.2). Mean of 0.016 million of autosomal SNPs were kept after being filtered, and were used for diversity and structure analysis. The number of quality variants within autosomal chromosomes varied greatly among the population studied, ranging from 0.06 to 0.18 million (Table 5.2).

Table 5.2. Number of SNP in millions for all the cattle populations

Population	Total No. of reads	Filtered	Population	Total reads	Filtered
AFE	4.44	0.09	HAM	5.19	0.17
AW	3.71	0.11	KER	4.10	0.11
BAD	4.55	0.12	MA	4.64	0.11
BAR	4.13	0.12	MENK	4.64	0.11
CAM	5.19	0.15	MENS	4.64	0.12
CF	4.74	0.14	REK	0.95	0.10
DOGO	4.11	0.11	SER	4.64	0.11
EMB	3.84	0.06	SEN	1.03	0.07
ENG	3.63	0.10	SHAH	4.01	0.08
FOR	3.08	0.08	SHE	4.71	0.16
GAL	3.67	0.08	SHES	4.05	0.14
GHE	4.09	0.11	TEKO	4.40	0.12
GOL	5.66	0.18	TEKR	3.88	0.12
HAB	3.88	0.10	-	-	-
Total				**109.61**	**3.08**
Mean				**0.058**	**0.016**

5.3.2. Genetic diversity of cattle populations

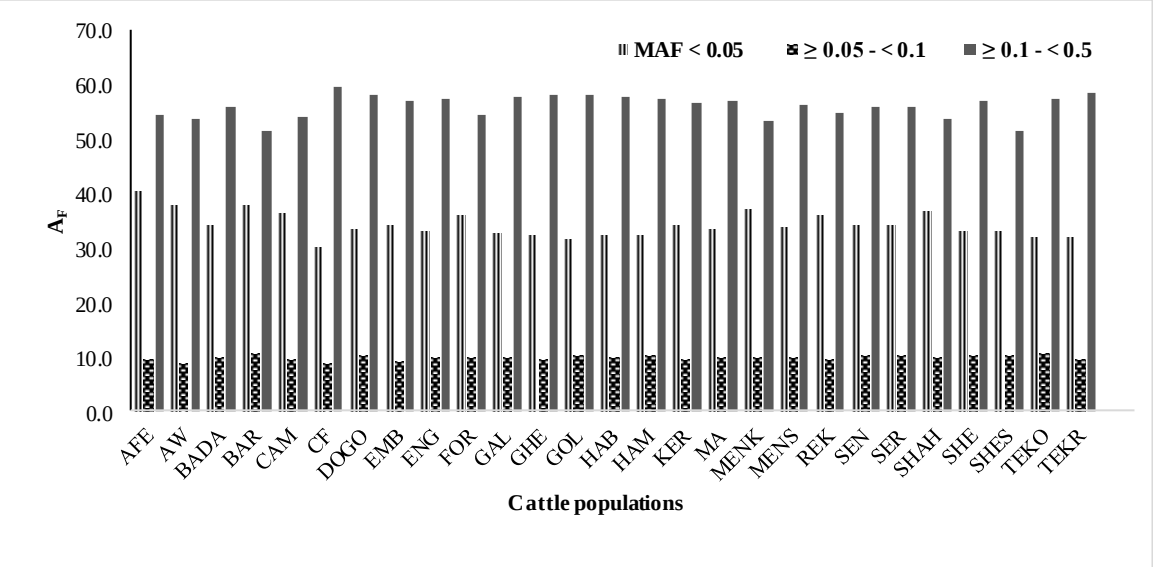
Genetic diversity indicators including the A_F, H_O, H_E and F_{IS} across all the loci were estimated and are presented in Table 5.3. The H_E per population ranged from 0.190 to 0.343 with the lowest value being recorded in FOR cattle population and the highest in BAR and AW cattle populations. The H_O ranged from 0.224 (AFE population) to 0.316 (AW and BAR populations). Slight deviation of 0.037 was detected between the means H_O and H_E. The mean F_{IS} for the studied populations was -0.089, but ranged from -0.175 in MENK population to 0.077 in AW population. An average of 0.157 A_F was obtained from an analysis with the AW and BADA cattle population have the highest A_F values.

Table 5.3. Heterozygosity, inbreeding and A_F (≥0.05) values for all the cattle populations

Cattle population	Heterozygosity		Inbreeding (F_{IS})	A_F (≥0.05)
	H_O	H_E		
AFE	0.224	0.209	-0.037	0.147
AW	0.316	0.343	0.077	0.227
BADA	0.299	0.321	0.070	0.227
BAR	0.316	0.343	-0.008	0.154
CAM	0.225	0.220	-0.005	0.159
CF	0.241	0.230	-0.010	0.164
DOGO	0.230	0.218	-0.044	0.154
EMB	0.230	0.215	-0.028	0.151
ENG	0.247	0.189	-0.137	0.146
FOR	0.247	0.190	-0.162	0.142
GAL	0.258	0.199	-0.137	0.149
GHE	0.247	0.190	-0.137	0.149
GOL	0.261	0.207	-0.122	0.154
HAB	0.257	0.200	-0.135	0.149
HAM	0.258	0.205	-0.121	0.153
KER	0.249	0.196	-0.123	0.146
MA	0.250	0.196	-0.127	0.145
MENK	0.251	0.192	-0.175	0.145
MENS	0.256	0.198	-0.148	0.145
REK	0.256	0.219	-0.089	0.157
SER	0.255	0.200	-0.117	0.150
SEN	0.256	0.219	-0.089	0.157

SHAH	0.244	0.198	-0.101	0.146
SHE	0.257	0.203	-0.126	0.157
SHES	0.255	0.201	-0.128	0.157
TEKO	0.261	0.206	-0.125	0.154
TEKR	0.255	0.201	-0.119	0.149
Grand mean	**0.255 ±0.01**	**0.218 ±0.01**	**-0.089 ±0.04**	**0.157 ±0.04**

The MAF was categorized into three groups; <0.05, ≥0.05 - 0.1 and ≥0.1 to <0.5. Mean of these proportions were 34.2%, 9.7% and 56.1% respectively (Figure 5.3). Considering monomorphic SNPs (<0.05), the lowest proportion (30.0%) was observed in the CF population. Similarly, considering polymorphic SNPs, the highest



polymorphic loci (59.7%) was observed in CF cattle population.

Figure 5.3. Distribution of A_F (%) in ICPs of Eritrea

Genetic distances were calculated using pairwise Nei standard genetic distances (Ds) to determine the genetic relatedness among the cattle populations. The genetic distances ranged from 0.32 to 0.38 between the cattle populations (Table 5.4). The lowest genetic distance (0.32) was recorded between HAB and GOL, HAM and KER, and TEKR and REK cattle populations. Conversely, the highest distance (0.38) was observed between CF versus AFE, BADA and BAR, and SEN versus HAM and HAB cattle populations. Mean pairwise F_{ST} values of all populations is 0.028 which is high genetic differentiation. The F_{ST} value between the populations ranged from 0.00 to 0.30 with the highest being recorded between DOGO and GAL cattle

populations (Table 5.4). The lowest F_{ST} value was obtained between AW and HAB, and CAM and SHE cattle population.

Table 5.4. Nei's genetic distances (Ds) below and population differentiation (F_{ST}) above diagonal based on 1autosomal SNPs of cattle populations[1]

REC	0.36		0.04	0.09	0.11	0.04	0.06	0.08	0.07	0.05	0.04	0.19	0.08	0.06	0.01	0.07	0.02	0.09	0.02	0.01	0.03	0.02	0.10	0.03	0.03	0.01	0.10
TEKR	0.37	0.32		0.14	0.01	0.01	0.06	0.04	0.02	0.03	0.05	0.06	0.29	0.02	0.04	0.11	0.03	0.12	0.01	0.08	0.11	0.13	0.08	0.00	0.07	0.07	0.09
TEKO	0.38	0.37	0.37		0.04	0.09	0.11	0.03	0.05	0.08	0.07	0.04	0.04	0.19	0.07	0.06	0.01	0.06	0.03	0.08	0.02	0.02	0.04	0.01	0.10	0.02	0.03
SHES	0.38	0.37	0.36	0.36		0.07	0.06	0.08	0.01	0.02	0.05	0.04	0.02	0.01	0.22	0.05	0.03	0.04	0.03	0.05	0.06	0.01	0.05	0.06	0.02	0.07	0.01
SHE	0.37	0.37	0.36	0.37	0.33		0.06	0.07	0.09	0.01	0.03	0.06	0.05	0.02	0.02	0.21	0.05	0.04	0.03	0.04	0.05	0.06	0.00	0.04	0.06	0.01	0.08
SHAH	0.38	0.37	0.36	0.37	0.37	0.33		0.14	0.01	0.01	0.07	0.05	0.02	0.03	0.06	0.06	0.29	0.03	0.04	0.11	0.04	0.12	0.02	0.08	0.12	0.13	0.09
SER	0.37	0.36	0.36	0.36	0.36	0.36	0.34		0.05	0.08	0.10	0.02	0.04	0.07	0.06	0.03	0.03	0.20	0.06	0.04	0.02	0.05	0.04	0.07	0.01	0.03	0.05
MENS	0.37	0.37	0.36	0.37	0.36	0.37	0.36	0.34		0.01	0.12	0.14	0.07	0.09	0.11	0.10	0.08	0.07	0.16	0.11	0.09	0.02	0.10	0.01	0.12	0.05	0.02
MENK	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.33		0.02	0.11	0.13	0.05	0.07	0.10	0.09	0.06	0.06	0.17	0.09	0.08	0.01	0.08	0.01	0.10	0.04
MA	0.37	0.36	0.37	0.37	0.37	0.37	0.36	0.36	0.36	0.34		0.06	0.07	0.09	0.02	0.03	0.06	0.05	0.03	0.02	0.21	0.06	0.04	0.03	0.04	0.04	0.07
KER	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36		0.13	0.01	0.02	0.05	0.03	0.01	0.02	0.04	0.05	0.28	0.01	0.03	0.10	0.02	0.11
HAM	0.38	0.37	0.36	0.36	0.37	0.37	0.37	0.36	0.37	0.37	0.36	0.32		0.02	0.11	0.13	0.06	0.08	0.10	0.09	0.07	0.06	0.17	0.10	0.08	0.01	0.09
HAB	0.38	0.37	0.36	0.37	0.36	0.37	0.36	0.36	0.36	0.37	0.36	0.37	0.34		0.10	0.03	0.05	0.03	0.01	0.02	0.01	0.02	0.02	0.25	0.01	0.00	0.07
GOL	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.32		0.03	0.10	0.12	0.04	0.06	0.09	0.08	0.06	0.05	0.18	0.09	0.07
GHE	0.38	0.37	0.36	0.37	0.37	0.37	0.37	0.36	0.37	0.37	0.36	0.37	0.37	0.36	0.35		0.10	0.03	0.05	0.02	0.01	0.02	0.01	0.01	0.02	0.25	0.02
FOR	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.35	0.36	0.36		0.12	0.02	0.03	0.04	0.02	0.00	0.01	0.03	0.04	0.27
GAL	0.37	0.36	0.36	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.37	0.36	0.36	0.37	0.36	0.33		0.15	0.28	0.30	0.23	0.24	0.27	0.26	0.24	0.23
ENG	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.35	0.36	0.36	0.36	0.33		0.08	0.05	0.07	0.01	0.01	0.04	0.03	0.01
EMB	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.35	0.36	0.36	0.36	0.36	0.35		0.09	0.05	0.06	0.01	0.01	0.03	0.02
DOGO	0.37	0.37	0.37	0.37	0.36	0.37	0.36	0.36	0.36	0.37	0.36	0.37	0.37	0.36	0.37	0.36	0.36	0.36	0.36	0.36		0.11	0.02	0.04	0.04	0.02	0.01
CF	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36		0.12	0.01	0.03	0.05	0.03
CAM	0.39	0.38	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.38	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.34		0.09	0.04	0.06	0.02
BAR	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.38	0.36		0.08	0.06	0.08
BADA	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.38	0.37	0.35		0.15	0.02
AW	0.37	0.36	0.36	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.37	0.37	0.36	0.37	0.36	0.36	0.36	0.36	0.36	0.36	0.37	0.37	0.37	0.33		0.13
AFE	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.38	0.37	0.37	0.37	0.35	
	SEN	**REK**	**TEKR**	**TEKO**	**SHES**	**SEN**	**SHE**	**SER**	**MENS**	**MENK**	**MA**	**KER**	**HAM**	**HAB**	**GOL**	**GHE**	**FOR**	**GAL**	**ENG**	**EMB**	**DOGO**	**CF**	**CAM**	**BAR**	**BAR**	**AW**	**AFE**

[1] BAR = Barentu, CAM = Cambo-10, AW = Awgaro, KER = Keru, CF = Akordet, SHE = Shambuko, GOL = Goluj, TEKO = Tekombia, TEKR = Tekreret, AFE = Afelba, GAL = Galanefhi, HAB = Habero, HAM = Hamelmalo, MA = Mai-Alba, SER= Serejeka, SEN = Senafe, SHAH = Sheha, SHES = Shieb-Seleba, BADA = Bada, DOGO = Dongolo, EMB = Emberemi, ENG = Enghel-Eila, FOR = Foro, GHE = Shieb-Gedged, MENK = Menkaneli, MENS = Shieb-Mensheb, REK = Rekumbedin

5.3.3. Populations differentiation

Clustering using UPGMA revealed three genetic groups (clusters) with their respective sub-groups (Figure 5.4). The first cluster has only one cattle population (CF) from WLL and the second cluster composed of most cattle populations from WLL (CAM, TEKO, TEKR, GOL, SHE and KER). The third cluster is big group included all cattle population from HL and ELL zones (Arado breed) with admixed cattle population from WLL (AW & BAR; Barka breed).

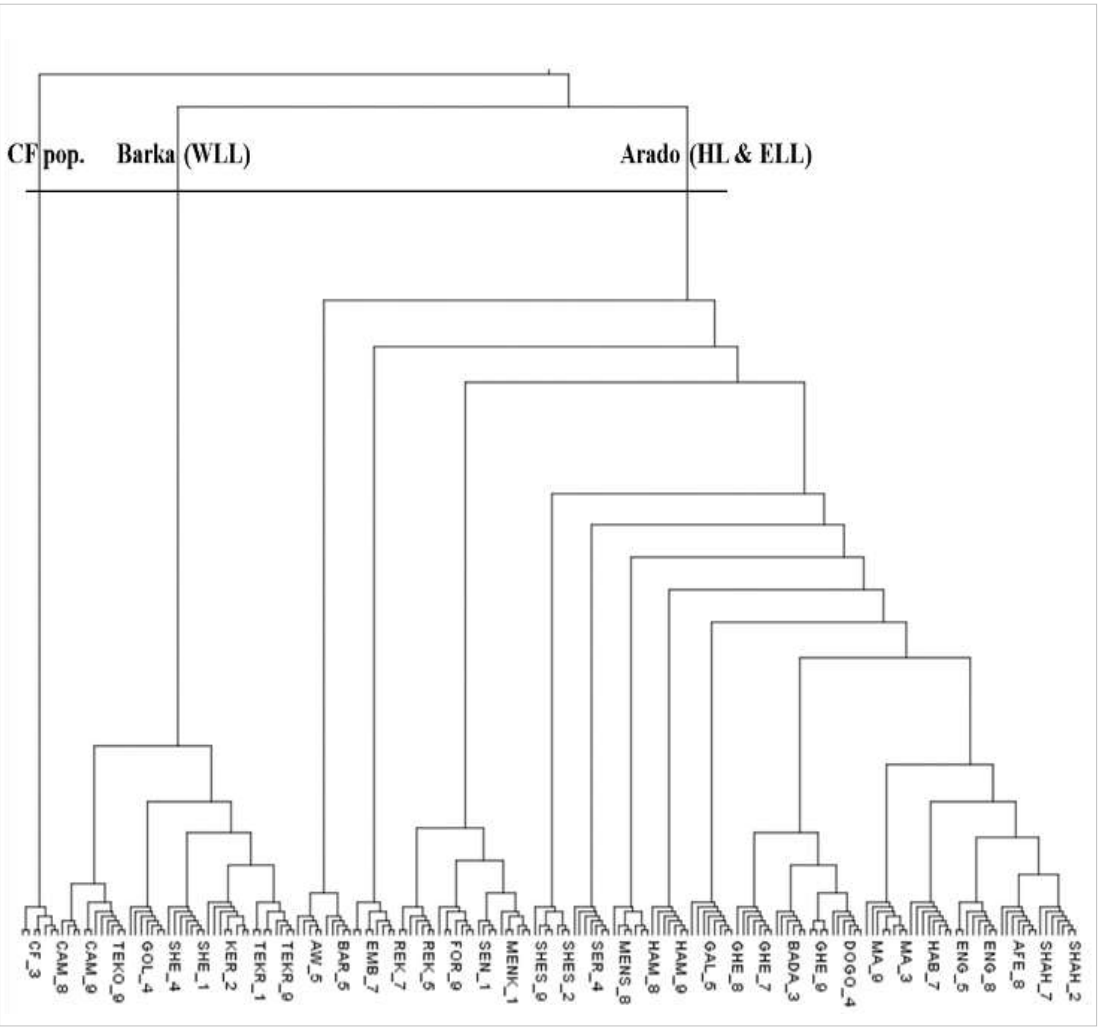


Figure 5.4. Classification of cattle population-based on UPGMA

5.3.4. Population structure and levels of admixture

Population classification and level of gene intermixing among the populations were studied using PCA and admixture analysis. The PCA was performed using allele frequencies of the SNP markers. The first and the second principal components (PC1 & PC2) explained 23.1% and 19.73% of the total genetic variation respectively (Figure 5.5). The PC1 resulted in a differentiation pattern between a cluster composed HL and ELL cattle populations commonly called Arado breed, and cluster composed of many cattle populations from WLL of Barka breed with intermixing HAB and HAM cattle populations. The PC2 did not classify most cattle population rather than separating between GOL, HAM and CF cattle populations, from KER cattle population from Dowhin ecotype of Barka breed.

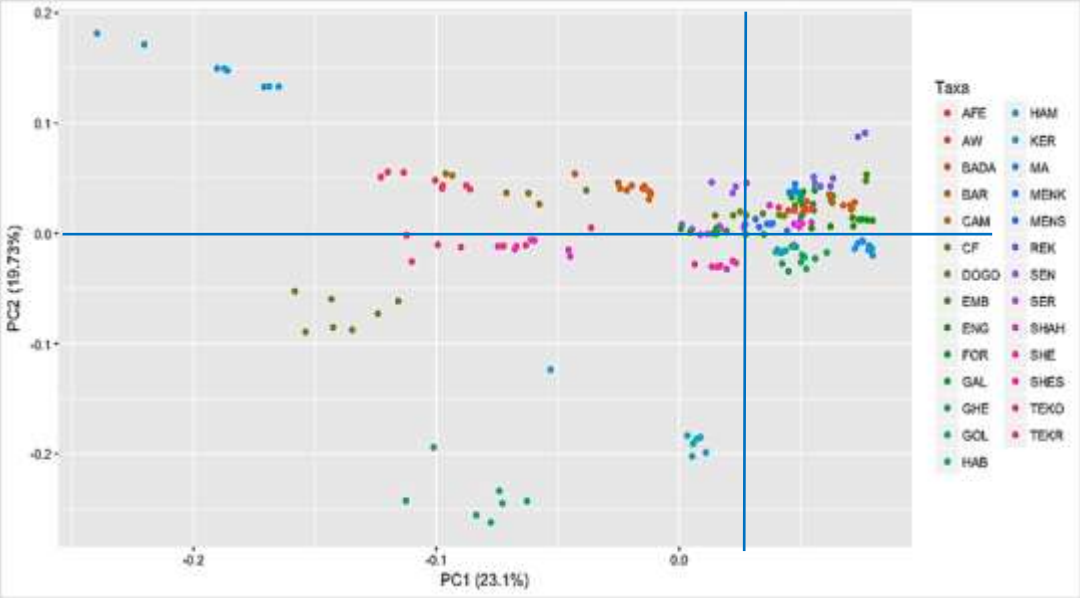


Figure 5.5. PCA of cattle populations based on autosomal SNPs for the first two PCs

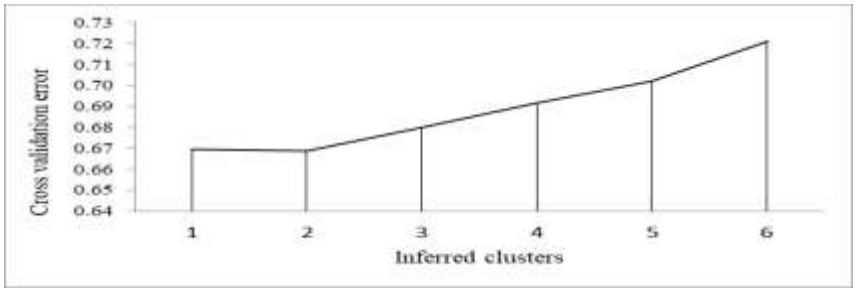


Figure 5.6. Optimum cluster based on cross validation error

Model-based clustering revealed population structure and levels of admixture among the populations (Figure 5.7). The lowest cross-validation error was at K=2 (0.669) indicating 2 as the most likely number of genetically distinct groups (clusters) (Figure 5.6). At K=2, a differentiation was observed between most mixed cattle populations from Barka and Arado breeds (red) and ELL populations (GHE, EMB, HAB, HAM and MENS) of Arebo ecotype (green). However, only two populations (GHE and AW) were clustered exclusively based on admixture software. At K=3, some of the cattle populations (REK, BADA and ENG) appeared as a group (blue) (Figure 5.7).

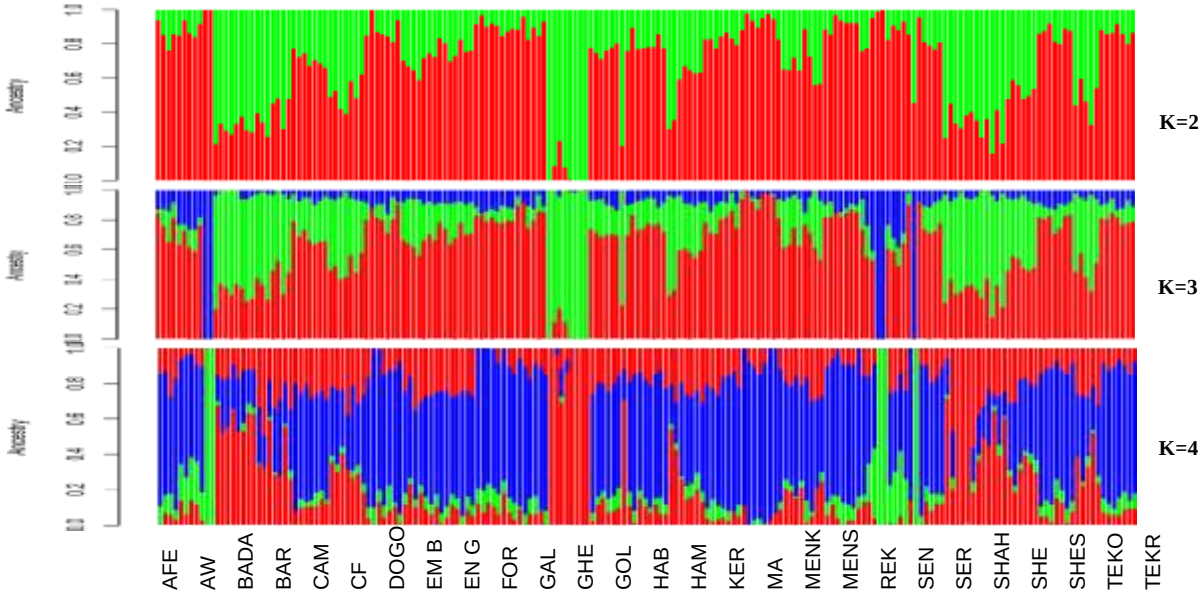


Figure 5.7. Clustering of the cattle populations. Each animal represented by a single white vertical line and colours indicates clusters of cattle populations, and the length of colours shows the magnitude of the proportion of admixture. Red = mixed Barka and Arado breeds, Green = ELL (Arebo ecotype) and Blue = Afar breed

Analysis of Molecular Variance (AMOVA) was conducted based on genetic distances. Genetic variation between the population was different (P <0.001), and accounted for 14.78% of the total variation, the remaining 85.22% is due to within population variation (Table 5.5).

Table 5.5. Analysis of molecular variance (AMOVA) based on genetic distances among cattle populations

Source of variation	DF	SS	MS	VR	*F* value
Among populations	26	0.167	0.0064	14.78	<.001
Within population	161	0.070	0.0004		
Total	**187**	**0.237**			

5.4. Discussion

5.4.1. Population diversity

Genetic characterization using SNP markers produced worthy information in terms of genetic diversity and differentiation of ICPs of Eritrea. Diversity analysis based on A_F showed medium to high variation among ICPs. Relatively, high A_F value (0.227) was observed in AW cattle population. The overall mean (0.157) of A_F of this study is higher than previous studies in indicine cattle (McKay *et al.* 2008; Edea *et al.* 2013; Kim *et al.* 2015), but lower than the result for Red Chittagong cattle (0.28) (Uzzaman *et al.* 2014). High differences in proportion of common SNPs among the cattle populations (51.6% - 59.7%) indicated that there is high variability among the cattle populations in Eritrea. Therefore, mean of proportion of common A_F (56.1%) showed that there is sufficient genetic diversity in Eritrean cattle populations though it is lower than the previous study (64%) of indigenous Zebu breeds in Pakistan (Hamid, 2017).

High heterozygosity indicates high genetic variation which is important for adaptation and improvement of cattle in light of the anticipated future fluctuation of environmental and market conditions. The mean expected value of heterozygosity of 0.218 obtained for the Eritrean *Bos indicus* cattle is lower compared to 0.39 - 0.41 in indigenous Ethiopian cattle (Edea *et al.* 2013), which are the closest cattle population comparable to Eritrean ICPs. Similarly, high levels of expected heterozygosity (0.40) was also reported in Sukuma,

Tarime and Maasai Zebu cattle in Tanzania (Msalya *et al.* 2017). In the current study the within-population genetic variation was highest in BAR and AW cattle populations. The high genetic variability in the two populations could have resulted from uncontrolled breeding (random mating) practised in pastoral and agro-pastoral production systems. Another possible explanation could be due to the frequent intermixing with other indigenous populations and mainly so with the BAR cattle population. This mostly occurs around major towns where animals from different areas are marketed. This movement of animals might have increased gene exchange among the cattle populations around regional town resulting in low level of inbreeding. The results of present study are consistent with finding in indigenous zebu populations of Tanzania that found high levels of admixture (Msalya *et al.* 2017). The low heterozygosity detected in MENK cattle population is expected as the population is mainly kept in an area characterised by extremely hot environmental conditions implying that the population is subjected to high natural selection pressure. This cattle population is found in the desert area of southern Red Sea where intermixing with other populations is also rare due to the environmental barrier. However, in most cattle populations there is sufficient heterozygosity implying low levels of inbreeding among most ICPs. In almost all cattle populations they have negative F_{IS} values suggesting that there is a high proportion of heterozygote genotypes. However, relatively, high inbreeding coefficient in AW cattle population could be due to the *Arem* cultural practice (slaughtering male calf) where very small number of bulls were used for breeding purpose.

5.4.2. Cattle population differentiation

Results from AMOVA indicated that there is a high proportion of variance between cattle populations. The result supports the findings obtained by Goitom *et al.* (2016) which showed that Eritrean ICPs had two distinct morphological categories. The standard estimate of pairwise Nei's genetic distances indicated that CF cattle population had a relatively greater genetic distance (0.38) from BADA, BAR and AFE cattle populations. This

observation is due to the long geographic distance between the cattle populations which could restrict the flow of genes. The other possible explanation is that farmers tend to select against CF (Akordet) cattle population as it's considered to be aggressive making it is routine management practices almost impossible to carry out. The low genetic distance between cattle populations found within HL (Arado) and ELL (Arebo) is due to the close proximity between the populations implying that there is more interbreeding. Population differentiation (F_{ST}), revealed high (>0.25) between BAR and GAL, BAD and GAL, AFE and FOR, DOGO and GAL, and BAR and HAB cattle populations. The high genetic differentiation suggests the populations examined do share much genetic diversity could be due to the existence of gene flow or short geographical distance between cattle populations.

Phylogeny and population structure analysis were done based on genome-wide SNPs from the 27 cattle populations. Multivariate PCA was performed to determine the cattle population structure. A multivariate analysis is essential in grouping correlated characters into uncorrected few new groups (Manly, 1994). This reduction of a set of original variables enables the maximum proportion of variance to be accounted for from a minimum number of new composite variables. The three clusters which were formed by UPGMA has consistency with the classification by PCA. PC1 separated cattle populations of Barka breed from most Arado cattle populations. It is expected to be formed two separate groups because they have differences in their origin. WLL populations (Barka type) is a member of Large East African Zebu while Arado type is from Zanga cattle type. The separation by contrasting some cattle populations (CF, HAM and GOL) from KER cattle population by the PC2. There is no possible explanation of this grouping because KER, CF and GOL are from the same breed but HAM cattle population admixed to the group. However, most cattle populations from Barka and Arado cattle populations were not distinctively separated by PC2. The proportion of genome in each population were inferred by the ADMIXTURE analysis. At K=2, the two clusters which were inferred by the lowest cross-validation were mostly comprised of HL and WLL cattle populations which formed the first cluster (red coloured), and the Arebo type

which formed the second cluster (green coloured). This is possibly explained that Barka breed is less related to Arebo cattle than Arado breed of HL. At K=3, three clusters were realized, which is consistent with the results of cluster analysis. The REK, FOR and ENG are cattle populations which are found in the South of ELL emerged as a separate cluster (blue coloured). The high value of admixture between Arado and Arebo types could be resulted due to recent intermixing populations, and has consistency with the result of the Eritrea ICPs morphological classification (Goitom *et al.* 2016). Moreover, the high admixture among most cattle populations could be due to common ancestry and high gene flow due to the migration of cattle populations due to small land area that characterizes Eritrea.

5.5. Conclusions

Results revealed patterns of genetic variations in the ICPs. The medium to high genetic diversity indicate high potential of ICPs of Eritrea for pure breeding and conservation programes. The model-based structure analysis produced two clusters namely, most Barka and HL Arado cattle populations and the ELL Arado (Arebo) . However, the distance-based UPGMA clustering approach identified three clusters (CF cattle population (Dowhin), most from Barka cattle populations and Arado cattle populations). While population differentiation based on PC2 identified two groups (cluster one comprising of CF, GOL and KER versus MENK cattle population). In conclusion, though ICPs of Eritrea relatively had high morphological diversity than genetic diversity. Nevertheless, the findings of high levels of within and between genetic variations serve as a resourceful information for understanding genetic variability, and will assist in proper establishment of future genetic improvement and conservation program for the ICPs.

CHAPTER SIX

LANDSCAPE GENOMICS AND SIGNATURE ANALYSIS TO UNDERSTAND THE GENETIC ADAPTATION OF ERITREAN INDIGENOUS CATTLE POPULATIONS

6.1. Introduction

6.1.1. Background

Indigenous cattle of Eritrea are found in different Agro Ecological Zones (AEZs). Their different characteristics could be explained by different production systems where they are existed over generations with the differences in environmental effects of the zones having impacted on their adaptation. Eritrea is found in arid and semi-arid northern Sahelo-Sudanian ecological zone of Africa. The dominant cattle population is indigenous (*Bos indicus*; Zebu) to the area. The ability of these indigenous cattle to produce and reproduce under different environment conditions makes them better in the utilization of different local inputs (Mdladla, 2016). Particularly, these cattle living on poor environmental conditions contribute to their unique ability to take advantage of marginal areas.

Currently, there is migration of animals from place to place in search of pasture (feed) and water. Such movement of animals encourages free flow of genes between the populations leading to similarity in their genetic make-up and the general population genetic composition. Manel *et al.* (2003) explained that the gene flow among natural populations depends on ecological features and human activities (mostly artificial selection). These activities limit the free flow of genes among cattle populations which has effect in changing allele frequencies. As geographical distances between the habitats of cattle population increases and high resistance of surface in terms of topography makes the gene flow is restricted due to reduced connectivity. This resulted to

acquisition of new adaptive and functional characteristics than similar cattle populations found in other habitats (Wang, 2015).

Manel *et al.* (2003) described landscape genetics as the study of identifying markers linked to adaptive candidate genes to geo-referenced samples collected across landscapes. In landscape genetics habitat, morphology and climatic features are considered as major causes of variation (Holderegger and Wagner, 2006). Therefore, through landscape genetics, cattle populations that show spatial patterns of genetic variability are identified and an investigation is carried out to explain the causes of such patterns. Isolation-by-Distance (IBD) can create genetic differentiation among populations because populations nearer each other share their gametes and specific environmental condition and thus tend to have a common pattern when compared to populations that are far apart.

Selection signatures is the genome that contain a mutation due to natural or artificial selection, and create a special patterns of DNA sequences (Mahmood, 2014). Therefore, detection of signatures within the genome of organisms is key step in understanding the proportion of a genome that are being shaped by the ongoing natural selection (Joost, 2007). This implies that the detection of signatures has a fundamental biological interest, because they can reveal the nature of adaptation and speciation (MacCallum and Hill, 2006). Particularly, a region of genome under selection is likely to be of functional importance, and inferences regarding selection may provide important information (Nielsen, 2005). Molecular markers specifically SNP is used to scan the genome for genetic signals in adaptation of environment (Orr *et al.* 2004). Positive selection of signatures could have relationships with the corresponding phenotype (disease resistance, heat tolerance and others) (Holderegger and Wagner, 2006).

Analysis for selection of signatures, landscape genetics employed different procedures. The IBD is tested based on Mantel's test which examines the regression between genetic differentiation and their geographical distance (Manel *et al.* 2003). This is the first method in exploring spatial patterns of gene flow and inference could be made for the presence of a barrier effects

(Manel *et al.* 2003; Guillot *et al.* 2009). The second method is the logistic regression models which determine the probability of allele presence/absence in a specific environment using the software Samβada (Joost *et al.* 2007). For this method, Geographical Information Systems (GIS) is importance to localize genetic resources on physical map (Longley *et al.* 2015). The geographic coordinates, via GIS tools, constitute additional descriptors or variables in the data sets. The data usually describes the X (longitude) and Y (latitude) coordinates (Mdladla, 2016).

A number of analysis of selection of signatures were used to scan the genome for signals identification. The analysis grouped into single-locus method the F-statistics of population differentiation (F_{ST}) and haplotype-based (Extended Haplotype Homozygosity; EHH) methods (Pybus *et al.* 2013). The F_{ST} method which uses the average F_{ST} value at each locus is robust and easy to implement (Barendse *et al.* 2009) and it is powerful in breed differentiation (Biswas and Akey, 2006). The most differentiated region represented by the largest F_{ST} value (Weir & Cockerham, 1984), could be considered as a region under selection. The second approach is based on EHH using integrated haplotype score to identify the selection of signatures in cattle population.

The current research was initiated from the information from a previous diversity study of the same Indigenous Cattle Populations (ICPs) (Goitom *et al.* 2018) in which the cattle populations showed medium to high genetic diversity. Moreover, cattle populations living in different environments have different functional characteristics in terms of morphological appearance and other adaptive characteristics. Figure 6.1 shows a sample of cattle populations from different regions and AEZs in Eritrea.

Figure 6.1. Morphological differences in terms of coat colour, horn, body size and other physical appearances of cattle populations. Source: (own photo)

The findings presented in this chapter were aimed at determining the relationships between the differences in the landscapes where different cattle breeds are reared and the respective genetic variability using IBD method. The signatures that cause adaptive variation in cattle populations found in different AEZs were also tested using F_{ST} method. Thus, findings of this chapter were aimed at testing the hypothesis that Eritrean cattle populations have no correlation with ecological landscapes where they live.

6.2. Materials and Methods

6.2.1. Study area

The landscape genomics and signature analysis study was carried out in three different environmental landscapes in Eritrea. The country has three AEZs which comprises six administrative zones. The three AEZs vary in geographical landscapes, which have differences in rainfall, temperature and vegetation. The altitude ranges from below zero to 3018m above Sea level East to West, and 680 m to 3018 m above Sea level West to East direction (Figure 6.2).

Figure 6.2. Agro-ecological zones of Eritrea (MoA, 2013 modified)

6.2.2. Data collection

Autosomal filtered SNPs data for 59, 63 and 66 of mature cows from the Western lowland (WLL), highlands (HL) and Eastern lowlands (ELL) AEZs respectively were used for the analysis. The respective regions' name and numbers of samples collected from each of the regions are presented in Table 6.1. The sample collection was majorly determined on the geographical origin of the cattle. Those cows morphologically and historically categorized into two breeds (Arado and Barka) (Goitom *et al.* 2018), however, it is not strongly separated genetically into these groups.

Table 6.1. Name and number (bracket) of sampled cattle populations in three AEZs

WLL		HL		ELL
CF Akordat (7)	Western escarpment	SEN Senafe (4)	Eastern escarpment	GHE Shieb-Gedged (8)
TEKR Tekreret (8)		AFE Afelba (7)		MENS Shieb-Mensheb (7)
KER Keru (8)		GAL Zighb (8)		DOGO Dongolo (8)
BAR Barentu (5)		MA Mai-alba (8)		EMB Emberemi (8)
SHE Shambuko (7)		SHEH Sheha (7)		FOR Foro (7)
TEKO Tekombia (7)		SER Serejeka (7)		ENG Enghel-Eila (8)
AW Awgaro (5)		SHES Shieb-Seleba (7)		MENK Menkaneli (6)
GOL Golul (7)		HAM Hamelmalo (7)		REK Rekumbedin (7)
CAM Cambo-10 (5)		HAB Habero (8)		BADA Bada (7)
Total 59		**63**		**66**

6.2.3. Data analysis

6.2.3.1. Genomic data analysis

DNA was extracted based on the standard protocol involving proteinase K digestion and phenol chloroform extraction (Sambrook, 1998). Quality analysis of the DNA extract was done by setting filtering criteria by removing any SNPs with more than 10% of missing genotypes, call rate >98%, removing SNPs with MAF >5%, pruning for LD set to 50 5 0.2 and HWE (P >0.001) of autosomal SNPs using PLINK v1.07 software (Purcell *et al.* 2007). The analysis of selection of signatures of this study was based on the effect of three AEZs of Eritrea.

6.2.3.2. Diversity analysis

Allele frequency (A_F), expected heterozygosity (H_E), observed heterozygosity (H_O) and inbreeding coefficient (F_{IS}) were calculated to compare the level of heterogeneity between and within cattle populations found in the three AEZs using PLINK v1.07 software (Purcell *et al.* 2007). The F_{ST} fixation indices (Weir and Cockerham, 1984) were calculated by Arlequin software version 3.5 (Excoffier *et al.* 2005). Standard Nei's genetic distances (Nei, 1972) between the three AEZs were estimated by employed genetic distance using TASSEL version 5.2.41 software (Edward *et al.* 2004).

6.2.3.3. Analysis of landscape genomics based on AEZs

Correlations between genetic variability and ecological landscapes were tested using Mantel regression analysis. Two possible barriers (western escarpment and eastern escarpment) between AEZs were used to separate cattle populations found in the three AEZs. The scores of first principal component were used for regression analysis in SPSS statistical software. A total of three comparisons were made; two comparisons were based on the two barriers while the third one is based on the two extreme geographical distance between AEZs. The comparisons were done between cattle populations in ELL and HL, HL and WLL, and WLL and ELL. The differentiation of cattle population across AEZs was analysed using F_{ST} values. Furthermore, analysis of molecular variance (AMOVA) was conducted to test the significance of variability between and within AEZs based on cattle population differentiation values.

6.2.3.4. Analysis selection of signatures

Signature analysis was done for cattle populations based on mean F_{ST} values. A three-step approach described in Ngeno (2015) was used to identify candidate regions affected by selection. The first step was estimation of F_{ST} values using BCFtools 0.1.13. The second step involved removal of windows with less than 10 SNPs and normalization of the mean F_{ST} (ZF_{ST}) data using Z-transformation

(Rubin *et al.* 2010). The last step was identification of candidate regions affected by selection. A threshold cutting of the upper 1% was used to delineate extreme values to determine candidate regions (Voight *et al.* 2006). A ggplot2 package in R software (R Core Team, 2013) was used to visualize the Manhattan plot based on autosomal chromosomes and the ZF_{ST} values.

6.2.3.5. Analysis of functional annotation and gene ontology

Positive signatures within the raised F_{ST} region were extracted for every chromosome of cattle populations found in the three AEZs using PLINK v1.07 software (Purcell *et al.* 2007). Functional annotations of genomic variants were determined by intersecting the candidate regions with the Variant Effect Predictor (VEP) file (Cow; *Bos taurus*) using web-based Ensembl VEP75 databases (McLaren *et al.* 2016). Only missense variants were extracted and their associations were determined using database genes for the Gene Ontology (GO) enrichment. BinGO v2.44 within Cytoscape v.2.8.3 software (Maere *et al.* 2005) was used for GO terms enrichment for biological process and other functions using *Bos taurus* cow.

6.3. Results

6.3.1. Discovery and quality of Single Nucleotide Polymorphism

The total reads produced from GBS procedure were 35.6, 40.5 and 33.5 million for HL, WLL and ELL respectively. Filtering read sequences above 50 base pairs were trimmed out for homogenous length of the sequences using sickle procedure of SAMtools software for best alignment score. Further, filtering SNPs were done for sex chromosomes, Insertion and Deletion (InDel), >99% (>20) quality, LD pruning (50 5 0.2), Minor Allele Frequency (MAF) (≤ 0.05) and Hardy Weinberg Equilibrium (HWE) ($P < 0.001$). Finally, 0.96, 1.20 and 0.92 million autosomal SNPs were remained for HL, WLL and ELL respectively with average 1.03 million polymorphic autosomal SNPs per AEZ were used for diversity and structure analysis of the cattle populations as shown in Table 6.2.

Table 6.2. Total FastQ reads and filtered SNPs of the three AEZs

AEZ	N	Total reads	Total read/population	Total filtered
HL	63	35.56	5.16	0.96
WLL	59	40.51	6.08	1.20
ELL	66	33.54	4.57	0.92
Total	188	109.61	15.81	3.08
Mean/AEZ		36.54	5.27	1.03
Mean/population		4.06	0.59	0.02

Agro Ecological Zones (AEZs); Western lowlands (WLL); Highlands (HL); Eastern lowlands (ELL); Number of animals (N)

6.3.2. Genetic diversity and population differentiation based on AEZs

Genetic diversity indicators (A_F, H_O, H_E and F_{IS}) of cattle populations based on the three AEZs were estimated and are shown in Table 6.3. The analysis was based on the total filtered autosomal SNPs for cattle population of HL, WLL and ELL. The respective monomorphic SNPs (MAF <0.05) were 0.168, 0.160 and 0.175. The highest polymorphic SNPs (MAF ≥0.05) was observed for WLL at 0.34 and lowest at 0.32 for ELL cattle populations. The comparison of the lowlands and HL cattle populations for both monomorphic and polymorphic SNPs showed similar trend.

Heterozygosity and inbreeding coefficients were estimated for cattle populations found in three AEZs. Means of H_E were found 0.26, 0.26 and 0.27 for cattle population found in HL, WLL and ELL respectively. The deviation between H_O and H_E in the studied cattle populations was found almost zero which follows HWE. Mean of F_{IS} estimates ranged from -0.019 to 0.023 among cattle found in three AEZs with the highest value at 0.023 in WLL as shown in Table 6.3.

Table 6.3. Measures of genetic diversity for cattle in three AEZs

AEZs	Heterozygosity			Inbreeding	Monomorphic SNPs	Polymorphic SNPs	
	N	H_O	H_E	(F_{IS})	(MAF <0.05)	(MAF ≥0.05)	%
HL	62	0.261	0.261	-0.001	0.168	0.332	66.4
WLL	60	0.269	0.264	0.023	0.160	0.340	68.0
ELL	66	0.262	0.268	-0.019	0.175	0.324	65.0
Mean	188	0.264	0.264	0.001	0.168	0.332	66.5

Genetic distances based on allele frequencies among cattle populations in AEZs were done using pairwise population differentiation (F_{ST}). The F_{ST} values were ranged from low of 0.021 between HL and WLL to high of 0.04 between HL and ELL as shown in Table 6.4.

Table 6.4. Pairwise F_{ST} based on average 1.03 million SNPs per AEZ

AEZs	HL	WLL	ELL
HL	0		
WLL	0.021	0	
ELL	0.040	0.024	0

Regression analysis based on genetic distance shows that there is no significant difference (t > 0.05) between cattle populations found in ELL and HL. However, there was highly significant (t <0.001) difference between cattle population of WLL and HL. The R^2 value (0.252) interpreted as 25.2% of the variation is explained by the barrier in the western escarpment or by the long distance between these AEZs while only 1.2% is explained by the barrier in the eastern escarpment or distance between HL and ELL. The comparison between ELL and WLL of AEZs which have both barriers between them, has significant effect (t <0.01) and this is due to either by both barriers or by the long distance between them as presented in Table 6.5.

Table 6.5. Comparison among AEZs based on the two barriers

Comparison	Boundary/barrier	r^2	b	*t* value
ELL & HL	Eastern escarpments	0.012	0.080 ±0.006	0.208
WLL & HL	Western escarpments	0.252	0.016 ±0.002	0.000
ELL & WLL	Both escarpments	0.093	0.024 ±0.007	0.001

Figure 6.2 shows the results for the three comparisons among cattle populations found in three AEZs (ELL, WLL and HL) which were carried out using standardized F_{ST} (ZF_{ST}) values. Three separate groups are identifiable with the highest distance (F_{ST} = 0.27 – 0.54) being observed among cattle populations found in WLL and HL in the first comparison. In the second comparison, relative low population differentiation (F_{ST} = 0.16 – 0.23) was observed between ELL and HL cattle populations. The third comparison (ELL vs WLL) has medium variability (0.2 – 0.47) when compared with the above comparisons. However, all comparisons within the range of high population differentiation (F_{ST}). The analysis of cattle population based on barrier/IBD clearly showed that cattle populations clustered according to their geographic distance and origin.

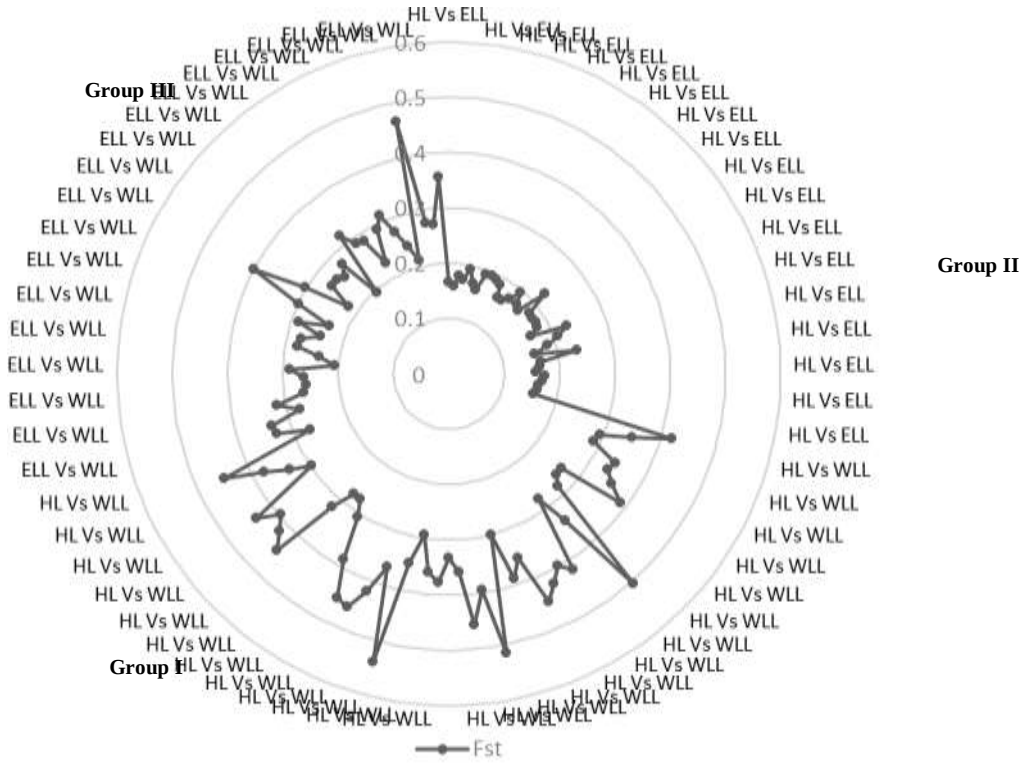


Figure 6.2. Comparison among cattle populations found in three AEZs based on ZF_{ST} values

Analysis of Molecular Variance (AMOVA) was conducted based on population differentiation. All the three comparisons between cattle populations found in different AEZs were different (P <0.001), and accounted for 14.53% from the total variation as shown in Table 6.6.

Table 6.6. Analysis of molecular variance among AEZs

Sources of variation	SS	DF	MS	VR	*F* value
Between AEZs	0.032	2	0.0161	14.53	<.001
Within AEZ	0.205	185	0.0011		
Total	**0.237**	**187**			

6.3.3. Selection of signatures

The 1.03 million SNPs (See Section 6.3.1.) were used in the analysis of selection of signatures which was carried based on F_{ST} method. Further, only chromosomes that had greater than 10 variants (SNPs) were included in the

selected window size and about 43,267 windows having below 10 SNPs were deleted from the analysis. Candidate genomic regions were selected based on ZF_{ST} values greater than 5.04 (1% cut off). Manhattan plot was constructed to visualize distribution of raised F_{ST} values for autosomal chromosomes as shown in Figure 6.3.

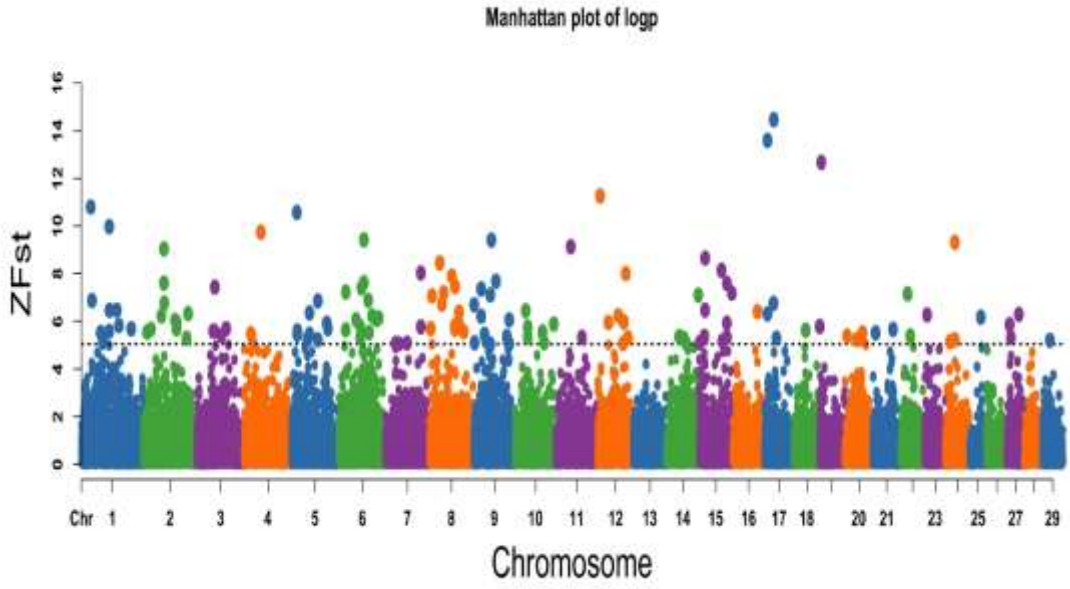


Figure 6.3. Distributions ZFst by autosomal chromosomes, and dash-lined indicates the cut-off (ZFst = 5.04) used for extracting outlier genomic regions

This was followed by identification of 108 genomic candidate regions based on the top 1% ZF_{ST} values across all cattle population. Chromosome one and three were selected for further analysis because they were significantly intersected with the Ensembl VEP database as seen in Table 6.7. The mean ZF_{ST} value across all the SNPs for chromosome one and three were 0.27 and 0.29 respectively which can be considered as representing high genetic differentiation (Wright, 1978). The selected genomic regions in the two chromosomes overlapped with a total 1061 Ensembl genes of which only 292 are missense variants. From all these missense variants only nine genes (*IFNAR2, IFNAR2, CASR, AHSG, ATP1B3, AIRE, ROBO2, SCHIP1* and *PARS2*) were significantly annotated for further gene ontology.

Table 6.7. Characteristics of identified genes in the elevated ZF_{ST} regions

Location	Mean F_{ST}	Allele change	Gene symbol	Gene	SIFT	Codons
1:1594307	0.27	C to G	*IFNAR 2*	ENSBTAG0015212	0.00	A**C**A/A**G**A
1:67342895	0.27	C to G	*CASR*	ENSBTAG003865	0.24	A**C**C/A**G**C
1:81202563	0.27	A to G	*AHSG*	ENSBTAG00522	-	**A**TT/**G**TT
1:128051363	0.27	G to A	*ATP1B 3*	ENSBTAG0014140	0.05	**G**TC/**A**TC
1:145750738	0.27	T to C	*AIRE*	ENSBTA0023393	1.00	**T**CC/**C**CC
1:6497335	0.27	A to G	*USP16*	ENSBTAG00000020122	1.00	**A**AA/**G**AA
1:25054614	0.27	G to A	*ROBO 2*	ENSBTAG00000010462	0.38	CGA/CAA
1:107560376	0.27	A to G	*SCHIP 1*	ENSBTAG00000014960	0.00	G**A**T/G**G**T
3:92099986	0.29	G to A	*PARS2*	ENSBTAG00000040313	0.00	**G**GA/**A**GA

The missense variants might be subjected to diversifying selective pressures due to different forces. Therefore, it is expected that the differentiation of the ICPs based on adaptive characteristics into different AEZs is the result of external environmental forces. Gene Ontology (GO) analysis revealed 179 significant GO-terms (P <0.05) involving the nine identified genes. Most candidate regions were neutral (tolerated), have no predicted variation effect on the protein function (SIFT ≥ 0.05) and shows normal pattern of gene flow of cattle populations. However, *SCHIP1* gene found in chromosome one at position 107560376 is involved in morphological specification particularly head and face profiles, locomotion, cell motility/migration, fibroblast migration, and skeletal and muscle system development with deleterious or damaging effect (SIFT = 0.0). Similarly, *ROBO2* gene is involved with cell morphogenesis with SIFT value of zero. The *IFNAR2* gene contained terms

involved in interferon alpha and beta receptors unit, and has a function in blood coagulation and wound healing. *CASR* gene is associated with mineral cellular homeostasis specifically calcium sensing receptors besides functioning in the circulatory system. The *AHSG* gene is involved in negative regulation of bone mineralization and response to stress and defence. *ATP1B3* gene has a function of energy related biosynthesis and metabolic processes as ATPase Na+/K+ transporting sub-unit beta 3 while the *AIRE* gene is involved in autoimmune regulatory function. The *PARS2* gene is involved in amino-acylation specifically and it is responsible for the attachment of proline to the 3' OH group of ribose of the appropriate tRNA. Lastly, the *USP16* (Ubiquitin specific peptidase 16) gene is associated with protein related activities specifically in deubiquitinates histone H2A.

Table 6.8. Genes identified in the elevated ZF_{ST} regions with their respective description of biological process

GO-ID	p-value	x	n	Description	Genes in test
6952	0.0279	1	233	defense response	AHSG
30502	0.0002	1	2	negative regulation of bone mineralization	AHSG
30279	0.0005	1	4	negative regulation of ossification	AHSG
51130	0.0085	1	71	positive regulation of cellular component organization	AHSG
6959	0.0034	1	28	humoral immune response	AIRE
6754	0.0095	1	79	ATP biosynthetic process	ATP1B3
6813	0.0073	1	61	potassium ion transport	ATP1B3
6814	0.006	1	50	sodium ion transport	ATP1B3
8015	0.0075	1	63	blood circulation	CASR
55074	0.0069	1	58	calcium ion homeostasis	CASR
6816	0.0078	1	65	calcium ion transport	CASR
5513	0.0004	1	3	detection of calcium ion	CASR
50880	0.0032	1	27	regulation of blood vessel size	CASR
51924	0.0026	1	22	regulation of calcium ion transport	CASR
60348	0.0069	2	58	bone development	CASR/AHSG
1503	0.006	2	50	ossification	CASR/AHSG
51049	0.0213	2	178	regulation of transport	CASR/AHSG
1501	0.0126	2	105	skeletal system development	CASR/AHSG
30001	0.0272	2	227	metal ion transport	CASR/ATP1B3
7596	0.0061	1	51	blood coagulation	IFNAR2

50878	0.0074	1	62	regulation of body fluid levels	IFNAR2
42060	0.0085	1	71	wound healing	IFNAR2
902	0.01089	1	91	cell morphogenesis	ROBO2
904	0.00826	1	69	cell morphogenesis involved in differentiation	ROBO2
48667	0.00634	1	53	cell morphogenesis involved in neuron differentiation	ROBO2
32990	0.00766	1	64	cell part morphogenesis	ROBO2
48858	0.00706	1	59	cell projection morphogenesis	ROBO2
1822	0.00431	1	36	kidney development	ROBO2
10171	0.00084	1	7	body morphogenesis	SCHIP1
48870	0.01448	1	121	cell motility	SCHIP1
6928	0.01723	1	144	cellular component movement	SCHIP1
10761	0.00012	1	1	fibroblast migration	SCHIP1
60322	0.00084	1	7	head development	SCHIP1
60323	0.00072	1	6	head morphogenesis	SCHIP1
40011	0.01879	1	157	locomotion	SCHIP1
61061	0.01209	1	101	muscle structure development	SCHIP1
60537	0.00802	1	67	muscle tissue development	SCHIP1
9887	0.02537	1	212	organ morphogenesis	SCHIP1
1501	0.01257	1	105	skeletal system development	SCHIP1
48705	0.00539	1	45	skeletal system morphogenesis	SCHIP1
48745	0.00048	1	4	smooth muscle tissue development	SCHIP1
51289	0.0012	1	10	protein homotetramerization	USP16
70646	0.0018	1	15	protein modification by small protein removal	USP16

6.4. Discussion

The present study focused on producing additional information on genetic characterization of cattle found in ecological landscapes of Eritrea. Eritrea had no information on genetic characterization of indigenous cattle resources. The only information available is these studies conducted in neighbouring countries having with the same cattle populations. Therefore, this study was focused on original places of indigenous cattle found in three AEZs. Almost all the ICPs in the country are reared under free communal rangelands characterized by limited human interventions and free movements within their habitat (ecology). Moreover, this study expected to produce reliable information because it used New Generation Sequencing (NGS) technique and high-throughput automated analysis in producing genomic SNPs in identifying genetic diversities and signatures.

6.4.1. Genetic diversity based on AEZs

Cattle characterization of previous chapter showed considerable amount of genetic variation among ICPs (Goitom *et al.* 2018). The hypothesis of this study was that AEZs are the causes for diversity, unique adaptability and other physical characteristics of ICPs. Findings from this study indicated that the AEZs are the causes of diversity specially, among cattle populations found between HL and WLL AEZs. Moreover, long distance (IBD) between and among cattle populations in different AEZs has led to acquisition of special characteristics that have enabled them to produce and reproduce in their respective environment. However, geographical barrier had less effect through the process of isolation by adaptation.

Diversity based on population differentiation (F_{ST}) indicates that there is medium to high variation among cattle populations found in different AEZs. However, relatively less differentiation was observed between cattle populations found in the HL and WLL which contradicted to the result of landscape based on regression analysis. Furthermore, the medium to high genetic diversity among cattle populations identified in the previous study (Goitom *et al.* 2018) has consistency with medium to high genetic diversity among the three AEZs. This substantial variability in genetic make-up could be explained due to outbreeding program of most ICPs of Eritrea.

Phenotypic variability of cattle population has been shaped by human intervention using artificial selection and introgression of exotic genes, and due to natural selection. The combined effect of these factors is change in the genetic make-up of cattle population which leaves a signature in their genetic constitution (genome). The genetic make-up of Eritrean cattle populations has wide variability in morphological structure (Goitom *et al.* 2016), production, reproduction and adaptation to different landscape environments. Some of these differences in characteristics among breeds or populations are associated with the effect of landscapes on genetic make-up of the animals which leaves signatures in their genome which at times has impact on the adaptability of the populations to the local environments. Recently, there are many studies

detecting signatures in cattle population associated to the adaptation to environmental factors (Flori *et al.* 2012; Ai *et al.* 2013).

6.4.2. Selection of signatures

The findings of the signature analysis reported here were based on all cattle populations found in three AEZs and was aimed at determining how divergent selection pressures may have affected the genomic pattern of the cattle populations. Various genes that are associated with different biological processes, cellular component and metabolic processes were identified. These included the *SCHIP1* gene in chromosome one which was associated with the facial morphology (head and face) with deleterious effect has consistency with the result of classification ICPs of Eritrea based on their morphological features (Goitom *et al.* 2016). The same function also found in human being though located in chromosome three. This gene has influence effect on motility and fibroblast migration. Further, affects bone and muscle systems development which is be important for pastoralists as they desire to select animals that are able to walk long distances in search of feed and water. Other genes identified included the *IFNAR2* and *AIRE* genes which are associated with immune mechanism that contribute to adaptation. The *AHSG* and *CASR* genes which are involved in biological processes mainly in the negative regulation of bone mineralization and calcium sensing respectively. Elena *et al.* (2011) found out that mutation on *CASR* gene in human and bovine chromosomes may cause hypo/hypercalcemia further confirming that the gene is involved in calcium balance. The condition is also common in indigenous cattle during early lactation especially high yielding cows fed on calcium deficient diets.

The *ATP1B3* gene which is found in chromosome one in cattle and chromosome three in human has the same function of transporting ATPase Na+/K+ across the plasma membrane which is essential for osmoregulation in hot environment. The *PARS2* gene which was also identified in this study is involved in mitochondrial amino-acylation. Sofou (2015) observed that

mutation on the *PARS2* gene specially in human is associated with a variety of neurodegenerative disorders such as the Alpers syndrome. Genes with effects on coat colour, body size and polledness have been identified (Biswas and Akey, 2006).

The genomic regions with high F_{ST} values overlapped with genes potentially favoured differently by selection of ICPs in different landscapes. The nature of genetic variation observed in the genomic region along environmental gradients interpreted as that caused by natural selection (Schmidt *et al.* 2008). Furthermore, detection of the genomic regions affected by selection has importance in understanding cattle breeding goals (Elferink *et al.* 2012). Therefore, these significant missense genes within the regions should be considered when setting up breeding programs of indigenous cattle found in different AEZs.

6.5. Conclusions

The findings from the study indicate that ecological landscapes have influence on the genetic differentiation of Eritrea cattle populations resulting from genotype by environment interaction. The differentiation is orientated towards adaptive characteristics and morphological appearances. It is remarkable that the diversity based on F_{ST} analysis showed the existence of less variability between cattle populations found in ELL and HL when compared to those between WLL and HL. This implies that the IBD principle (based on AEZs) has more influence on the genetic differentiation than on the IBR (barrier). The study detected new signatures thus providing important information on genomic regions that are under the effect of environment driven selection. The findings indicated that the regions under selection host genes with importance in various biological processes including facial morphology, immune regulation, body fluid regulation, skeletal and muscle systems development which have importance in adaptation.

CHAPTER SEVEN

GENERAL CONCLUSIONS AND RECOMMENDATIONS

7.1. General conclusions

Morphological and genetic characterization of indigenous cattle of this study produced important information on cattle husbandry in Eritrea. Study on the production systems and status of the indigenous cattle indicated that much of the practices are still based on traditional interventions with indigenous cattle dominating all the production systems. Breed improvement through selection is unstructured with good performing animals being selected irrespective of their genetic relationships. The predominant system was found to be low input low output extensive system with much of the animals serving subsistence needs of the livestock keepers' households.

Morphological characterisation distinguished two distinct cattle groups (breeds). However, within each group there is huge amount of diversities. There is evidence that a sub-cluster of Dowhin cattle population formed within Barka breed. This Dowhin cattle is found in the WLL (Akordet, Keru and Tekreret sites) which is relatively drier. Further, Dongolo cattle population sub-cluster was found to be within the Arado breed. The sub-clustering can be attributed to the climatic condition of the area where they are kept which is characterised by two rainy seasons resulting to higher supply of forage. Indigenoues cattle in the ELL did not form sub-clusters. Genetic characterization based on diversity and structure analysis mainly focused on confirming the morphological classification besides determining the genetic diversity between/within cattle populations. Genetic analysis identified a third cluster that was not detected through morphological characterisation. The cluster was identified from the CF cattle population. The genetic analysis revealed high levels of genetic diversity between and within cattle populations.

Morphological and genetic variability of cattle populations have been influenced by human intervention and natural selection. These interferences could result to identifiable signatures in the genetic constitution (genome) of ICPs. Eritrean cattle genetic resources have wide variability in morphological structure, production, reproduction and adaptation when reared in the diverse landscapes. For instance, the Afar cattle breed and Arebo ecotype are adapted to high temperature and low altitude around sea coasts. Similarly, Dowhin ecotype is adapted to hot and arid environmental condition in WLL. The landscape genomics in this study has revealed cattle populations differentiation resulting from the effect of environment in they are kept. This is clearly demonstrated by separation of cattle populations found in WLL from HL by western escarpment (West barrier) or by long distance (IBD) between the populations. The analysis further revealed a total of nine genes were identified with importance in cattle adaptation to the biophysical challenges and opportunities in the environments they are reared and have potential for identification and selection.

7.2. General recommendations

The results obtained from this characterization study will provide valuable information that could be used in improvement and conservation plan of ICPs of Eritrea. There exists sufficient genetic variability between and within indigenous cattle populations in Eritrea which should be considered in design of cattle genetic improvement programme. More emphasises should be in within population selection as findings from this study have shown that cattle populations have evolved to adapt to the harsh environmental conditions in the areas where they are reared. Genes identified based on genetic and AEZs variabilities should be considered in future Genome Wide Association Study (GWAS) programme in identification of QTLs in cattle genetic improvement plan.

REFERENCES

Abdulmojeed, Y., Kingsley, O., Hadiza, S., Matthew, W. and Samuel, A. (2010). Multivariate analysis of phenotypic differentiation in Bunaji and Sokoto gudali cattle. *Agris category code*, 111-121.

Ai, H., Huang, L., and Ren, J. (2013). Genetic diversity, linkage disequilibrium and selection signatures in Chinese and western pigs revealed by genome-wide SNP markers. *PLoS ONE* 8: e56001.

Alexander, D.H., November J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 19: 1655–1664.

Alsiddig, M.A., Babiker, S.A., Galal, M.Y. and Mohammed, A.M. (2010). Phenotypic characterization of Sudan Zebu cattle (Baggara type). *Veterinary Sciences*, 5: 10-17.

Arsham, H. (2005). Questionnaire design and surveys sampling, 9th edn. http://home.ubalt.edu/ntsbarsh/statdata/Surveys.htm.

Baral, K., Coulman, B., Biligetu, B. and Fu, Y. (2018). Genotyping-by-Sequencing enhances genetic diversity analysis of crested wheatgrass. 19: 2587.

Barcaccia, G., Felicetti, M., Galla, G., Capomaccio, S., Cappelli, K. and Albertini, E. (2013). Molecular analysis of genetic diversity, population structure and inbreeding level of the Italian Lipizzan horse. *Livest Sci.* 151: 124-33.

Barendse, W., Harrisson, B.E., Bunch. R.J., Thomas, M.B. and Turner, L.B. (2009). Genome wise signature of positive selection: The comparison of independent samples and the identification of region associated to trait. *BMC Genomics*, 10: 178.

Barton, N. and Bengtsson, B.O. (1986). The barrier to genetic exchange between hybridising populations. *Heredity*, 57: 357–376.

Biswas, S. and Akey, J.M. (2006). Genomic insights into positive selection. *Trends Genet*, 22: 437–46.

Breuil, Christophe, Grima and Damien. (2014). Baseline report Eritrea: Smart fish programme of the Indian Ocean commission, Fisheries Management FAO component, Ebene, Mauritius. 25.

Chali, Y. (2014). In Situ phenotypic characterization and production system study of Arsi cattle type in Arsi HL of Oromia region, Ethiopia. *In:* Thesis, Haramaya University, Ethiopia.

Chebo, C., Ayalew, W. and Wuletaw, Z. (2014). Traditional breeding practices and trait preferences of cattle farmers in Gamo Goffa zone, southern Ethiopia. *Animal Genetic Resources*, 55: 19-27.

Chen, S., Lin, B.Z., Baig, M., Mitra, B., Lopes, R.J., *et al.* (2010). Zebu cattle are an exclusive legacy of the South Asia Neolithic. *Mol Biol Evol*, 27: 1–6.

Childers, D., Corman, J., Edwards, M. and Elser, J. (2011). Sustainability challenges of phosphorus and food: Solutions from closing the human phosphorus cycle. *Bioscience*, 61(2): 117.

Cock, P.J.A., Gruning, B.A., Paszkiewicz, K. and Pritchard, L. (2013). Galaxy tools and work flows for sequence analysis with applications in molecular plant pathology. *Peer J*, 167.

Cribb, J. (2010). The coming famine: The global food crisis and what we can do to avoid it. *University of California Press*, Berkeley, CA, USA.

Dackson, N.Z. (2008). Genetic characterization of Zambian native cattle breeds. *In:* Thesis. In Animal and Poultry Sciences, Faculty of the Virginia Polytechnic Institute and State University.

DAGRIS, (2007). Domestic Animal Genetic Resources Information System (DAGRIS). Eds. Kemp, S., Mamo, Y., Asrat, B. and Dessie, T., International Livestock Research Institute, Addis Ababa, Ethiopia.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R. (2011). 1000 Genomes Project Analysis Group,

Bioinformatics. Variant Call Format and VCFtools. http://dx.doi.org/10.1093/bioinformatics/btr330.

De Donato, M., Peters, S.O., Mitchell, S.E., Hussain, T. and Imumorin, I.G., (2013). Genotyping-by-sequencing (GBS): A novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. 8: e62137.

Delgado, C., Rosegrant, M., Steinfeld, H., Ehui, S. and Courbois, C. (2001). Livestock to 2020: The next food evolution. Food, agriculture and environment discussion paper 28. International Food Policy Research Institute, Food and Agriculture Organization, and International Livestock Research Institute.

Dessalegn, G., Mekonnen, H. and Kelay, B. (2012). Morphometric characteristics and livestock keeper perceptions of Arado cattle breed in northern Tigray Ethiopia: *Livest. Res. Rural Dev*, 24.

Desta, T., Ayalew, W. and Hedge, B. (2011). Breed and trait preferences of Sheko cattle keepers in southern Ethiopia. *Trop. Anim. Health Prod*, 43: 851–856.

Dossa, L., Wollny, C., Gauly, M. (2007). Spatial variation in goat populations from Benin as revealed by multivariate analysis of morphological traits. *Small Rum Res*, 73: 150–159.

Drinkwater, R.D. and Hetzel, D.J. (1991). Application of Molecular Biology to Understanding resources in a changing world. Resources, 10: 564–567.

Edea, Z., Dadi, H., Kim, S., Dessie, T., Lee, T., Kim, H., *et al.* (2013). Genetic diversity, population structure and relationships in indigenous cattle populations of Ethiopia and Korean Hanwoo breeds using SNP markers. *Front. Genet*. 4: 35.

Edward, B. et al. (2004). TASSEL (Trait Analysis by Association, Evolution and Linkage). Version 5.2.41.

Elena, L., Renata, S., Catherine, R., Silvia, V., *et al.* (2011). Mutations of calcium-sensing receptor gene: two novel mutations and overview of impact on calcium homeostasis. *Euro J Endocrinology*, 165: 353–358.

Elferink, M.G., Megens, H.J., Vereijken, A., Hu, X., Crooijmans, R.P. and Groenen, M.A. (2012). Signatures of selection in the genomes of commercial and non-commercial chicken breeds. *PloS one*, 7: e32720.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E. (2011). A robust, simple Genotyping-By-Sequencing (GBS) approach for high diversity species. *PLoS One*, 6: 1–9. (CrossRef)

Endashaw, T., Tadelle, D., Aynalem, H., Wudyalew, M., and Okeyo, M. (2012). Husbandry and breeding practices of cattle in Mursi and Bodi pastoral communities in Southwest Ethiopia. *AFR J AGR RES*, 7(45): 5986-5994.

Epstein, H. and Mason, I.L. (1984). Evolution of domesticated animals. *Longman*, New York.

Excoffier, L. and Lischer, H.E. (2010). ARLEQUIN suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology*.

Excoffier, L., Laval, G. and Schneider, S. (2005). Arlequin Version 3.0: An integrated software package for population genetics data analysis. Switzerland: Computational and Molecular Population Genetic Laboratory (CMPG), Institute of Zoology, University of Berne.

Falconer, D.S. and Mackay, T.F. (1996). Introduction to quantitative genetics. 4th ed. Harlow, England, Longman.

FAO, (1986). Animal genetic resources data banks 2. Descriptor lists for cattle, buffalo, pigs, sheep and goats. *Animal Production and Health Paper*, 2(59): Rome.

FAO, (1994). Eritrea agricultural sector review and project identification Report III. FAO, Rome.

FAO, (1996). World livestock production systems. Current status, issues and trends, by Seré, H. Steinfeld & J. Groenewold. *FAO Animal Production and Health Paper No. 127*, Rome, Italy.

FAO, (2000). World watch list for domestic animal diversity. 3rd ed., FAO, Rome, Italy.

FAO, (2002). Cattle and small ruminant production systems in sub-Saharan Africa. A systematic review, by M.J. Otte and P. Chilonda. Rome, Italy.

FAO, (2004). Review of the livestock sector in the horn of Africa. *Animal Production and Health Paper No. 88.* Rome.

FAO, (2006). Livestock numbers, meat and total milk production of Eritrea 2001-2005. FAOSTAT 2006, Rome, Italy.

FAO, (2011). Draft guidelines on phenotypic characterization of animal genetic resources. Rome, Italy.

FAO, (2012). Phenotypic characterization of animal genetic resources. *FAO Animal Production and Health Guidelines. No.11*, Rome, Italy.

FAO, (2015). The second report on the state of the worlds animal genetic resources for food and agriculture, edited by B.D. Scherf and D. Pilling. FAO commission on genetic resources for food and agriculture assessments. Rome, 415-450.

Feder, J.L. and Nosil, P. (2010). The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution*, 64: 1729–1747.

Finch, V.A., Bennett, L.L. and Holmes, C.R. (1984). Coat colour in cattle: Effect on thermal balance, behaviour and growth, and relationship with coat type. *Journal of Agricultural Sciences*. 102: 141-147.

Firas, R.A., Abdulkareem, A.A. (2015). Molecular Markers: An introduction and applications. *European Journal of Molecular Biotechnology*, 9: 118-130.

Flori, L., Gonzatti, M.I., Thevenon, S., Chantal, I., Pinto, J., Berthier, D., Aso, P.M. and Gautier M. (2012). Aquasi-exclusive European ancestry in the genepool tropical cattle breed highlights the importance of the slick locus in tropical adaptation. *PLoS One* 7: e36133.

Frankham, R. (1999). Quantitative genetics in conservation biology. *Genet. Res, Camb,* 74: 237–244**.**

Ftiwi, M. and Tamir, B.J. (2015). On-farm phenotypic characterization of indigenous Begait cattle in western Tigray, northern Ethiopia. *Anim. Pro. Adv*, 5(7): 718-732.

Gamaniel, I.B. and Gwaza, D.S. (2017). Molecular characterization of animal genetics resources, its potential for use in developing countries. *Journal of Genetics and Genetic Engineering*, 1: Issue 1, 43-57.

Gavrilets, S., Cruzan, M.B. (1998). Neutral gene flow across single locus clines. *Evolution*, 52: 1277–1284.

Genstat Release 4.2. (2014). 14th ed. Lawest Agricultural Trest, UK.

Genzebu, D., Tamir, B. and Berhance, G. (2016). Study reproductive and productive performance of crossbred dairy cattle under smallholder's management system in Bishoftu and Akaki towns. *International Journal of Advances in Research Biological Sciences.* 3: 118-123.

Gizaw, S., Komen, H., Hanote, O., van Arendonk, J.A.M., Kemp, S., Haile, A., Mwai, O. and Dessie, T. (2011). Characterization and conservation of indigenous sheep genetic resources: A practical framework for developing countries. ILRI Research Report No. 27, Nairobi, Kenya, ILRI.

Goitom, S., Gicheha, M.G. and Teclehimanot, G. (2016). Morphological characteristics of indigenous cattle in Eritrea. Ruforum Biennial Conference: Linking Universities with private sector, governments and other stakeholders in support of agricultural development in Africa 17th–2st October 2016, Cape Town, South Africa.

Goitom, S., Gicheha, M.G., Francis, K.N. and N'geno, K. (2018). Genetic diversity, population structure and admixture analysis in Eritrean indigenous cattle. *South Africa Journal of Animal Science*, South Africa. [Accepted for publication]

Government of Eritrea, 2012. The state of Eritrea ministry of land, water and environment department of environment Eritrea's five years' action plan (2011-2015). Ministry of land, water and environment, Asmara, Eritrea.

Guillot, G. and Santos, F. (2009). A computer program to simulate multi-locus genotype data with spatially auto-correlated allele frequencies. *Molecular Ecology Resources*, 9: 1112–1120.

Güven, G., Bilal, A. and Okan, E. (2010). Ankara use of RAPD-PCR for genetic analyses on the native cattle breeds in Turkey. *Üniv Vet Fak Derg*, 57: 167-172.

Hamid, M., Kim, E., Huson, J., Adeela, A., David, R., Talat, N., Afzal, A., Khalid, J. and Tad, S. (2017). Genome-wide SNPs analysis of indigenous zebu breeds in Pakistan. *Biotechnology in Animal Husbandry*, 33 (1): 13-25.

Hanotte, O., Bradley, D.G., Ochieng, J.W., Verjee, Y., Hill, E.W. and Rege, J.E. (2002). African pastoralism: Genetic imprints of origins and migrations, *Science* 296*:* 336–339.

Hans, E.J. (1982). Livestock production systems and livestock development in tropical Africa. Kieler wissen schaftsverlag vauk, Addis Ababa, Ethiopia.

Hendry, A.P. (2004). Selection against migrants contributes to the rapid evolution of ecologically dependent reproductive isolation. *Evolutionary Ecology Research*, 6: 1219–1236.

Herlocker, D. (1999). Rangeland resources in eastern Africa: Their ecology and development. GTZ, German technical co-operation, Nairobi, Kenya, 213.

Holderegger, R. and Wagne, H.H. (2006). A brief guide to landscape genetics. *Land Ecol 21:* 793–796.

Joost, S., Bonin, A., Bruford, M.W., Després L., Conord, C., Erhardt, G. and Taberlet, P. (2007). A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, 16: 3955–3969.

Joost, S., Duruz, S., Rochat, E. and Widmer, I. (2016). Open computational landscape genetics. *Peer J*, 155-165.

Kanai, E.T, Wamagi, I.T. and Zagi, I. (2013). Phenotypic characterization of white Fulani (Bunaji) and Bunaji x Friesian breed of cattle from national animal production research institute (NAPRI) cattle herd from Nigeria. *World Journal of Agricultural Sciences*, 1 (5): 185-189.

Khan, F.A., Phillips, C.D., Baker, R.J. (2014). Timeframes of speciation, reticulation, and hybridization in the bulldog bat explained through

phylogenetic analyses of all genetic transmission elements. *Syst Biol*, 63: 96-110.

Kim, E., Elbeltagy, A., Aboul-Naga, A., Rischkowsky, B., Sayre, B., Mwacharo, J. and Rothschild, M. (2016). Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity*, 116: 255-264.

Kim, E.S. Sonstegard, T.S. and Rothschild, M.F. (2015). Recent artificial selection in U.S. Jersey cattle impacts autozygosity levels of specific genomic regions. *BMC Geno.* 6: 302.

Kosgey, I.S. (2004). Breeding objectives and breeding strategies for small ruminants in the tropics. *In:* Thesis. Wageningen University, the Netherlands.

Kosgey, I.S., van Arendonk, J.A.M. and Baker, R.L. (2003). Economic values for traits of meat sheep in areas of the tropics with medium to high production potential. *Small Rum. Res.* 50: 187–202.

Kremen, C., Iles, A., and Bacon, C. (2012). Diversified Farming Systems: An agro-ecological, systems-based alternative to modern industrial agriculture. *Ecology and Society,* 17(4): 44.

Kugonza, D.R., Nabasirye, M., Mpairwe, D., Hanotte, O. and Okeyo, A.M. (2011)**.** Productivity and morphology of Ankole cattle in three livestock production systems in Uganda. *Animal Genetic Research*, 48: 13–22.

Laercio, R., Tad, S., George, E., Derek, M., Marcos, V.B., Marco, A., Yuri, T., Jose, F., Cedric, G. and Curtis, P. **(**2013). Genomic divergence of zebu and taurine cattle identified through high-density SNP genotyping. *BioMed Central Genomics*, 14: 876.

Laura, S. (2014). Agro-ecology: What it is and what it has to offer. IIED Issue Paper. IIED, Longman, London, UK.

Lesosky, M., Dumas, S., Conradie, I., Handel, I.G., Jennings, A., Thumbi, S., Toye, P. and Bronsvoort, B.M. (2012). A live weight-heart girth relationship for accurate dosing of east African shorthorn zebu cattle. *Tropical Animal Health Production*, 45: 311–316.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics,* 25: 2078–2079.

Lirón, J.P., Ripoli, M.V., De Luca, J.C., Peral-García, P. and Giovambattista, G. (2002). Analysis of genetic diversity and population structure in argentine and Bolivian creole cattle using five loci related to milk production. *Genetics and Molecular Biology*, 25(4): 413-419.

Lischer, H.E.L., and Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics.* 28: 298–299.

Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W. (2015). Geographic information science and systems. Chichester: Wiley, Pietermaritzburg, South Africa.

Lukuyu, M.N., Gibson, J.P., Savage, D.B., Duncan, A.J., Mujibi, F.D. and Okeyo, A.M. (2016). Use of body linear measurements to estimate live-weight of crossbred dairy cattle in smallholder farms in Kenya. *Springer Plus*, 5: 63.

MacCallum, C. and Hill, E. (2006). Being positive about selection. *PloS Biology*, 4: 293–295.

Maere, S., Heymans, K. and Kuiper, M. (2005). BiNGO: A Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, 21(16): 3448-3449.

Mahmood, G. (2014). Selection signature detection in a diverse set of chicken breeds. *In*: thesis Program for Agricultural Sciences in Goettingen (IPAG), Georg-August-University Göttingen, Germany.

Manel, S., Joost, S. and Epperson, B.K. (2010). Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, 19: 3760–3772.

Manel, S., Schwartz, M.K., Luikart, G. and Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. Trends in ecology and evolution, 18: 189–197.

Manly, F.J. (1994). Multivariate statistical methods. A primer, 2nd ed., Chapman and Hall, Boundary Row, London, UK.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res*. 27: 209–220.

Margaret, O.A. (2002). Characterization of genetic diversity in indigenous cattle of East Africa: Use of micro satellite DNA techniques. International Livestock Research Institute (ILRI), Nairobi, Kenya.

Mason, I.L. and Maul, J.P. (1960). The indigenous livestock of eastern and southern Africa. Common Wealth Bureau of Animal Breeding and Genetics, CAB, Farnham Royal, UK. 179.

Mckay, S.D., Schnabel, R.D., Murdoch, B.M. *et al.* (2008). An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. *BMC Gene*. 9: 37.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R., Thormann, A. *et al*. (2016). The Ensembl Variant Effect Predictor. *Genome Biol.* 17: 122.

Mdladla, K. (2016). Landscape genomic approach to investigate genetic adaptation in South African indigenous goat populations. *In:* Thesis, University of KwaZulu-Natal, South Africa.

Meghen, C., Machugh, D.E. and Bradley, D.G. (1994). Genetic characteristics of West African cattle. *World Anim. Rev*. 78; 59-66.

Mehler, A., Melber, H., Van Walraven, K. (2014). Africa Yearbook 2013: Politics, Economy and Society South of the Sahara. Leiden: Brill.

Mekonnen, A., Haile, A., Dessie, T. and Mekasha, Y. (2012). On farm characterization of Horro cattle breed production systems in western Oromia, Ethiopia. *Livestock Research for Rural Development*, 24: 6-17.

Metzker, M.L. (2010). Sequencing technologies the next generation. *Nat Rev Gene*, 11: 31–46.

Miguel, A.A., Fernando, R.F. and Paulo, P. (2012). Agro-ecologically efficient agricultural systems for smallholder farmers: Contributions to food sovereignty. *Agron. Sustain. Dev.* 32: 1–13.

MoA, (2001). Annual report of Ministry of Agriculture. Asmara, Eritrea.

MoA, (2009). Eritrea: The Ministry of Agriculture striving to combat desertification. Relief web.

MoA, (2011). Annual report of Gash Barka. Barentu, Eritrea.

MoA, (2012). Department of regulatory services plant health division. Asmara, Eritrea.

MoA, (2013). Annual report of Ministry of Agriculture. Asmara, Eritrea.

Msalya, G., Lutatenkwa, D. and Chenyambuga, S.W. (2017). Possibilities of utilizing biotechnology to improve animal and animal feeds productivity in Tanzania–review of past efforts and available opportunities. *J Dairy Vet Anim Res*, 5(5): 155.

Msanga, Y.N., Mwakilembe, P.L. and Sendalo, D. (2012). The indigenous cattle of the southern HL of Tanzania: Distinct phenotypic features, performance and uses. *Livestock Research for Rural Development*, 24: 7.

Mwacharo, J.M., Okeyo, A.M., Kamande, G.K. and Rege, J.E.O. (2006). The small East African shorthorn zebu cows in Kenya. 1: Linear body measurements. *Tropical Animal Health Production.* 38: 65-76.

Mwai, O., Hanotte, O., Young, J.K. and Seoae, C. (2015). African indigenous cattle: Unique genetic resources in a rapidly changing world. *Asian-Australasian Journal of Animal Sciences (AJAS)*, 28(7): 911-921.

National Livestock Development Project (NLDP), (2007). Project completion report. Agriculture and agro industry department Osan.

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

Ngeno, K. (2015). Breeding program for indigenous chicken in Kenya. *In:* Thesis, Wageningen University.

Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics*, 39: 197–218.

Nosil, P., Funk, D.J. and Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology,* 18: 375–402.

Notter, D.R. (1999). The importance of genetic diversity in livestock populations of the future. *J Anim Sci.* 77(1): 61-89.

Obert, T., Voster, M. and Kennedy D. (2013). Preferential traits for breeding Nguni cattle in low-input *in-situ* conservation production systems. *PMCID*, 12: 195.

Oleksyk, T.K., Smith, M.W. and O'Brien, J. (2010). Genome-wide scans for footprints of natural selection. *Philos Trans R Soc Lond B,* 365: 185-205.

Orr, H.A, Masly, J.P. and Presgraves, D.C. (2004). Speciation genes. *Curr. Opin. Genet. Dev.* 14: 675–679.

Parker, P.G., Snow, A.A., Schug, M.D., Booton, G.C. and Fuerst, P.A. (1998). What molecules can tell us about populations: choosing and using a molecular marker. *Ecology*. 79(2): 361–382.

Payne, W.J.A. (1970). Cattle production in the Tropics. Volume1: Breeds and breeding population structure. *Evolution*, 1358-1370.

Perez, O., Brien, A.M., Utsunomiya, Y.T., Mészáros, G., Bickhart, D.M., Liu, G.E., Van Tassell, C.P. *et al.* (2014). Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. *Genet. Sel. Evol*. 46: 19.

Pilling, D. (2010). Threats to animal genetic resources for food and agriculture: Approaches to re-cording, descriptions, classification and analysis. *Animal Genetic Resources*. 47: 11–22.

Powell, R.L., Norman, H.D., and Sanders, A.H. (2003). Progeny testing and selection intensity for Holstein bulls in different countries. *J. Dairy Sci.*, 86: 3386–3393.

Przeworski, M., Coop, G. and Wall, J.D. (2005). The signature of positive selection on standing genetic variation. *Evolution*, 59: 2312-2323.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A. and Bender, D. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *J. Hum. Genet.* 81: 559–575.

Putman, A.I. and Carbone, I. (2014). Challenges in analysis and interpretation of microsatellite data for population genetic studies**.** *Ecol Evol.* 4(22)**:** 4399-4428.

Pybus, M., Dall'Olio, G.M., Luisi P., Uzkudun, M., Carreño-Torres, A., Pavlidis, P. *et al.* (2013). 1000 Genomes Selection Browser 1.0: a

genome browser dedicated to signatures of natural selection in modern humans. *Nucleic Acids Res*, 42: 903–909.

R Core Team, (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (URL http://www.R-project.org/).

Rege, J.E.O and Gibson, J.P. (2003). Animal genetic resources and economic development: Issues in relation to economic valuation. *Ecological Economics* 45: 319–330.

Rege, J.E.O. (1994). Issues and current developments in the conservation of indigenous African domestic animal diversity. Proceedings 5th World congress genetics applied to livestock production (Guelph, Canada), 21: 439-446.

Rege, J.E.O. (1999). The State of African Cattle Genetic Resources I. Classification framework and identification of threatened and extinct breeds. International Livestock Research Institute (ILRI), *AGRI*, 25: 1-25.

Rege, J.E.O. and Tawah, C.L. (1999). The state of African cattle genetic resources II. Geographical distribution, characteristics and uses of present-day breeds and strains. Animal Genetic Resources Information Bulletin, 26: 1–25.

Robinson, A. (2015). QualityTrim. https://bitbucket.org/arobinson/qualitytrim.

Roderic, D.M. (2003). Tangled trees: phylogeny, co-speciation, and co-evolution, The University of Chicago Press, Chicago and London.

Ron, M. and Weller, J.I. (2007). From QTL to QTN identification in livestock winning by points rather than knock-out. *Anim Genet*, 38(5): 429-39.

Rubin, C., Zody, M., Eriksson, M., Meadows, J.R. *et al.* (2010). *Nature*, 464: 587–591.

Saidu, O. and Omedo, B. (2010). Climate change, genetics of adaptation and livestock production in low-input systems 2nd international conference: Climate, sustainability and development in semi-arid regions, Brazil.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning a laboratory manual. 2nd ed, Cold Spring Harbor Laboratory Press, New York, USA.

Scherf, B.D. (2000). World watch list for domestic animal diversity, 3rd eds. 3rd edition. Food and Agriculture Organization of the United Unions, Rome.

Schmidt, P.S., Serrao, E.A., Pearson, G.A. *et al.* (2008). Ecological genetics in the North Atlantic: Environmental gradients and adaptation at specific loci. *Ecology*, 89: 91–107.

Seo, K., Mohanty, R.T., Choi, T. and Hwang, I. (2007). Biology of epidermal and hair pigmentation in cattle: A mini review. *Vet. Dermat.* 18: 392–400.

Sheskin, D.J. (2011). Handbook of parametric and non-parametric statistical procedures. 5th ed., Chapman & Hall, London.

Sneath, P.H.A. and Socal, R.R. (1973). Numerical taxonomy. Software package for population genetics data analysis. Evolutionary Bioinformatics, W.H. Freeman and Co, San Francisco, California, USA.

Sofou, K. *et al.* (2015). Whole exon sequencing reveals mutation in NARS2 and PARS2, encoding the mitochondrial Asparaginyl-tRNA synthesis and prolyl-tRNA synthesis, in patient Alpers syndrome. *Mol. Genet Genomic Med, PMID*, 256-270.

Solomon, G., Komen, H., Hanote, O., van Arendonk, J.A.M., Kemp, S., Aynalem Haile Mwai, O. and Tadelle, Dessie. (2011). Characterization and conservation of indigenous sheep Genetic Resources, a practical framework for developing countries. ILRI Research Report No. 27. Nairobi, Kenya.

SPSS, (2006). Statistical Package for the Social Sciences. SPSS, Inc, Version 23, New York.

Tebug, S.F., Missohou, A., Sabi, S.S., Juga, J., Poole, E J., Tapio, M. and Marshall, K. (2016). Using body measurements to estimate live weight of dairy cattle in low-input systems in Senegal, *Journal of Applied Animal Research*. 46: 87-93.

Uzzaman, M.R., Zewdu, E., Bhuiyan, M.S.A., Jeremy, W., Bhuiyan, A.K., Kwan, S.K. (2014). Genome-wide single nucleotide polymorphism analyses reveal genetic diversity and structure of wild and domestic cattle in Bangladesh. *Asian Australian Journal of Animal Science*, 27 (10):1381-1386.

Voight, B.F., Kudaravalli, S., Wen, X. and Pritchard, J.K. (2006). A map of recent positive selection in the human genome. *PLoS Biol.* 4: 72.

Wang, Y. (2015). Genetic and geographic diversity of Gyr (*Bos-indicus*) cattle in Brazil. University of Natural Resources and Life Sciences, Vienna, Austria.

Weir, B.S. and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution,* 38: 1358–70.

William, P. (2003). Current protocols in bioinformatics. John Wiley & Sons, Inc, 391-412.

Williams, J.L. (2008). Meat Biotechnology. Parco Tecnologico Padano, Polo Universitario, F. Toldr´a (ed.), *Springer Science*, 50: Italy.

Williams, J.L., Haley, C.S. and Haley, C.S. (1999). Discriminating among cattle breeds using genetic markers. *Heredity*, 82: 613-619.

Wright, S. (1951). The genetically structure of populations. *Ann. Eugen.* 15: 323–354.

Wright, S. (1978). Evolution and the genetics of population. In variability within and among natural populations. 4, Chicago, University of Chicago Press.

Yakubu, A., Ogah, D.M. and Idahor, K.O. (2009). Principal component analysis of the morpho-structural indices of White Fulani cattle. *Trakia Journal of Science*. 7: 67-73.

Zechner, P., Zohman, F., Solkner, J., Bodo, I., Habe, F., Marti, E. and Brem, G. (2001). Morphological description of the Lipizzan horse population, *Livest. Prod. Sci.* 69: 163–177.

Zerabruk, M., Vangeo, M. and Haile, M. (2011). The status of cattle genetic resources in North Ethiopia: On-farm characterization of six major cattle breeds. *Anim. Genet. Resources. Info. Bull,* 40: 15-32.

APPENDICES

Appendix 1. Questionnaires for survey data collection

Questionnaire No._____ Date of interview _____/_____/2016

Enumerator name Name_____Code no.
1. Agro-ecological zone Name _____ Code no.
2. Sub Agro-ecological zone Name_____Code no.
3. Sampling site Name _____ Code no.
4. Production system Name _____ Code no.
5. GPS reading _____ *(to be filled in later)*

A. General information

1. Name of head of the family_____
2. Sex
   1. Male__ 2. Female__
3. Age_____
4. Education background
   1. Literate_____ 2. Illiterate____ 3. Read and write_____
5. If your answer is literate, what is your level of education? _____
6. Marital status
   1. Married___ 2. Single___ 3. Divorced ___4. Widow___
7. How many family members do you have? Total _____Male_____ Female_____
8. What is your family livelihood (source of income)?
   1. Agriculture___ 2. Trade___ 3. Employee ___4. Other (specify)_____
9. What is your major farming activity?
   1. Crop ___2. Livestock ___3. Both____
10. What is your family total area of land?
    Crop land_____ Local measurement =_____ha

11. What are the major objectives of cattle production in your family?

Uses	Tick	Rank (top three)
Income		
Home consumption (Meat)		
Home consumption (Milk)		
Saving		
Wealth status		
Manure		
Other (specify)		

12. Household income contribution of different farming activities (in ranking order)

Farming activities	Rank
Cattle production	
Field crop production	
Sheep production	
Goat production	
Apiculture	

Vegetable production	

13 Which species of livestock are more important for your livelihoods (in ranking order)

Species	Rank
Cattle	
Sheep	
Goat	
Equine	
Poultry	
Bee	

B. Reproductive performance (Specific to Eritrean cattle)

1. What is the average age of 1st mating of cows? _____
2. What is the average age of 1st mating of bulls? _____
3. What is the average age of 1st calving? _____
4. What is the average calving interval? _____
5. What is the frequent birth?
   1. Single ___2. Twin ___3. Triple___
6. How long the average reproductive age of the cows? _____Year
7. How many calves are born in the life time of one cow? _____

8. Please tell us the months where frequent calving is happening.

Month	Tick	Rank (top three)
September		
…		
August		

C. Mating and breeding management

1. Do you have your own bull? 1. Yes ___2. No___
2. If yes, how many bull do you have? _____
3. What is your purpose of keeping Bull?
   1. Mating ___ 2. Socio-cultural ___ 3. Fattening ___ 4. Other Specify_____
4. Do you practice control mating? 1. Yes ___2. No___
5. If you answer is yes, how?
   1. Introduction of bull at fixed time (Bull isolation) ___
   2. Castrate unwanted bull ___ 3. Others (Specify)_____
6. If your answer is no, why?
   1. Cattle graze together___ 2. Lack of bull___ 3. Lack of awareness___4. Other specify___
7. Where do you get replacement bull?
   1. From young calves of my own herd ___ 2. From young calves of other herd___
   3. Purchased from market ___ 4. Others (specify)_____
8. Do you select best cows as parent of the next generation with in your cattle? 1. Yes ___2. No___
9. If your answer is yes, what are your selection criteria for cows (cows)?

Criteria	Tick as mentioned	Rank (Top three)

Size/appearance		
Colour		
Calf growth/ Survival		
Udder size		
Calving frequency		
Mothering ability		
Milk yield		
Age at first maturity		

10. Do you select best bulls as parent of the next generation with in your cattle?
    1. Yes___ 2. No___

11. If your answer is yes, what are your selection criteria for bulls (Bull)?

Criteria	Tick as mentioned	Rank (Top three)
Appearance/Conformation		
Colour		
Character (temperament)		
Growth		
Libido		
Age at first maturity		
Pedigree		
Yield of milk/meat		
Adaptability		

12. Do you practice cattle castration?
    1. Yes ___ 2. No___
13. If your answer is yes what are the reasons?
    1. Control mating ___2. Fattening ___3. (specify)____
14. At what age do you castrate your cattle? _____
15. What method do you use for cattle castration? _____
    1. Traditional _____ (specify)_____
    2. Modern _____ (specify)_____

D. Culling
1. Do you practice culling of cows?
    1. Yes ___ 2. No ____
2. If your answer is yes, what is the reason?
    1. Disease ___ 2. Old age ___ 3. Sterility ___ 4. Poor physical condition___
    5. Low milk yield 6. Poor mothering ability 7. Other (specify)_____
3. Do you practice culling of bulls? 1. Yes ___2. No___
4. If your answer is yes, what is the reason?
    1. Disease ___ 2. Old age ___ 3. Poor physical condition ___4. Bad colour___
    5. Poor libido___ 6. Poor horn___
5. At what age cows and bull culled?
    1. Cows_____year 2. Bull _____year
6. What is the use of culled animals?
    1. Sold ___2. Slaughtered ___3. Exchange ___4. Others (specify)_____

E. Market
1. What is average market age of bulls? _____
2. What is average market age of cows? _____

3. Which class of cattle do you sell first in case of cash needed?

Class	Rank	Estimated price
Male calf (<6 month)		
Female calf (<6 month)		
Male (6 to 12 months)		
Female (6 to 12 months)		
Breeding cow		
Breeding bull		
Old cow		
Castrated		

4. What is the cost of the following inputs?

Category	Type	Cost (Nakfa)	Remarks
1.Variable costs	Cost of feed for weaned calf		
	Cost of feed for cow		
	Cost of market for weaned calf		
	Cost of market for culled cow		
	Cost of market for slaughter cow		
	Cost of none-feed for weaned calf		
	Cost of none-feed for cow		
	Cost of reproduction for weaned calf		
	Cost of reproduction for cow		
2. Fixed costs			

5. Do your family sells milk and milk products from cattle?

1. Yes ___ 2. No___

6. If sold, how much was the average prices (in the last 12 months) in Nakfa/kg?

1. Raw milk_____2. Yogurt _____3. Cheese _____4. Butter_____

F. Feeds and Feeding

1. What are the major cattle feed resource in your area?

Feed Resource	Dry season	Rank (Tope 3)	Wet season	Rank (Tope 3)
1 Communal grazing land				
2 Private grazing land				
3 Grazing after harvest				
4 Grazing fallow land				
5 Crop residue				
6 Cut grass and browses				
7 Improved forage				
8 Concentrate				
9.Hay				
10 *Hatella*				
11 Other (specify)				

2. What are the grazing methods in your area in different seasons?

Grazing Methods	Wet season	Dry season
Free grazing		
Herded		
Cut and carry		

3. Do you provide concentrate for your cattle? Yes _____ No _____
4. If your answer is yes,

Type of concentrate	Class of cattle				Name	Rank
	Calves	Cow	Bull	Castrate		
Homemade grain						
Bran						
Oil seed cake						
Hatella (local brewery by product)						

5. When do you provide concentrate for your cattle?
   1. Dry season ___2. Wet season___ 3. Both ___4. Calving ____5. Other___
6. If you don't provide concentrate feed, what are the reasons?
   1. Expensive ___2. Not Available ___3. Not want to offer (specify)_____

G. Housing

1. What shelter do you have for your cattle?
   1. No shelter__ 2. Separate house for cattle main house__ 3. Shelter constructed in side__
   4. Shelter constructed expansion of the main houses. ___5. Open barn ___6. Other ___
2. Are calves housed together with adult cattle?
   1. Yes ___ 2. No ___
3. Are cattle housed together with other animals?
   1. Yes ___ 2. No___
4. If your answer is yes which animals housed together with cattle?
   1. Sheep ___ 2. Goat ___ 3. Equine___ 4. All species___

H. Herding and other management activities

1. How are your cattle herded during grazing time?
   1. With other species___ 2. Separately___3. No control___
2. If they are herded separately; in which season and the reason?

	Season	Reason
1.	_____	_____
2.	_____	_____

3. If they are graze together with other species, with what species,

Species	Season	Reason
1. Goat _____	_____	_____
2. Sheep _____	_____	_____
3. Equines _____	_____	_____
4. All species _____	_____	_____

4. Who do the different tasks and decides on benefits obtained from cattle?

Task	Husband	Wife	Girl	Boy	Hired labour
Herd herding					
Care for calf					
Animal sealing					
Watering					
Milking					
Cleaning					
Product processing					
Castration					
Cut and carry grasses					

5. Do you practice mixing of your cattle herd with other herds?

1. Yes ___ 2. No___

6. If your answer is yes, how many households mix their cattle together? _____

I. Health

1. Please specify (describe) the major cattle disease, their symptoms, season of occurrences, and cultural treatment (please include the predator)

Local name	Symptoms	Season of occurrences	Reason of occurrences	Is it contagious (yes/no)	Age group affected	Local treatments	Service provide

Service: 1. Vaccination 2. Diagnosis 3. Treatment 4. Others (Specify)

2. Do you get vaccination service for your cattle?
   1. Yes ___ 2. No___
3. If your answer is yes, when the service is given?
   1. When disease out brake occur__2. Any time in a year__3. Others (specify)__
4. Where you get treatment and vaccination?
   1. Agricultural office__ 2. NGO__ 3. Private veterinary house__ 4. Others (specify)__
5. How many cattle are died in the last year (Previous 12 months) in your herd?

Category	Reason of death				
	Number of death	Disease	Predator	Mechanical	Other (specify)
Cow					
Bull					
Young Cow					
Young Bull					
Calves					
Castrate					

L. Product utilization

1. Do you slaughter cattle for household consumption?
   1. Yes ___2. No___
2. If your answer is yes how frequent?
   1. festivals ___2. Whenever slaughter old animals available ___ 3. Wedding ___
   4. Births in family ___5. For guests ___6. circumcise specify_____
3. Which sex you usually slaughter?
   1. Intact Male ___2. Castrated___3. Female ___
4. What is the average age of slaughter? Male _____
5. Do you use cattle milk for consumption?
   1. Yes ___2. No___
6. Do you process milk into other product?
   1. Yes ___2. No___
7. If your answer is yes, what are the products?
   1. Yogurt ___ 2. Cheese ___ 3. Butter ___ 4. Others specify_____
8. What is the milk production per day per cow (in litres)?
   Maximum _____Minimum _____Average _____

9. What is the lactation length (in months)?
Maximum _____Minimum _____Average _____
10. Frequency of milking
1. Once a day ___ 2. Twice a day ___ 3. Three times a day___
11. Do you practice weaning? 1. Yes ___ 2. No___

12. Average weaning age of calves?
1. <3 months ___ 2. 3–4 months___ 3. 5–6 months___ 4. >6 months ___
13. Milk feeding up to weaning
1. Unrestricted suckling__2. Restricted suckling__3. Packet feeding__4. Others (Specify)__

M. Production constraint
1. What are the major problems of cattle production in your area? (Rank according to their severity)?

Constraint	**Tick as mentioned**	**Rank (Top three)**
Disease		
Feed shortage		
Water Shortage		
Labour shortage		
Market problem		
Predator		
Lack of input		
Lack of extension service		
Drought		

Recording format for body measurements and physical description

ID	Breed (local name)	Sex	Age	Dentition	Body weight	Body length	Chest girth	Height at wither	Coat	Hair type	Head profile	Dewlap	Tail length		Horn		Ear	
									Colour	Pattern				Shape	Orientation	Length	Formation	Length

Codes for body measurement and physical description

S.No.	Character	Level	Code
	Breed	Barka	1
		Dowhin	2
		Arado	3
		Arebo	4
		Afar	5
		Bwadir	6
		Others	7
	Sex	Male	1
		Female	2
	Dentitions	0 pair of PI lost	0
		1 pair of PI lost	1
		2 pair of PI lost	2
		3 pair of PI lost	3
		4 pair of PI lost	4
		5 broken teeth	5
	Head profile	Straight/flat	1
		Concave	2
		Markedly convex	3
		Slightly convex	4
	Dewlap	Long (>30 cm)	1
		Medium (20 cm-30 cm)	2
		Short (,20 cm)	3
	Ear formation	Rudimentary	1
		Short ear	2
		Long ear	3
		Erect	4
		Pendulous	5
		Semi-pendulous	6
	Horn type	Horned	1
		Polled	2
		Loose (*dombay*)	3
	Horn shape	Straight	1
		Curved	2
		Scars	3
	Horn orientation	Obliquely (curved) upward	1
		Obliquely (curved) front	2
		Obliquely (curved) backward	3
		Lateral	4
	Hair type	Short and smooth	1
		Long and course	2
		Short and course	3
	Cot colour	White	1
		Brown	2
		Red	3
		Black	4
		Grey	5
		When mixed dominant colours	6

S.No.	Character	Level	Code
	Coat colour pattern	Plain	1
		Patchy	
		• Patchy black and white	2
		• Patchy red and white	3
		• Others specify	4
		Spotted	
		• Spotted black and white	5
		• Spotted red and white	6
		• Other specify	7

Milk production performance of a herd

	Cow's name	Top milk Producer (litres)	Herd Average (litres)	Top heavy weight (kg)	Herd average weight (kg)
1					
…					

Appendix 2. Statistical outputs

1. Principal Component Analysis (PCA)

[PRINT=roots, loadings, scores, tests; NROOTS=6; METHOD=ssp] (body length, chest girth, Coat colour, dewlap
width, ear formation, ear length, hair type, head profile, height at wither, Horn orientation, horn type, tail length)
* Principal components analysis *
* Percentage variation *

1	**2**	**3**	**4**	**5**	**6**
74.48	13.20	7.94	3.32	0.50	0.14

*Latent Vectors (Loadings) *

Characteristics	**1**	**2**	**3**	**4**	**5**	**6**
Body length	-0.51484	-0.76677	-0.35328	0.14515	0.02808	-0.01139
Chest girth	-0.70369	0.63309	-0.29015	0.13746	0.02536	-0.00013
Coat colour	-0.03174	-0.00800	0.01144	0.07384	-0.99145	0.04132
Dewlap width	-0.17479	-0.02566	-0.09657	-0.97559	-0.06354	0.02072
Ear formation	-0.01333	-0.00478	0.00072	-0.01287	0.02322	0.41161
Ear length	-0.01609	-0.00536	0.00962	-0.00332	-0.00058	-0.02291
Hair type	0.00120	0.00402	0.00975	0.00435	-0.04772	-0.03101
Head profile	-0.00678	-0.00161	0.00529	-0.00182	-0.02939	-0.02528
Height at wither	-0.45443	-0.10170	0.88349	-0.00286	0.02868	0.01518
Horn orientation	0.00565	-0.00266	-0.0137	0.01659	0.01972	0.86542
Horn type	-0.00465	0.00073	0.00530	0.00073	-0.01486	0.19126
Tail length	-0.03407	0.01069	0.02353	-0.04820	-0.08033	-0.20106

* Significance tests for equality of final K roots *

No. (K) Roots	**Chi-squared**	**df**
2	6.90	2
3	22.26	5
4	74.13	9
5	106.21	14
6	149.19	20
7	201.82	27
8	731.13	35
9	2743.66	44
10	4281.75	54
11	5370.44	65
12	8186.68	77

2. Discriminal analysis: Group statistics

AEZs		Mean	Std. Deviation
Central HL	Body length	112.6557	8.92354
	Chest girth	135.9836	10.38347
	Height at wither	110.7213	4.79629
	Coat colour	3.4918	1.43321
	Hair type	1.2623	.68073
	Head profile	1.0656	.24959
	dewlap	11.0656	4.39647
	Tail length	1.1639	.37329
	Horn type	1.1639	.52219
	Horn orientation	1.5082	.86839
	Ear formation	4.1311	.49918
	Ear length	1.9836	.12804
WLL	Body length	126.4865	4.98008
	Chest girth	156.1351	6.01214
	Height at wither	125.2568	5.78155
	Coat colour	4.7568	1.22542
	Hair type	1.1892	.58930
	Head profile	1.3378	.50415
	dewlap	18.9459	3.57653
	Tail length	2.7432	.46994
	Horn type	1.3108	.59509
	Horn orientation	1.1622	.37112
	Ear formation	4.5946	.92038
	Ear length	2.5676	.49880
ELL (sea coasts)	Body length	111.3333	9.13647
	Chest girth	138.8642	8.52313
	Height at wither	108.0741	4.60374
	Coat colour	3.2346	1.41661
	Hair type	1.1728	.56547
	Head profile	1.0247	.15615
	dewlap	14.2716	2.66464
	Tail length	1.5185	.69121
	Horn type	1.0494	.26932
	Horn orientation	1.5185	.83832
	Ear formation	4.1975	.60041
	Ear length	1.8025	.43069

Variables in the Analysis				
Step		Tolerance	Sig. of F to Remove	Wilks' Lambda
1	height at wither	1.000	.000	
2	height at wither	.994	.000	.394
	tail length	.994	.000	.302
3	height at wither	.993	.000	.310
	tail length	.987	.000	.219
	dewlap	.992	.000	.199
4	height at wither	.993	.000	.229
	tail length	.985	.000	.183
	dewlap	.991	.000	.167

	ear length	.998	.000	.164
5	height at wither	.985	.000	.193
	tail length	.984	.000	.154
	dewlap	.949	.000	.149
	ear length	.986	.000	.143
	coat colour	.935	.000	.139
6	height at wither	.936	.000	.148
	tail length	.972	.000	.139
	dewlap	.908	.000	.122
	ear length	.981	.000	.128
	coat colour	.929	.000	.125
	chest girth	.885	.000	.120
7	height at wither	.933	.000	.140
	tail length	.972	.000	.131
	dewlap	.908	.000	.116
	ear length	.971	.000	.123
	coat colour	.928	.000	.118
	chest girth	.885	.000	.113
	head profile	.987	.004	.107

Classification Results

			Predicted Group Membership			
		AEZ	Central HL	Western lowland	ELL (sea coasts)	Total
Original	Count	Central HL	47	1	13	61
		WLL	0	74	0	74
		ELL (sea coasts)	18	0	63	81
	%	Central HL	77.0	1.6	21.3	100.0
		WLL	.0	100.0	.0	100.0
		ELL (sea coasts)	22.2	.0	77.8	100.0
a. 85.2% of original grouped cases correctly classified.						

3. Logistic regression analysis

1. Cow selection preference

Classification Table[a,b]					
	Observed		Predicted		
			Cow selection 1 dependent		Percentage Correct
			not selected first	first selection	
Step 0	Cow selection 1 dependent	not selected first	122	0	100.0
		first selection	121	0	.0
	Overall Percentage				50.2
a. Constant is included in the model.					
b. The cut value is .500					
Classification Table[a]					

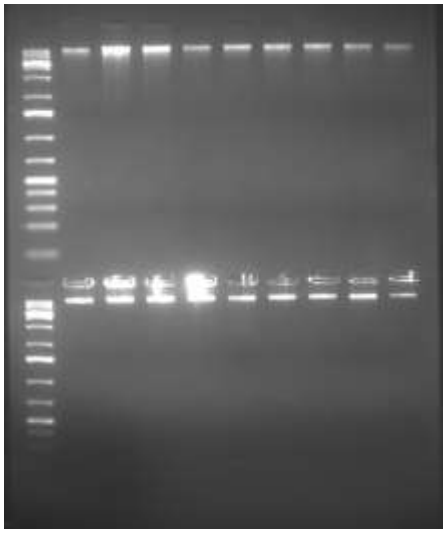
			Predicted		
			Cow selection1depen		
	Observed		not selected first	first selection	Percentage Correct
Step 1	Cow selection 1 dependent	not selected first	84	38	68.9
		first selection	78	43	35.5
	Overall Percentage				52.3
a. The cut value is 0.50					

2. Bull selection preferences

Classification Table[a,b]					
			Predicted		
			Bull rank 1		
	Observed		others	Appearance	Percentage Correct
Step 0	Bull rank 1	others	0	116	.0
		appearance	0	127	100.0
	Overall Percentage				52.3
a. Constant is included in the model.					
b. The cut value is 0.50					

Classification Table[a]					
			Predicted		
			Bull rank 1		
	Observed		others	Appearance	Percentage Correct
Step 1	Bull rank 1	others	38	78	32.8
		appearance	43	84	66.1
	Overall Percentage				50.2
a. The cut value is 0.50					

Appendix 3. Results of some DNA extraction and quality check-up



1. Gel electrophoresis of genomic DNA for Afelba (AFE) and Foro (FOR) cattle populations respectively

2. Some of the result of quality of DNA based on Nonodrop spectrophotometer and others

Module: Nucleic acid, Path: 10 mm, Software: 3.7.1 and Frim ware

Sample ID	ng/ul	A260	A280	260/280	260/230	Config.
FORO_1	112.73	2.255	1.461	1.54	0.7	0.964108/-0.07/128/24
FORO_2	52.76	1.055	0.623	1.69	0.92	0.964108/-0.07/128/24
FORO_3	124.61	2.492	1.507	1.65	0.88	0.964108/-0.07/128/24
FORO_4	116.62	2.332	1.479	1.58	0.77	0.964108/-0.07/128/24
FORO_5	82.13	1.643	1.003	1.64	0.75	0.964108/-0.07/128/24
FORO_6	287.05	5.741	3.767	1.52	0.99	0.964108/-0.07/128/24
FORO_7	161.21	3.224	2.084	1.55	0.67	0.964108/-0.07/128/24
FORO_8	197.76	3.955	2.65	1.49	0.71	0.964108/-0.07/128/24
FORO_9	150.64	3.013	2.137	1.41	0.66	0.964108/-0.07/128/24
ENG_1	275.78	5.516	3.251	1.7	1.09	0.964108/-0.07/128/24
ENG_2	201.82	4.036	2.4	1.68	0.94	0.964108/-0.07/128/24
ENG_3	243.35	4.867	6.405	0.76	-41.09	0.964108/-0.07/128/24
ENG_4	92.62	1.852	1.097	1.69	0.87	0.964108/-0.07/128/24
ENG_5	176.58	3.532	1.905	1.85	1.93	0.964108/-0.07/128/24
ENG_6	127.56	2.551	1.809	1.41	0.6	0.964108/-0.07/128/24
ENG_7	15.49	0.31	0.2	1.55	0.86	0.964108/-0.07/128/24
ENG_8	147.78	2.956	1.811	1.63	0.92	0.964108/-0.07/128/24
ENG_9	499.23	9.985	6.125	1.63	1.12	0.964108/-0.07/128/24
…						

Appendix 5. Some of results of fastQ and analysis population parameters in Bio-informatics pipelines

1. FASTQC of some animals

#FOR 9
* Reads: 3208028
+ Trimmed (pass): 2677178 (83.45%, 1338589 pairs)
+ Singleton (pass): 203556 (6.35%)
- Chastity (fail): 0 (0.00%)
- Length (fail): 327073 (10.20%)
- Average Quality (fail): 0 (0.00%)
- N base count (fail): 221 (0.01%)

#GAL 1
* Reads: 2771584
+ Trimmed (pass): 2324110 (83.85%, 1162055 pairs)
+ Singleton (pass): 172650 (6.23%)
- Chastity (fail): 0 (0.00%)
- Length (fail): 274656 (9.91%)
- Average Quality (fail): 0 (0.00%)
- N base count (fail): 168 (0.01%)

2. Some of the analysis of genomic region based on standardized F_{ST}

CHROM	BIN_STAR	BIN_END	N_VARIAN	WEIGHTEI	MEAN_FS	u(overall r	MEAN_FS	stdev of MEAN	ZFST
chr15	70150001	70250000	19	0.187815	0.106925	0.000629	0.106296	0.015533157	6.843166445
chr1	81900001	82000000	16	0.147256	0.10592	0.000629	0.105291	0.015533157	6.77846614
chr15	70200001	70300000	24	0.162928	0.10344	0.000629	0.102811	0.015533157	6.618807676
chr10	79200001	79300000	34	0.136448	0.102402	0.000629	0.101773	0.015533157	6.551982884
chr21	53200001	53300000	33	0.083224	0.096958	0.000629	0.096328	0.015533157	6.201474616
chr2	68000001	68100000	27	0.112629	0.092894	0.000629	0.092265	0.015533157	5.93986006
chr1	1.15E+08	1.16E+08	29	0.116469	0.090895	0.000629	0.090266	0.015533157	5.81120624
chr17	34050001	34150000	35	0.122006	0.089579	0.000629	0.08895	0.015533157	5.72645206
chr15	12250001	12350000	25	0.105207	0.08763	0.000629	0.087001	0.015533157	5.600978533
chr1	1.22E+08	1.22E+08	26	0.117717	0.084694	0.000629	0.084065	0.015533157	5.411950638
chr21	47900001	48000000	14	0.123822	0.084495	0.000629	0.083866	0.015533157	5.399184399
chr15	76900001	77000000	14	0.143714	0.08421	0.000629	0.083581	0.015533157	5.380785048
chr15	47500001	47600000	44	0.10267	0.08336	0.000629	0.082731	0.015533157	5.326102025
chr11	39600001	39700000	30	0.093013	0.081496	0.000629	0.080867	0.015533157	5.20608135
chr1	87700001	87800000	44	0.126174	0.08052	0.000629	0.079891	0.015533157	5.143260895

3. Genetic parameters estimation (MAF, observed and expected heterozygous and inbreeding coefficients (f))

CHR	SNP	A1	A2	MAF
1	chr1:57052	1	2	0.2143
…				

28	chr1:88747	1	2	0.07143

CHR	SNP	A1	A2	O(HET)	E(HET)	P
1	chr1:57052	1	2	0.7143	0.5	0.5105
…	chr1:127331	1	2	0.2857	0.4898	0.4406
	chr1:242804	2	1	0.7143	0.5	0.5105
28	chr1:271139	1	2	0.1429	0.4592	0.1049

FID	IID	O(HOM)	E(HOM)	N(NM)	F
BADA_1	BADA_1	85220	8.36E+04	123220	0.03991
BADA_2	BADA_2	87403	8.36E+04	123220	0.09506
BADA_3	BADA_3	86434	8.36E+04	123220	0.07058
BADA_4	BADA_4	86292	8.36E+04	123220	0.06699
BADA_5	BADA_5	86329	8.36E+04	123220	0.06793
BADA_8	BADA_8	84934	8.36E+04	123220	0.03268
BADA_9	BADA_9	88159	8.36E+04	123220	0.1142