MOLECULAR CHARACTERIZATION OF ETHIOPIAN INDIGENOUS GOAT POPULATIONS: GENETIC DIVERSITY AND STRUCTURE, DEMOGRAPHIC DYNAMICS AND ASSESSMENT OF THE KISSPEPTIN GENE POLYMORPHISM

Getinet Mekuriaw Tarekegn

A dissertation submitted to the department of Microbial, Cellular and Molecular Biology

Presented in fulfillment of the requirements for the Degree of Doctor of Philosophy in Applied Genetics

Addis Ababa University
Addis Ababa, Ethiopia

June 2016

Department of Microbial, Cellular and Molecular Biology

College of Natural Sciences

Addis Ababa University

This is to certify that the dissertation prepared by Getinet Mekuriaw Tarekegn entitled: "**MOLECULAR CHARACTERIZATION OF ETHIOPIAN INDIGENOUS GOAT POPULATIONS: GENETIC DIVERSITY AND STRUCTURE, DEMOGRAPHIC DYNAMICS AND ASSESSMENT OF THE KISSPEPTIN GENE POLYMORPHISM**" submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy in Applied Genetics complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Melkaye G. Melka (PhD)	----------	----------
External examiner	Signature	Date
Gurja Belay (PhD)	----------	----------
Internal examiner	Signature	Date
Kassahun Tesfaye (PhD)	----------	----------
University Supervisor	Signature	Date
Tadelle Dessie (PhD)	----------	----------
Supervisor	Signature	Date
Okey Mwai (PhD)	----------	----------
Supervisor	Signature	Date
Appolinaire Djikeng (PhD)	----------	----------
Supervisor	Signature	Date
Joram Mwacharo (PhD)	----------	----------
Supervisor	Signature	Date

---------- ----------

Chair of Department or Graduate Program Coordinator

i Acknowledgement

First and above all I would like to thank the **Almighty GOD** with His Mother **St. Virgin Mary** for His absolute protection and mercifulness, for the courage and patience He has unreservedly offered me during the study and all the time. Next, I would like to provide my sincere appreciations to my supervisors: Kassahun Tesfaye (PhD), Joram M. Mwacharo (PhD), Okeyo A.Mwai (PhD), Tadelle Dessie (PhD) and Appolinaire Djikeng (PhD) for their unreserved support and guidance throughout my study.

I would also like to extend my gratitude to the funding organizations of the project. The project was part of the BecA-ILRI Hub livestock productivity program (entitled: Harnessing genetic diversity for improving livestock productivity: goats) funded by the Sida (Sweden Government offices decision UF2011/55504/UD/UP). This project was also supported by the BecA-ILRI Hub through the Africa Biosciences Challenge Fund (ABCF) program. The ABCF Program is funded by the Australian Department for Foreign Affairs and Trade (DFAT) through the BecA-CSIRO partnership; the Syngenta Foundation for Sustainable Agriculture (SFSA); the Bill and Melinda Gates Foundation (BMGF); the UK Department for International Development (DFID) and; the Swedish International Development Cooperation Agency (Sida).

Prof.Morris Agaba: this work would have never come up to end in the absence of your all rounded support. I am very much confident that I am somewhere much better today than where I was. The secret is you. Thank you!!! I would also like to extend my appreciation to Firew Tegegn (PhD) for your all rounded and unlimited advice and reviewing the entire document.

Workneh Ayalew (PhD): I have no words to explain; I would rather like simply to say "your name explains far better than I would say". Yes you are right that *"effecting trans-period work helps for tomorrow's to whom be"* though I am not quite sure whether my efforts demonstrated in this work qualifies to your iconic words. Gash Alemayehu Mengistu: your frequent "back stoppings" were very much helpful, and I am very much grateful for that.

My special thanks goes to Aynalem Haile (PhD), Solomon Gizaw (PhD), Azage Tegegne (PhD), Girma Tesfahun (PhD) and Wubishet Dessie for their unreserved advice and support during my study. I would also like to express my heartfelt thanks to Rahel Misganew, Tigist Endeshaw, Goshu Cherinet, Yetinayet Mamo, Michael Temesgen, Grum Gebreyesus, Wondimeneh Esatu (PhD), Tatek Woldu, Temesgen Jemberie, Alubel Alemu, Netsanet Zergaw and Ayda Tegenu for your all rounded support in my stay at ILRI. Eshetu Zerihun, the field and highway manager, wow! Your manoeuvring ideas and advices fetched from your long years experience have helped me to tillage the road map of my career like '*augur buzzard'* birds. Thank you!

Sarah Osama, you are really honest and always close to fellows to support. Your contribution for my career is really a lot. Thank you once again. Wellington Ekaya (PhD), Valeria, Marvin, Mercy, Immaculate Shiro, Collins, Ben, Lucy, Joyce, Martina, Moses, Francesca Stomeo (PhD) and all BecA family and BecA alumni, thank you all!. Ben and Wellington (PhD): thank you very much for your close support, understanding and advice.

Prof. Wenguang Zhang: you have brought me many steps forward in genomics. Thank you very much!!! Bin Liu, Bai Xue, Ji Xiao Yang and all family of college of Animal Science at Inner

Mongolia Agricultural University for your unreserved support, guidance and all your welcoming. I enjoined my stay at Hohhot. Thank you.

Amare Dejenu and Worku Wubie: you are always in my heart; you have proven that any orphan can contribute something good for the world if someone puts something in the life of the orphan. Please receive my limitless appreciation and thanks. Yayehirad Fekadu (my elementary school teacher: 28 years ago), you know how to sow good seed at infant brains. Thank you! Mr.Gebrie Alebachew (my high school biology teacher): trust me that you have not retired in genetics and genomics. I proudly speak that you are seeing your needs and aspiration on genetics and genomics via your student. Thank you!

My appreciation also goes to the department of Microbial Cellular and Molecular biology Addis Ababa University and its staffs offering me the admission and required support; department of Animal production and Technology Bahir Dar University offering me study leave and its staffs and colleagues/friends: Dr.Hailu Mazengia, Dr.Mussie H/Melekot, Dr.Assamnew Tassew, Dr.Yeshambel Mekuriaw, Dr.Kefyalew Alemayehu, Damtie Kebede and Fentahun Mihret for their unreserved support and encouragement during my study.

Colleagues and my PhD study classmate: Mengistie Taye, Adebabay Kebede, Shigdaf Mekuriaw, Demeke Mewa, Zewdu Edea (PhD), Solomon Abegaz (PhD), Tesfaye Getachew (PhD), Berhanu Kumbi and Fekadu Gadissa, I have no words to explain. I would rather say let God pay you back. Thank you!

Finally, I would like to extend my deep appreciation to my family: my lovely wife, w/ro Tsigereda Ayalew, you took all the responsibility of family affairs throughout my study. Let God help me to pay you back. My lovely kids, Anania and Azaria, I feel sorry that I did not share the love you were supposed to get from your dad during the study period. W/ro Meazashwork and Kirubel Ayalew, your contribution for my study has special place in my heart. Thank you!!!

ii Dedication

In **memory** of the 21st century Martyrs (30) who scarified their lives in Libyan dessert in May 2015: only because of their Honesty to God, for the faith, who became witness to Ethiopian Orthodox Tewahido Church

,

my father, Mekuriaw Tarekegn (1940-2005) and my mother, Yetemegn Ayenew (1950-1995);

Dedicated to my lovely wife: W/ro Tsigereda Ayalew

,

my Kids: Anania Getinet and Azaria Getinet

and

to the poor farmers who are keeping the indigenous goats for the rich!

iii. Acronyms

AMOVA	Analysis of MOlecular Variance
Arc	Aarcuate nucleus
AVPV	Anteroventral periventricular nucleus
BMP15	Bone morphogenetic protein 15 gene
BMPRIB	Bone Morphogenetic Protein Receptor-IB gene
CAD4	Cinnamyl Alcohol Dehydrogenase
CHIP	Chromatin Immunoprecipitation
CV	Cross-validation
Cyt b	Cytochrome *b*
D-loop	Displacement loop
DNA	Deoxyribo Nucleic Acid
DnaSP	DNA Sequence Polymorphism
Dyn	Dynorphin gene
ERE	Estrogen Response Element
ERα	Estrogen receptor α isoform
FAO	Food and Agricultural Organization
FSH	Follicular Stimulating Hormone
GDP	Gross Domestic Product
GH	Growth Hormone
GnRH	Gonadotrophin Releasing Hormone
GPR54	G-protein coupled receptor ligand 54 gene
HVI	Hyper Variable region I
HWE	Hardy-Weinberg Equilibrium
IBC	Institute of Biodiversity and Conservation
INHA	Inhibin Alpha-subunit
IUPAC	International Unit for Pure and Applied Chemistry
KISS1	Kisspeptin 1 gene
LD	Linkage Disequilibrium
LH	Luteinizing Hormone
LMP	Livestock Master Plan

LSM	Least Square Mean
MAS	Marker Assisted Selection
MHC	Major Histocompatibility Complex
MJ	Median- Joining
MNA	Mean Number of Allele
MRCA	Most Recent Common Ancestor
mtDNA	Mitochondrial DNA
NJ	Neighbour-Joining
PCA	Principal Components Analysis
PCR	Polymerase Chain Reaction
PIC	Polymorphic Information Content
QTL	Quantitative Trait Loci
SCA	South and Central America
SD	Sequence Divergence
SE	Standard Error
SNP	Single Nucleotide Polymorphism
SSD	Sum of Squares
SSR marker	Simple Tandem Repeat marker

iv. Table of content

v. List of Tables

vi. List of Figures

vii. Abstract

MOLECULAR CHARACTERIZATION OF ETHIOPIAN INDIGENOUS GOAT POPULATIONS: GENETIC DIVERSITY AND STRUCTURE, DEMOGRPHIC DYNAMICS AND ASSESSMENT OF THE KISSPEPTIN GENE POLYMORPHISM

Getinet Mekuriaw Tarekegn

Addis Ababa University, 2016

Ethiopia lies within the earliest geographic center for the dispersal of livestock into the African continent. It therefore comprises a unique repository of livestock genetic diversity in the continent. However, information relating to the origin and divergence of domestic goats in the North-East African region, and Ethiopia in particular, remains unknown despite the availability of genomic tools to assess the same. Here, the complete mitochondrial DNA *D-loop* region of 309 individuals representing 13 Ethiopian indigenous populations of goats was sequenced to investigate the maternal historical demographic dynamics of Ethiopian indigenous goats. Similarly, genotype data were generated from 14 populations by the 50K *Caprine* SNP CHIP array to investigate the genetic diversity and structure of Ethiopian indigenous goats. The Chinese Cashmere, Cashmere-Ibex hybrid and the wild Swiss Alpine Ibex populations were also included as outgroup for comparison. In addition, analysis of polymorphism of the KISS1 gene and its association with litter size, and haplotype and patterns of linkage disequilibrium (LD) analysis were conducted in exon regions of the kisspeptin gene of Gondar and Woyto-Guji goat populations. A total of 173 and 242 Sanger sequences were employed for further analysis of exon1 (1210 bp) and exon2 (325 bp) regions of the kisspeptin gene, respectively. Sequencing of the mtDNA, 174 variable sites that generated 231 haplotypes were observed in the *D-loop* region of mtDNA sequences and these haplotypes defined two haplogroups lacking a phylogeographic

structure but with a high genetic diversity. In a combined analysis of reference haplotypes representing the six globally defined *Caprine* haplogroups revealed out two haplogroups to be A and G, with the former being the most predominant. The lack of phylogeographic structure and a weak population differentiation (overall F_{ST} = 0.0245) suggest extensive gene flow amongst indigenous Ethiopian goats. In the genome wide survey, high genetic diversity ($H_E \geq 0.35$) but low level of genetic differentiation among the populations is revealed. The 14 populations were grouped into six clusters and none of them was specific to one population or geographic region. Interestingly, population phylogenetic analysis did not support the classification of Ethiopian goats based on their agro-ecological location, associated production system and phenotypic family as suggested previously, but differentiated the Kaffa goat from the rest of 13 Ethiopian goat populations. Based on the admixture and phylogenetic network analyses, the 14 Ethiopian goat populations can be re-grouped into seven goat types. On the other hand, polymorphism analysis of the KISS1 gene revealed five complete substitutions and 15 polymorphic sites in both exon regions. The overall average H_E was 0.18863±0.21 for exon1 and 0.03155±0.01 for exon2. Among detected polymorphic sites only four SNPs contributed 18%-31% multiple birth. Similarly, a total of 29 and three haplotypes were detected in exon1 and exon2, respectively. The overall haplotype diversity was 0.8703 for exon1 and 0.0703 for exon2. Lowest (0.083 for Woyto-Guji and 0.081 for Gondar goats) and modest (0.656 for Woyto-Guji and 0.635 for Gondar goats) average estimates of R^2 and $|D'|$ were obtained in LD decay analysis in exon1, respectively, and most of the R^2 and $|D'|$ values suffer *floor* and *ceiling* effects. The neutrality tests showed significant and negative values of F_S for Woyto-Guji (F_S = -8.098) and for Gondar goats (F_S = -12.08); whereas, the Tajima's *D* test was positive and non-significant. Overall, there was high goat population dynamics that caused to have highest level of population admixture.

The KISS1 gene polymorphism analysis showed importance of the gene for multiple birth in Ethiopian goats and can be suggested for marker assisted selection breeding program.

Key words: Ethiopia, Genetic diversity, Haplogroup, High density SNP Chip, Indigenous goat, Kisspeptin gene, Population expansion, Neutrality test

CHAPTER I

1 General introduction

In Ethiopia, there are more than 29 million goats (FAOSTAT, 2014; accessed on February 25, 2016) which are kept for milk and meat, mainly for family consumption, and inhabit a wide range of environments, extending from tropical to cool temperate climates. Goats in the lowlands are kept in large flocks by pastoralists; whereas, there is a strong complementary relationship between small ruminant keeping and cropping in high lands of the country (Tibbo, 2006). Goats account for a large contribution of livestock to the countries' economy estimated at 47% of the agricultural GDP (Behnke, 2010).

Goats contribute about 12% of the total annual meat production and together with sheep, they contribute about 90% of the live animal/meat and 92% of the total skin export trade value (FAO, 2004). In spite of this contribution, relative to the livestock population size, the foreign exchange earnings from livestock and livestock products at large are still lower than expected (Berhanu et al., 2007). In comparison, China ranks as the largest producer of Cashmere in the world with an annual production of 11,057 tones; about 30% of this production comes from 13 million Cashmere goats found in the Inner Mongolia which is home to one of the ten Cashmere goat populations in China (Zhou et al., 2003). Despite the huge goat population, on top of insufficient efforts on the identification and structure of the goat populations, one possible contributing factor for minimal benefit could be the absence of a clear strategy to improve livestock production and productivity in Ethiopia (LMP, 2015).

With regard to the study of origin and genetic diversity, investigation of maternal DNA based genetic diversity and nuclear DNA information using high density SNP panel of the indigenous goat populations in Ethiopia remains uncovered till now. Two decades ago, information compiled on physical description and management system of goats revealed 13 goat types (FARM-Africa, 1996) and later, the microsatellite marker analysis showed only eight goat types (Tesfaye, 2004). However, the latter report was based on D_A distance with low bootstrap values; moreover, the admixture analysis, which is not well resolved, did not support the eight goat types classified. Despite the classification difference, the existence of such a large gene pool is important potential for future breed preservation and development of a sustainable animal production system (Mahmoudi et al., 2011).

Archaeological evidences indicated that goats were the first ruminant animal to be domesticated in 10,500 years ago around the Fertile Crescent (Zeder and Hesse, 2000; Luikart et al., 2001; Fernández et al., 2006; Zeder, 2008). However, there are still some studies which suggest second origin of domestication (Chen et al., 2005). There were two suggested wild species of the genus *Capra* (*C. aegagrus* and *C. falconari*), with the closest candidate *C. aegagrus*, which domestic goat gene pool was derived from (Mannen et al., 2001). Analyses of the control region (the displacement-loop) of mitochondrial DNA (mtDNA) and nuclear DNA are most useful examinations and are informative genomic elements for explicating the origin, diversity, genetic relationship and diversification of livestock including goat (Dorji et al., 2010). The former is because of its variable nature and structured enough across the geographic range of species (Naderi et al., 2007), and evolve at a constant rate (Bruford et al., 2003). Moreover, it allows maternal lineages to be followed and is less sensitive to introgression from wild species than

nuclear DNA (Luikart et al., 2001). Whereas, the latter has highest proportions of non-coding and coding regions which are appropriate for the study of genetic diversity and population structure (FAO, 2007a).

It is believed that genetic diversity has been shaped by past population processes and will also affect the sustainability of species and populations in the future (Soule, 1987). Maintenance of genetic diversity in livestock species requires adequate implementation of conservation priorities and sustainable management programs, which should be based on comprehensive information regarding the structure of the populations, including sources of genetic variability among and within breeds (Mahmoudi et al., 2011). It is also a key to the long-term survival of most species (Hall and Bradley, 1995; Väli et al., 2008) and widely used to categorize animals in the world (Cardellino and Boyazoglu, 2009). Farm animal genetic diversity, in particular, is required to meet current production needs in various environments, to allow sustained genetic improvement, and to facilitate rapid adaptation to changing breeding objectives (Kumar et al., 2006; Glowatzki-Mullis et al., 2008; Kevorkian et al., 2010). However, classifying the genetic diversity based on historical, anthropological and morphological evidences (Ali, 2003) as well as their geographical origin is not satisfactory and consequently is not enough for the purpose of conservation and utilization of these resources. Hence, comprehensive knowledge of the existing genetic variability is the first step for the conservation and exploitation of domestic animal diversity (Li et al., 2002).

Goats, among the livestock species, are considered the most prolific ruminant especially under harsh climatic conditions (Yadav and Yadav, 2008). The high versatility, moderate size and

hardy nature of goats made them ideal as a food resource in the lengthy commercial and exploratory journeys that took place in the old world long time ago (Amills et al., 2008). Besides, goat has good adaptability, wide geographic distribution, and very abundant breed resources (Li et al., 2006). Amills et al. (2008) and Luna-González et al. (2012) also stated that *Caprine* breeding represents an alternative for meat, milk and fur production mostly in arid, semi-arid and mountainous areas of the world. However, following the limited information on goat populations globally acquired, the absence of well-managed conservation genetics programmes and the uncontrolled introgression between indigenous as well as foreign breeds are seriously threatening the future of many populations in various parts of the world (Pariacote, 2006). The absence of fundamental information on the genetic diversity, structure, origin and demographic dynamics is significantly seen on Ethiopian goat populations in contrast to the country which served as main gate for livestock entry into the continent.

It is believed that, in Ethiopia, social/historical as well as natural events happened at various times and contributed to the wider, and mixed, coverage of livestock species including goats. For instance, massive physical movement of the people/tribes (due to war, expansion and settlement) together with their animals frequently happened since the end of the first millennium (Yilma, 1967). Moreover, the recurrent droughts the country faced in 13-14 years cycle for the last 600 and before (Girma, 1988) obliged the people to move from center of their residence to new environments. These two major sources of events could have influenced the current genetic architecture of the animals and implies that the goat populations might have gone through a continuous change of their genetic structure (Getinet et al., 2005). However, in connection to this, no information has been reported using maternal DNA and high density SNP panel.

On the other hand, it is known that reproduction is crucial economic trait in animal husbandry (Zhang et al., 2011). Genetic studies have indicated that reproduction traits can be genetically determined by the action of genes (Deldar-Tajangookeh et al., 2009). However, the identification of candidate genes that are responsible for variation in fitness (e.g. litter size) and continuous traits have been a challenge in modern genetics (An et al., 2013). As effect, little has been known on the major genes associated with litter size and other reproductive traits in livestock species. There are only limited number of reports conducted on polymorphism analysis of candidate genes and their association with multiple births in goat and sheep. These are kisspeptin (KISS1) gene with its receptor G-protein coupled receptor ligand (GPR54) gene (Messager et al., 2005; Feng et al., 2009; Cao et al., 2010; Hou et al., 2011; Huijbregts et al., 2012; Christina, 2013), inhibin alpha-subunit gene (INHA) (Hua et al., 2008; Wu et al., 2009), the gonadotrophin releasing hormone receptor gene (GnRHR) (An et al., 2009), the bone morphogenetic protein receptor-IB gene (BMPRIB) (Polley et al., 2009) and the bone morphogenetic protein 15 gene (BMP15) (Chu et al., 2007). In Ethiopia in particular, there is no any report carried out at gene level associated with the production or reproduction traits in all the livestock species to support the breeding schemes intervened with molecular techniques, like marker assisted selection, except a recent report which was conducted on α_{S2}-casein gene polymorphism analysis on 10 Ethiopian goat populations (Mestawet et al., 2013).

1.1 Statement of the problem

Despite the significant contribution of livestock to the economy of the country, little attention has been given to identify, characterize and conserve the diversity of indigenous livestock types in Ethiopia (Fedlu et al., 2007). As a result, only limited activities have been conducted on genetic characterization of the indigenous goats (IBC, 2004). Tesfaye (2004) characterized 11

Ethiopian indigenous goat populations using SSR markers; Chenyambuga et al (2004) and Hassen et al (2012a) focused on genetic diversity study of only two (Western highland and Arsi-Bale goats) and six (Agew, Gumuz, Bati, Abergelle, Central Abergelle, Begemidir goats) Ethiopian indigenous goat populations with the same marker type, respectively. Recently, Solomon (2014) carried out genetic diversity and population structure of two (Gumuz and Abergelle goats) indigenous goat populations using high density SNP CHIPs array. At gene level, Mestawet et al (2013) conducted α_{S2}-casein gene polymorphism analysis on 10 Ethiopian goat populations.

Therefore, these limited works imply that little is known about the genetic diversity, structure and degree of admixture within and between Ethiopian indigenous goat populations. Consequently, they do not allow comparative analyses; the fragmented information is subjective and perhaps inaccurate which can make the implementation of rational and effective conservation and utilisation strategies difficult (Rege, 1999). On the other side, this has also an implication of the presence of terrible risk that most breeds may perish before they have been exclusively recognized and exploited (Hoda et al., 2012). In connection to this, the genetic diversity of many goat populations is being rapidly eroded globally due to the consequence of increasing farmland abandonment in marginal areas (FAO, 2012) and 18% of local goat breeds over the world were threatened or already extinct (http://faostat.fao.org/).

Moreover, except the recent work conducted on α_{S2}-casein gene (Mestawet et al., 2013), there is no any effort carried out on polymorphism analysis at gene level yet that could eventually provide useful information for further marker assisted selection (MAS) breeding.

Overall, extensive survey on the genetic architecture of the indigenous goat populations is paramount, and this will provide further application of various molecular approaches for better and sustainable utilization of the animal genetic resources.

1.2 General objective

- To characterize Ethiopian indigenous goat populations using molecular techniques

Specific objectives:

1. To assess demographic expansion and identify maternal origins of Ethiopian indigenous goat populations using mitochondrial DNA control region (*D*-loop)
2. To evaluate genome-wide genetic diversity and structure of Ethiopian indigenous goat populations using high density SNP CHIP array
3. To assess polymorphic sites on kisspeptin gene and their contribution for multiple births in selected Ethiopian indigenous goat populations
4. To evaluate the haplotype diversity and extent of linkage disequilibrium of detected loci in the kisspeptin gene

CHAPTER II

2 Literature review

2.1 Genetic diversity and its importance

In a given population, genetic structure patterns are commonly explained by various factors. Some of these factors, which disrupt the gene flow, are isolation-by-distance (Wright, 1943), historical and geological factors (Gübitz et al., 2000), physical barriers (Nicholls and Austin, 2005; Trizio et al., 2005) and ecological factors through morphological adaptation to local conditions (Brown and Thorpe, 1991; Whiteley et al., 2004). In most cases, especially in domestic animals, the gene flow disruption is overseen more by human intervention than by physical barriers (Gizaw et al., 2007). Genetic isolation of populations leading to reduced effective population size and further divergence might be also caused by local management (Cañón et al., 2006) and cultural separation (Rege, 2002). What so ever contributing factors are present, genetic diversity is the basis for the present diversified living organisms. This diversity needs to be properly utilized, improved and conserved. Conservation and improvement strategies ought to be based on proper genetic characterization in association with phenotypic evaluation (Tadelle, 2003; Halima, 2007).

Genetic characterization requires knowledge of genetic variation that can be effectively measured within and between populations; and it is considered as an initial step in considering the sustainable management or conservation of a particular population (http://www.arc.agric.za/home.asp?pid=567). Differences in ancestral origins and migration

events are important causative factors explaining genetic differences between current populations (Alvarez et al., 2004; Rendo et al., 2004).

The most widely used methods to quantify genetic diversity are phenotypic characters, and biochemical and molecular markers (Msoffe et al., 2001 and 2004). However, though phenotypic characters are cheap and easy to apply, they are subjected to environmental influences due to the nature of the qualitative and quantitative traits to be considered (Karp et al., 1997; http://www.fao.org/biotech/logs/c13logs.htm). In addition, the appropriateness of phenotypic traits to study the genetic variation between populations is very limited (Meghen et al., 1994). Similarly, biochemical markers, like isoenzymes, are poor in the nature of polymorphism (Meghen et al., 1994). As a consequence, in the absence of proper genetic identification, the risk of loss of surviving genetic diversity is very high (Fedlu et al., 2007; Hoda et al., 2012).

A decade ago, it was suggested that outcomes of morphological characterization need to be complemented by genetic characterization which involves the description of breeds in terms of the relative allelic frequencies, genetic distances, degree of polymorphism using a set of neutral and non-neutral reference molecular markers (FAO, 2007a). Since early 1990s, molecular markers have played a leading role in the characterization of diversity and provided relatively rapid and cheap assays in the absence of quality phenotypic measures (Toro et al., 2006). Thus, following the advent of polymerase chain reaction techniques, the use of DNA has become an alternative for the research of various genetic, breeding and physiological questions in animal sciences (Hoffman et al., 2004).

As a general guiding decree, ideal genetic markers for population and evolutionary studies should be abundant and distributed widely across the genome (Sunnucks, 2000).Genotyping and sequencing technologies have played significant role for better understanding and utilization of genomic information for different purposes in various ways. Thus, the analysis of genomic DNA variation among and within populations serves as an important initial guide to develop conservation strategies (Davila et al., 2009). This allows the optimum utilization of farm animal genetic resources and permits efficient genetic improvement for production and conservation needs. Moreover, such studies help to design and implement improvement programs in the context of unique quality/ies a population has (Dadi et al., 2008).

2.2 Origins and domestication of goat

2.2.1 Origin and source of goat domestication

2.2.1.1 Origin of goat domestication

Goat domestication was an integral part of the rise of agriculture (Fernández et al., 2006) and the adoption of agricultural practices throughout much of the world (Luikart et al., 2006). Goat, the *''poor man's cow''* (MacHugh and Bradley, 2001), was certainly the first ruminant to be domesticated along with their close relative sheep (Devendra and Mcleroy, 1982; Melinda et al., 2006). It is believed that the goats might have been domesticated in high, rocky mountain regions extending from the Taurus Mountains of Turkey into Pakistan (Epstein, 1971) about 10,500 years ago (Zeder and Hesse, 2000), and then spread quickly following patterns of human migration and trade (Luikart et al., 2001; Fernández et al., 2006). However, the exact location of domestication still remains uncertain (Sardina et al., 2006). Payne and Wilson (1999) reported South-west Asia (Iran and Iraq: the most likely origin of domestication area of goats), south of

Levant (Horwitz et al., 2000) and Mehrgarh (Sultana et al., 2003) to be the ancient centers of goat domestication. The analysis of ancient goat DNA from Inner Mongolia region was closely genetically related to Chinese modern goats suggesting China is also considered the possible center of domestication particularly for sub-haplogroups B1 and B2 (Han et al., 2010). A recent study conducted on mtDNA hypervariable (HVI) region of ancient DNA indicated that Central Zagros has possibly played a key role for domestication of *C.hircus* (Mazdarani et al., 2014).

In contrast, the Balkans or Carpathian Mountain regions of Romania and Southern France were also suggested to be the origin of goats following a divergent lineage C and distinct lineage 3 found in Switzerland and Slovenia (Fernandez et al., 2006; Luikart et al., 2006; Pereira et al., 2009). However, the limited sample size casts doubts on these suggested regions to be the other centre of domestications of *C.hircus*. In addition, this is contrary to the hypothesis of domestication stated by Luikart et al. (2001) for lineage C, and is far from putative domestication centers (Naderi et al., 2008) and questions the previous premises of domestication in general. Moreover, presence of lineage A and C South-east and Central Europe could be accompanied by the first Neolithic migration waves (Colli et al., 2015).

2.2.1.2 Sources of wild gene pool of domestic goat (*Capra hircus*)

Historical and archeological evidences indicate that the domestic goats could have been domesticated from two wild *Capra* species (*C.aegagrus* and *C. falconeri*) (Epstein, 1971), and from markhor (*C. falconeri*) in West Asia and the ibex in East Asia (Harris, 1962). However, it was, earlier, proposed that bezoar (*C.aegagrus*) is the most likely ancestor of domestic goats (Harris, 1962; Zeuner, 1963). The mtDNA analysis strengthened this idea: at least four different strains of wild *Capra* might have been the source of the modern domestic goats (Sultana et al.,

2003). The three species of the wild goat, bezoars (*C. aegagrus*), markhors (C. *falconeri*) and ibex (*C. ibex*), are closely related to the modern domestic goat (*C. hircus*) and especially the bezoar goat is the closest and likely be matriarchal ancestor of domestic goats (Mannen et al., 2001; Sultana et al., 2003). A recent extensive whole mitochondrial genome analysis confirmed the bezoar (*C. aegagrus*) is the most contributor for formation of the identified haplogroups of *C.hircus* (Colli et al., 2015). Sindh Ibex (*C. aegagrus blythi*) was also indicated as a possible contributor to the genetics of domestic goats (Sultana et al., 2003). Luikart and his colleagues also concluded the presence of multiple maternal origins of goats (Luikart et al., 2001); however, the monophyletic and paraphyletic trees obtained (Naderi et al., 2008) do not support multiple origins.

On the other hand, three goat lineages arose from genetically discrete populations rather than from a single wild population and the possible multiple maternal lineages could have been originated via introgression rather than separate domestication events (Luikart et al., 2001). This idea strengthens the paraphyletic tree nature rather than monophyletic tree. The three distinct lineages could be related to either (*i*) three separate maternal origins from genetically distinct populations, or (*ii*) one origin from an extremely large population containing three highly divergent lineages. However, all the domestic goat lineages (A, B, C and D) examined in Indian goat populations fall into a single monophyletic group that is distinct from all available wild goat sequences (Joshi et al., 2004), and the authors hypothesized that the contributing lineages found in India were derived from an unknown population that might have become rare or extinct.

There are still discordances between mtDNA and Y-chromosome phylogenies in which the intimations are explained. According to Pidancier et al. (2006) the following remain unresolved: i) amplification of nuclear-mtDNA copies; i.e. laboratory artifacts and in most cases authors use many bone samples, for which nuclear amplification is less probable than mitochondrial amplification because of DNA quantity and quality ii) selection, iii) lineage sorting of ancestral polymorphisms or iv) horizontal transfer of genes which may result from hybridization and introgression in mammals. In relation to the latter hypothesis, the mtDNA control region and Y-chromosome analyses indicated the possible case of recent introgressive hybridization in *Capra* between *C. cylindricornis* individuals from Daghestan groups and Daghestan *C. aegagrus* rather than with its conspecifics (Pidancier et al., 2006).

In conclusion, in spite of some contradiction, the closest possible wild source of the present domestic goat is the bezoars (*C. aegagrus*). This is also supported by the following evidences: the branch length between the *Cyt b* and mtDNA control region is shortest from domestic goats to the wild goat, *C. aegagrus* (Manceau et al., 1999; Luikart et al., 2006). The mtDNA analysis revealed that the domestic goat originated from Bezoar goat (*C.aegagrus*) (Takada et al., 1997; Manceau et al., 1999; MacHugh and Bradley, 2001; Colli et al., 2015). These are consistent with the Y-chromosome and autosomal (microsatellite) marker based findings of Luikart et al. (2006) and Pidancier et al. (2006), as with those of morphological studies, archaeological data, and inferred geographical distribution of wild *Capra* species (Smith, 1998). In addition to the above maternal and paternal origin evidences, the paleontological evidence also supports the *C. aegagrus* to be the closest ancestor of domestic goats (Porter, 1996). The second-closest taxon to

domestic goats, based on the Y chromosome, is *C. falconeri*, which is a species separated from both domestic goats and *C. aegagrus* by two to three mutations (Luikart et al., 2006).

2.2.2. Identification of lineage, dispersion routes and global coverage of domestic goats

2.2.2.1 Identification of lineages

There is no clear definition between lineage and haplogroup. While Luikart et al. (2001) and Sultana et al. (2003) put both alternatively. Nomura et al (2013) indicated lineage is source of wild ancestor whereas haplogroup is common ancestor. For this paper, both terminologies have been used interchangeably. Bearing this in mind, various scholars have identified six lineages of domestic goat which are dispersed throughout the world following various routes of dispersion at different times. Luikart et al. (2001) identified three lineages (A, B and C) by sequencing hypervariable region I (HVI). Sultana et al. (2003) revealed four lineages (A, B, C and D) by sequencing both D-loop and *Cyt b* regions in Pakistan's goats. Joshi et al (2004) revealed five lineages (A, B, C, D, E) in Indian goats. Naderi et al (2007) identified six lineages (A, B, C, D, F and G) by sequencing HVI and disproved existence of haplogroup E rather those haplotypes which were named by this haplogroup created sub-haplogroup B1 and B2 which were moved to North, East and South East Asia. The presence of sub-haplogroups was confirmed later (Han et al., 2010, Nomural et al., 2013; Akis et al., 2014 and Colli et al., 2015). The ancient DNA analysis indicated goats from haplogroup B were detected in the Swiss Alps which later replaced by haplogroup A and C (Schlumbaum et al., 2010). Chen et al. (2005) had also found four mitochondrial lineages (A, B, C and D) in Chinese goat breeds. Similarly, by amplifying HVI, Nomura et al. (2013) confirmed the presence of all previously identified lineages/haplogroups except lineage G. This could be because of: i) the divergence regions of the latter lineage was

only towards to South-west Asia and Europe, ii) their limited focus to South, East and South east Asia. A recent and extensive study conducted on whole mitochondrial genome revealed various clades of haplogroups (A1-A7, B1, C1a, D1 and G) (Colli et al., 2015).

However, based on the microsatellite markers, the three lineages identified by Nomura et al. (2012) differ slightly from those reported by Luikart et al. (2001) particularly for the Asian goat populations. Nomura et al. (2012) investigated two different lineages, which were dispersed to South and South-East Asian countries. However, these haplogroups were considered earlier as a lineage (Luikart et al., 2001). On the other hand, the two different lineages identified by Luikart et al. (2001) which were spread to North (Mongolia) and East Asia regions were merged and reported as a lineage (Nomura et al., 2012). Besides, an additional lineage which was moved to South-East Asia (including Taiwan, Japan and Korea) was also identified. In general, wider wild origins/lines is found in Asia than other parts of the world, as a result the regional genetic diversity is also comparatively higher as discussed below in detail.

2.2.2.2 Dispersion routes and global coverage of domestic goat

The domestic goats had been dispersed following various routes of divergence globally from their initial domestication areas. It had followed Mediterranean and Danubian routes to arrive in Europe and was aligned with the routes of Neolithic culture diffusion in the region (Fernández et al., 2006). Civilizations like Phoenicians, Greece, Romans and Berbers probably introduced new species of animals and new breeds of livestock in South-west Europe following the sea route (Pariset et al., 2009b). The archaeological data and radio carbon dates on bones in Western Europe indicated that goats had arrived earlier through Mediterranean route compared with the

Danubian route (Zilháo, 2001; Voruz, 1999; Guilaine, 2003). In Asia, dispersion of the three types of lineages from the domestication center followed two main routes (the *Silk* Road and the *Khyber* Pass) (Devendra and Nozawa, 1976). The latter route was one of the known *Silk* Roads in the world found between Afghanistan and Pakistan and served for the migration of the Nubian goat type, which had descended from the Savannah type, to Indian sub-continent. Similarly, the former route to Asia served for expansion of both Bezoar-type and Savannah-type goat (Devendra and Nozawa, 1976).

Based on the microsatellite evidence, the East Asian cluster corresponded morphologically to the Bezoar type and the Mongolian cluster corresponded to the Savannah type (Nomura et al., 2012). Taiwan goats are direct descendants of Chinese indigenous goats during the seventeen century by immigrants, and the Savannah type reached back to Mongolia from the Indian subcontinent and China (Nomura et al., 2012). The genetic subdivisions of East Asian goats were consistent with the migration history of goats and also with morphological and geographical classifications (Nomura et al., 2012). Amills et al. (2008) had fairly addressed wide geographical distribution of the populations and reported the existence of genetic variation at continental level despite smaller sample sizes used in many of the populations studied. The haplogroups of the wild bezoar did not decline in population size since the Early Holocene suggesting the bezoar populations were not modified so much by humans (Naderi et al., 2008).

Despite the inherent and unavoidable bias of sampling, haplogroup A is the earliest (~10,000 YA) expanded lineage and is known to occur throughout the world including Africa and parts of Asia; haplogroups F is linked to Europe (particularly in Sicily) and haplogroup D is limited to

Asia (Luikart et al., 2001; Naderi et al., 2007 and 2008; Pereira et al., 2005 and 2009; Han et al., 2010; Hughes et al., 2012). However, the existence of haplogroup E has been disproved and considered to be haplogroup B (B1 and B2) (Naderi et al., 2007). The global coverage of haplogroup A is 89% in Asia, 98% in Europe (Pereira et al., 2005).

However, though Pereira (*ibid*) reported 100% pre-dominance of haplogroup A in Middle East and Africa, Naderi et al (2007) detected haplogroup G (in Egypt, Saudi Arabia Turkey and Iran), haplogroup B in Namibia and South Africa together with haplogroup A. This haplogroup was also detected in Canary Islands and southern and eastern Asian countries: Pakistan, India, Malaysia, China and Mongolia (Amills et al., 2004; Pereira et al., 2005; Luikart et al., 2006; Han et al., 2010; Nomura et al., 2013). The ancient DNA showed that goats from sub-haplogroup B1 were present in alpine areas of Switzerland in 4500 YA (Schlumbaum et al., 2010); and this haplogroup is the result of a second domestication event (Luikart *et al*., 2001) and represents a relatively recent expansion (Pereira et al., 2005). Haplogroups A and C show conspicuous rapid expansion and haplogroups B and G show slow expansions; population size of haplogroup F has been slowly declined (Nomura et al., 2013).

Recent study also indicated that haplogroup A and G are reported in Kenya (Kibegwa et al., 2015). Akis et al (2014 and 2016) also observed haplogroups A, B1, C, D and G in Anatolia region. There is an absolute predominance of lineage A in the Atlantic archipelagos and South and Central American (SCA) (Amills et al., 2008). Lineages B, C, D, F and G are absent in SCA goats (Amills et al., 2008), and are also very rare or even absent in Europe (e.g. haplogroup D) (Luikart et al. 2001; Joshi et al. 2004; Amills et al. 2004; Azor et al. 2005; Pereira et al. 2005;

Naderi et al. 2007). The ancient DNA showed existence of Haplogroup B in Swiss Alps in former times (Schlumbaum et al., 2010). In general, the contribution of haplogroup B, D, F and G in domestic goats is very low (7.69%) (Naderi et al., 2008).

Though the origin and evolution of haplogroup C still remains controversial, it is present with very low frequencies in Europe (<5%) (e.g. Iberian Peninsula, Slovenia and Switzerland), Asia (1%) and in Mongolia which represent recent secondary expansion (Luikart et al., 2001; Pereira et al., 2005).It is also found in Near Eastern populations except in Pakistan (Luikart et al., 2001; Sultana et al., 2003), and recently in Corsica (Hughes et al., 2012) and Anatolia (Akis et al., 2014). This dispersion may suggest older origin (Pereira et al., 2005); however, the sampling employed was less comprehensive. Fernández et al. (2006) also explained both lineages A and C coexist in Europe, and were represented among the first populations of domestic goats that entered into Western Europe. This coexistence of lineages A and C in South-west Europe, since as early as the beginning of the Neolithic, may have resulted from either the succession of different waves of goats bearing different haplotypes between the first Impressa (7,700–7,500 B.P.) and Cardial (7,500–7,000 B.P.) time periods, or from one wave bearing all of the diversity as early as the first Impressa steps (Fernández et al., 2006), which is the first arrival of goats to this region. This finding is consistent with the first waves of arrival of Neolithic farmers (7,500YA) through the Mediterranean route.

Unlike the absence of a strong phylogeographic structure in the Spanish peninsula, European, African and Asian populations, the ancestral Canarias goat mitochondrial haplotypes are still highly ubiquitous in some of the breeds providing a recognizable population structure (Amills et

al., 2004). On the other hand, from the historical perspective Iberian livestock were extensively transported from South of Spain and Portugal to America, and similarly from Portugal, Africa and Canary Islands to Cape Verde by Portuguese sailors during the 15th century (Rodero et al., 1992). The similar haplotypes obtained in Cape Verde with Canary Islands (Amills et al., 2008) can be a very good witness despite the limited contribution of the Atlantic archipelagos to the large-scale population process (Rodero et al., 1992).

However, the mtDNA analysis indicated that the initial goats (i.e. variant B) arrived in the Canary Islands by the first settlers 3000YA (Amills et al., 2004). Capote et al. (2004) had also reported the first inhabitants of the Canary Islands settled at the archipelago carrying a small number of domestic animals in 2200YA. Despite the time variation seen in these reports, the first settlers of the Islands are believed to be the Berber people of Morocco though there is no clear evidence reported yet. Especially the *Caprine* breeds of Canary Islands are likely to have North African origin, and were isolated for 1700 years until Spanish colonization, however, had an important influence in the constitution of the American mosaic of breeds and breed types (Capote et al., 2004). It is also reported that the majority of the Canarian domestic animals prior to the colonization are of virtually unknown origin but assumed to most probably be from the African continent, for instance, the three types of Canarian *Caprine* (Fresno et al., 1992) look like the African relatives. However, mtDNA analysis of Pereira et al. (2009) could not substantiate this assumption of gene flow into the Canary Islands from the Maghreb (North West African countries except Egypt) rather the Y-chromosome analysis. The latter analysis revealed presence of three main haplotypes (Pereira et al., 2009) with the most frequent haplotype Y2 reaching 76.09% frequency in Morocco. Haplotypes Y1A and Y1B occur at 19.57% and 4.35%,

respectively, which is consistent with findings of Amills et al. (2004) though it contradicts the mtDNA analysis of Pereira et al. (2009). In support of the mtDNA analysis, the plot of pair wise F_{ST} genetic distances indicates that the Canary goats are closer to Middle East goat than North Africa goat (Pereira et al., 2009) suggesting the Canary goats diverged from the center of origin via Mediterranean Sea instead of terrestrial routes. This idea can be strengthened by the presence of strong phylogeographic relationships among Canary island populations compared with other regions (Amills et al., 2004).

In general, despite the discordance of inference between the mtDNA and Y-chromosome, male flocks from Asia might have moved via Morocco to Canary Islands. But, still it does not necessary mean the origin of Canarian goat population is only from Africa. The maternal origin has also strong implication about the other origin of Canarian goats to be directly from the center of origin via Mediterranean Sea.

The presence of variant A found in some of the breeds in the Canary Islands (Amills et a., 2004) might be because of the introgression between the native goats (variant B) with other European and African breeds around 500-600YA following the Spanish colonization (Capote et al., 1999). Y-chromosome analysis also supported the presence of bidirectional gene flow between Africa and southern Iberia (Pereira et al., 2009). However, there is no any genetic footprint of Iberian goats rather that of Canarias's obtained in SCA. It is argued that the Iberian populations had a poor phylogeographic structure at the time of the American colonization, and the Canarian goats contributed to the foundation of the current genetic pool of SCA goat breeds (e.g. two Andean populations of Chile and Argentina have descended from Canarian goats) (Amills et al., 2008).

Morphological similarity between Canarian and American goats is the other supporting evidence about the contribution of the Canarian goats to their American counterparts (Capote et al., 2004). In connection with this, there is a high diversity of mtDNA lineages in Moroccan populations with 54 different (all belong to haplogroup A) haplotypes (Pereira et al., 2009) which are similar in number and type of South and Central American goat haplotypes (Amills et al., 2008). A recent study on mtDNA confirmed that all haplotypes detected in Moroccan goats belong to haplogroup A (Benjelloun et al., 2015). Moreover, Pereira et al. (2009) did not report the existence of this variant B in Morocco. However, there is no concrete evidence about either the transportation of goats from the Canary Islands to SCA at a considerable scale or rapidly disseminated in SCA with one or few introduction events (Amills et al., 2008).

Still the point which needs to be clear is that if variant B is found in Canary Islands and all haplotypes found in SCA that belong to lineage A are descendants of the Canary Islands, why variant B is not found in SCA? This could possibly be due to the limited coverage of the study populations in SCA and small sample size used and/or might be because of the absence of examining the divergence from paternal perspective that could probably indicate the connection it would have had with Africa. The other possible reasons might be variant B could have been extinct in SCA or it could have been only lineages A which was transported to SCA. On the other hand, the regional analysis of genetic diversity suggests nucleotide and haplotype diversities are particularly reduced into two Andean populations located in Chile and Argentina compared to Cape Verdean goats implying these two populations descended from Canarian goats (Amills et al., 2008).

Despite the limited molecular data report, the archaeological data indicated that domestic goats were first introduced into the African continent through i) Mediterranean coast ii) Red Sea Hills iii) over land via the Sinai Peninsula and Nile Delta in 7,000YA (Hassan, 2000; Gifford-Gonzalez and Hanotte, 2011) (Figure 1). Similarly, the archaeological data suggested that goats and sheep spread rapidly from the Near East into the Central Sahara and Ethiopian highlands between 6,500 BP and 5,000 BP (Clutton-Brock, 2000) and later expanded to south because of, besides the tsetse barrier, the increasing aridity of North Africa (Smith, 1992). Radiocarbon dates of goat and sheep bones from various archaeological sites along the North African coast (dated 6,000 BP at Grotte Capeletti in Algeria or 6,800 BP at Haua Fteah in Cyrenaica, Libya) are similar to those excavated in the eastern Sahara, suggesting a very rapid dispersal of small ruminants from Southwest Asia into North Africa between 7,000 BP and 6,000 BP (Hassan, 2000). In contrast to this, mtDNA diversity (lineage A) suggested recent time of expansion (<3,000YA) in the African continent via south of Saharan desert (Luikart et al., 2006). The route of introduction into the African region is believed either through the present-day Sahara desert by overland diffusion or along the Mediterranean coast (Hassan, 2000) (Figure 2). The mtDNA and Y-chromosome analyses strengthened the use of both Mediterranean route in the east-to-west movement of domestic goats and the terrestrial transport along the North African continent (Pereira et al., 2009).

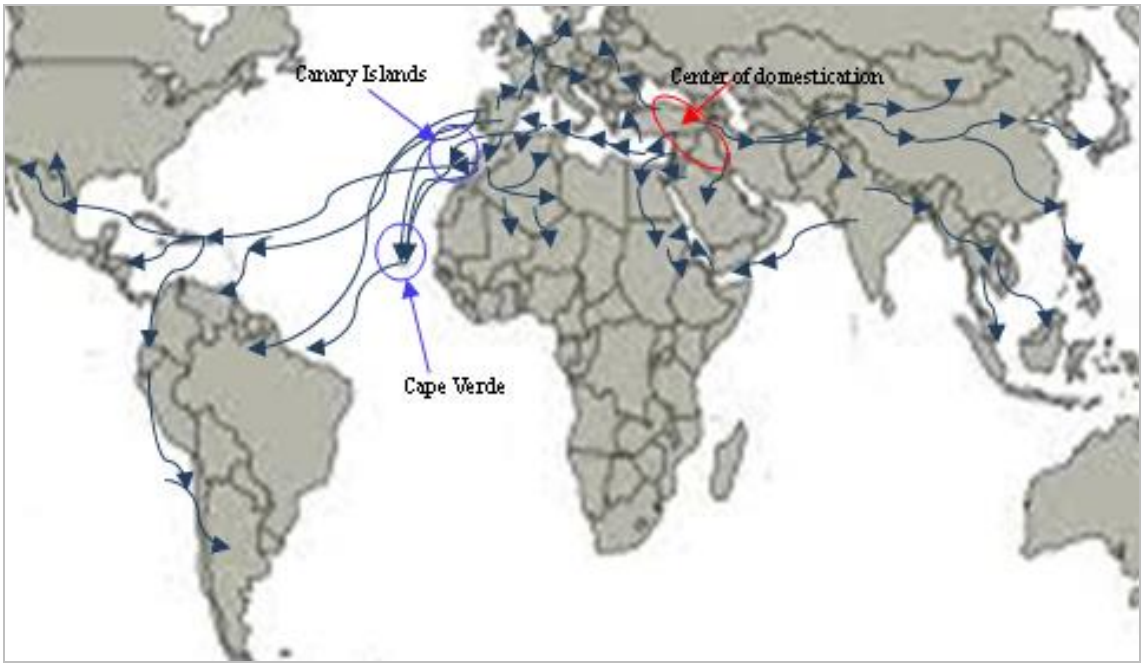
However, there is no indication of median joining network on the movement of domestic goat from Egypt to North Africa towards Morocco rather this route could be extended from Egypt directly to Ethiopia following the Nile Valley. The absence of Egypt's route to North West

Africa seems contradictory with the archeological findings. In general, Figure 2 summarizes the global dispersion routes of *C. hircus* from center of domestication areas.

Source: Gifford-Gonzalez and Hanotte, 2011

Figure 1. Origin and divergence of goat into Africa:based on archaeological information

Figure 2. Summary of global dispersion routes of domestic goat



Source: Getinet et al., nd; *in press*

2.2.2.3 Divergence of *C. hircus* haplogroups: the time and expansion

In most cases, fossils are exposed in regions of geological activity (Pagani et al 2012) and as a consequence, do not always point to the real divergence time. It is because of the fact that the first stratigraphic appearance of taxa in the fossil record may be subject to sporadic sedimentary disruptions due to erosion or lack of sedimentation during regression and/or irregular sedimentary processes (Nomura et al., 2013). In addition, only very few paleontological data are available for species of the genus *Capra* because their preferred mountainous habitats are not favorable for fossil preservation (Simpson, 1945). Instead, scholars have employed various molecular techniques to be able to know time of divergence of organisms since last decade.

Molecular techniques have been used to date divergence of the wild progenitors (Nomura et al., 2013). Sequence divergence (SD), the estimated divergence times between the mtDNA lineages

A and D for goats have been estimated to be from 260,483 to 371,052YA (Sultana et al., 2003). Given the time variation, the estimated times for most recent common ancestors (MRCAs) of each haplogroup (32,300 to 90,950 YA) and the times of nodes with star-like branching pattern (17,210 and 90,950YA) can be indicative for prior expansion of goats (10,000YA) before domestication (Zeder and Hesse, 2000; Nomura et al., 2013). Lineage A is believed to have been expanded initially about 10,000YA, and then the less abundant lineages may have expanded about 6,000 to 6,110YA (lineage C), and 2,130 to 2,600YA (Asian lineage B) (Luikart et al., 2001 and 2006). However, Fernández et al. (2006) indicated the third event of domestication (haplogroup C) dated 7,500YA in Southern France. The latter is however inconsistent with the previous premises in both place of origin and time of domestication.

For Indian goat populations, the MRCA calibrated against the fossil record was 103,000 to143,000 or 201,000 to 280,000YA (Joshi et al., 2004) which agrees with Luikart et.al (2001) report particularly for lineage A. For the goat populations in Pakistan, the new lineage D revealed high sequence diversity (SD) from lineage A and may be the oldest branch under domestication, while lineages B and C showed lower SD and could have been domesticated during an advanced stage of the domestication process (Sultana et al., 2003). Apart from this, the four lineages (A, B, C, D) of C*yt b* indicated that the estimated MRCA of the domestic goat lineages was 427,006 to 597,806YA (Sultana et al., 2003); however, lineage D diverged from lineage A more recently (265,038 to 371,052YA). This finding is strengthened by the D-loop average SD value estimation (4.59%) of the four lineages (A, B, C and D) and 2.8% of lineages A and D, which is the most recent divergence (Sultana et al., 2003). However, this seems too early compared to the mtDNA HVI analysis of Joshi et al. (2004) that showed the three lineages

(A, B and C) had diverged over 200,000YA. Bear in mind that the mt-lineages exhibiting lower (higher) SD could have been captured and adopted at later (earlier) periods of domestication (Sultana et al., 2003). The ancient divergence time and the different geographical localizations of the lineages suggest the likelihood of either multiple domestication events or introgression of additional lineages after the original domestication (Joshi et al., 2004).

2.2.3 Goat populations differentiation and gene flow

Despite the huge global goat population size, the genetic diversity of goats as revealed by maternal mitochondrial and nuclear (microsatellite) DNA marker studies contemporary domestic goats (*C. hircus*) show far weaker intercontinental population structuring than other livestock species (Luikart et al., 2001). The highest proportion (90%) of the current domestic goat mtDNA haplotypes belongs to haplogroup A which could not have been changed dramatically in the expanding goat population since domestication (Naderi et al., 2008). This suggests that haplogroup A goats, could have been dispersed more often, more successfully and more extensively than other livestock (Luikart et al., 2001; Fernández et al., 2006; Nomura et al., 2012). The genetic distance between the Portuguese goat breeds is not positively correlated with the geographical distribution of these breeds (Pereira et al., 2005). This therefore is a very good example for the above argument from a microgeography perspective. Geographically most distant breeds (Algarvia and Bravia) show the lowest genetic distance (F_{ST}=0.020), while the most divergent breeds are Serpentina and Charnequeira (F_{ST}=0.083) with a closer geographical distribution. In addition, the genetic variation estimated by mean number of alleles (MNA) and allelic richness within-country populations of Asian goats was lower than that of European breeds (Nomura et al., 2012). The average F_{ST} (0.13, 0.07) estimates of Asia (Nomura et al., 2012) and European goat breeds (Cañón et al., 2006) strengthened this notion. This lack of

relationship between genetic distances and microgeography can be interpreted to mean and to have resulted from complex and diverse female stocks in the origins of Portuguese breeds and/or extensive successive introduction of extraneous female individuals (Pereira et al., 2005).

From the macrogeography perspective, like horse (Equus caballus) (Kim et al., 1999), there is low mtDNA population structure in domestic goats compared to cattle (Luikart et al. 2001; MacHugh and Bradley, 2001). Only about 10% of the total mtDNA variation in domestic goats (*C. hircus*) was due to differences among continents (Luikart et al., 2006). It is far lower than estimates of 54 to 80% intercontinental variation in cattle for the same mtDNA region (HVI). Investigation of positive values, the converse is also true for the negative values, in all bezoars made in Tajima's *D* (Tajima, 1989b) estimates indicate the presence of population expansion events of bezoars that are closer to domestic goat since recently (Nomura et al. 2013). These all findings imply that geographical location has little relevance to the mtDNA type that a particular animal possesses or the absence of clear tie between the genetic make-up of goats and geography rather at within-population level (MacHugh and Bradley, 2001; Amills et al., 2008). This might also be due to the extensive intercontinental dispersion and high gene flow of goats compared with cattle (Luikart et al., 2006). A relative lack of breed standardization, herdbook breeding, parentage control and rigorous management might have facilitated gene flow between geographically nearby breeds (Cañón et al., 2006).

Moreover, the founder effect has also contributed to the decreased genetic diversity. For instance, upon the conquest and colonization of the New World by the Spanish and Portuguese, goats and other livestock species were massively transported through the Atlantic Ocean for food

in exploratory and military expeditions (Rodero et al., 1992). In Brazil, goats were first introduced by Portuguese settlers during the beginning of 16th century (Machado et al., 2000). This depicts the current gene pool of South and Central America goats were founded in the last five centuries (Amills et al., 2008). Similarly, Mongolian goat populations have the lowest genetic distance in contrary to the geographical distances (Takahashi et al. 2008; Nomura et al., 2012). However, Pariset et al. (2009b) revealed significant and positively correlated genetic and geographic distances.

On the other side, the within population variation estimated values are the other indications of the weak structuring of goat breeds that support the utilization of domestic goats as a portable food resource accompanying human migratory movements (Amills et al., 2008). Hence, around 69% of the genetic variation corresponds to the within-population component for South and Central American goats and almost similar to Iberian and European breeds, but haplotype diversities were somewhat lower (Amills et al., 2008). Apparently, almost similar estimates (78.7% and 77%) of within breed genetic variations were reported in European, African and Asian goat mitochondrial sequences (Luikart et al., 2001; and Naderi et al., 2007). Still this estimation is higher (83%) for Indian goats (Joshi et al., 2004). From the AMOVA analysis 96.65% of variation occurs within breeds, the remaining 3.35% from among breed variation, for Portuguese goat populations (Pereira et al., 2005). However, for the latter report, comparatively high within breed diversity found in all breeds and the sharing of some haplotypes with other foreign breeds is consistent with the repeated introduction of exotic animals into the Portuguese gene pool in the last centuries (Pereira et al., 2005). All the above estimates are very high

compared with the within breed (45%) and amongst-group components of the total variation in cattle (Luikart et al. 2001).

On the other hand, despite the above reports which revealed the weak phylogeographic structure in goat compared to other domestic animals, there is significant mtDNA variation among Indian goat breeds (Joshi et al. (2004). However, this study was limited only to mtDNA and did not include autosomal and/or Y-chromosomal markers, and hence was unable to show the overall gene flow from paternal perspective. The overall diversity and population structure of domestic goat reported based on autosomal markers are summarized in the next section.

2.3 Status of molecular characterization of domestic goats (*Capra hircus*) using autosomal markers

Following the presence of various molecular tools, persuading efforts have demonstrated in the identification and structure of domestic goat populations found in various parts of the world. However, most of the efforts carried out in the past may demand further works to suit for designing appropriate conservation and breeding management intervention (Getinet et al., 2016). This is because of various factors that include the technical fissures observed during the analysis which could have been pursued from the resource limitations during the study. Therefore, studies at genome wide level using high density SNP array is unquestionably required. Some of the studies conducted using microsatellite (for more than 120 goat populations) and SNP markers have been summarised below.

2.3.1 Genetic diversity of goat populations using microsatellite markers

2.3.1.1 Genetic diversity and polymorphic information content (PIC)

Genetic diversity refers to the total number of genetic characteristics in the genetic makeup of a species that serves as a way for populations to adapt to changing environments. It represents diversity within a population (Tesfaye, 2004). With more variation, it is more likely that some individuals in a population will possess variations of alleles that are suited for the environment. Those individuals are more likely to survive to produce offspring bearing that allele. Choosing the appropriate breed or population for conservation is one of the most important problems in the conservation of genetic diversity in domestic animals. Some of the methods which help for the study of genetic diversity within a population are heterozygosity (expected and observed) estimates and allelic distribution; and they are good indicators of genetic polymorphisms within a population (Tesfaye, 2004; Ramamoorthi et al., 2009; Hassen et al., 2012a). On the other hand, the precision of estimated genetic diversity is a function of the number of loci analyzed, the heterozygosity of these loci and the number of animals sampled in each population (Barker, 1994).

The expected heterozygosity is the proportion of heterozygotes expected in a population; whereas, observed heterozygosity is the percentage of loci heterozygous per individual or the number of individuals heterozygous per locus (Ojango et al., 2011). Various reports confirmed the status of genetic variability of different goat populations (Table 1). Hence, relatively lower genetic diversity estimations, expected heterozygosity (H_E) and observed heterozygosity (H_O), were observed in goat of Sri Lanka South (H_E=0.48), Sri Lanka N-Central (H_E=0.49) and Australian goat (H_E=0.45) (Barker et al., 2001), some Korean goats (H_E=0.38, H_O=0.36; Kim et

al., 2002) and Jamunapari goat breed of India (H_O=0.42; Gour et al., 2006). These low estimates imply that there could be high selection pressure, small population size, minimal or null immigration of new genetic materials into the population.

Table 1 Estimation of genetic heterozygosity of indigenous goats

Breed (N)	**Country**	H_E	H_O	**MS**	**Author**
Sri Lanka and Australian goats (12)	Sri Lanka-Australian	0.45-0.49		22	Barker et al. 2001
Korean goats	Korean	0.38	0.36	9	Kim et al., 2002
Indian goat populations	India	0.54-0.79	0.505	17- 25	Fatima et al., 2008; Pramod et al., 2008; Dixit et al., 2009; Kumar et al., 2009
Swiss goats (11)	Swiss3-24	0.66		47	Glowatzki-Mullis et al., 2008
Canary Island goats	C. Islands	0.62		27	Martínez et al., 2004
Kalahari Red goats		0.63		8	Kotze et al., 2004
Sub-Saharan breeds	*	0.54	0.56	11	Muema et al., 2009
Spanish Guadrrama goat	Spain	0.81	0.78	10	Serrano et al., 2009
Croatian spotted goat	Croatia	0.77	0.76	20	Jelena et al., 2011
Chinese ten goat populations	China	0.54-0.64	0.55-0.62	14	Verma et al., 2007; Di et al.. 2010
Brazilian goats and herds	Brazil	0.50-0.70	0.61-0.70	11	Araújo et al., 2006
Guinea Bissau goat	W. Africa	0.60	0.61	14	Di et al.. 2010
Iranian goat populations	Iran	0.65-0.80		13	Mahmoudi et al. 2010
Ardi	S.Arabia	0.68	0.55	11	Aljumaah et al., 2012
Twelve Chinese breeds	China	0.61-0.78	0.60 - 0.78	17	Li et al., 2002
Three Egyptian and two Italian goat breeds	Egypt and Italy	0.67-0.79		7	Agha et al., 2008
Tswana goat	Botswana	0.16	0.12	12	Maletsanake et al., 2013
Ethiopian goat populations	Ethiopia	0.55-0.69	0.52-0.68	15	Tesfaye, 2004; Hassen et al., 2012a

Key: N = number of populations**;** MS = number of microsatellite; * = Uganda (4), Tanzania (5), Kenya (2), Mozambique (2), Nigeria (3), Mali (1) and Guinea Bissau (1)

However, moderate and higher estimates of genetic diversity were reported for some of the goat populations (Table 1). The estimates indicate that the populations studied have substantial and high amount of within population genetic diversity. This might be due to low selection pressure, large population size and immigration of new genetic materials (Aljumaah et al., 2012). In most

of the above diversity estimates, the observed (H_O) and expected heterozygosity (H_E) estimates for each locus and goat population are closer to each other indicating no overall loss in heterozygosity (allele fixation) (Araújo et al., 2006). On the other side, Saitbekova et al. (1999) evaluated diversity among nine domestic Swiss goat herds, Wild Ibex goats and Bezoar goats; and the heterozygosity was higher in the domestic breeds than in the wild goats, with the mean H_E ranging from 0.51 to 0.58 for domestic herds and from 0.17 to 0.19 for the wild species. The lowest heterozygosity estimates in the wild goats are comparable with the mean observed (H_O=0.12±0.16) and expected heterozygosity (H_E=0.16±0.20) values of Tswana goat breed (Maletsanake et al., 2013) which is because of the effects of inbreeding and selective breeding in small and closed population. Population size with heterozygosity estimates are positively correlated (r = 0.35) (Caňón et al., 2006). Low amounts of genetic diversity increases the vulnerability of populations to catastrophic events such as disease outbreaks, indicates presence of high levels of inbreeding with its associated problems of expression of deleterious alleles or loss of over-dominance, can destroy local adaptations and break up co-adapted gene complexes that ultimately leading to the probability of population or species extinction (Mahmoudi et al., 2011).

The allelic distribution is the other measure of genetic variability in a given population (Ramamoorthi et al., 2009; Aljumaah et al., 2012; Maletsanake et al., 2013). However, in the absence of rarefaction technique, the allelic distribution is highly influenced by sample size: large samples contain more alleles than small samples (Kalinowski, 2004). The following table (Table 2) summarizes mean number of alleles (MNAs) and polymorphic information content (PIC) observed in domestic goats found in various parts of the world.

Hence, most of the goat populations addressed showed MNAs above the suggested minimum estimation. However, the MNAs in some goat populations in Ethiopia, Brazil, Egypt, Italy and Iran goat were below the minimum suggested value. This could be because of: i) using very few sires, e.g. 3-5 sires per year for Tswana goat for 16 years of almost closed breeding program at BCA farm (Maletsanake et al., 2013); ii) directional selection for parasite resistance/tolerance coupled with increased productivity (Nsoso et al., 2001) that possibly accumulates inbreeding, iii) insufficient microsatellite screening techniques: e.g. from 26 loci employed for twelve Chinese goat diversity analysis nine of them had less than four loci or non specific PCR products (Li et al., 2002). For studies like genetic distance, microsatellite loci should have no fewer than four alleles to reduce the standard errors of distance estimates (Barker, 1994).

Moreover, heterozygosis deficiency may be resulted because of the presence of a null allele (the allele that fails to multiply during PCR using a given microsatellite primer due to a mutation at the primer site (Callen et al., 1993; Pemberton et al., 1995), due to small sample size where rare genotypes are likely to be included in the samples (Mahmoudi et al., 2011), due to the Wahlund effect: that is presence of fewer heterozygotes in population than predicted on account of population subdivision and due to the decrease in heterozygosity because of increased consanguinity (inbreeding) (Kumar et al., 2006). Higher heterozygosity provides better assignment performance (Manel et al., 2002) and the loss of alleles is probably the consequence of repeated founder effects during migration events (Cymbron et al., 2005).

Table 2 Estimated mean number of alleles and polymorphic information content

Breed (N)	Country origin/Region	MNA per breed	MNA per MS	PIC per locus	Reference	MS
Egyptian and Italian goat breeds (5)	Italy	6.48	3.8-9.8	0.22 -0.87	Agha et al., 2008	7
Indian goat breeds (10) f	India	6.33-9.7	4-24	0.08-0.90	Aggarwal et al., 2007;Pramod et al., 2008; Kumar et al., 2009; Ramamoorthi et al., 2009	17-25
Taleshi goat	Iran	6.7	2.4-5.2	0.54-0.81	Mahmoudi and Babayev, 2009	9
Iranian goat breeds (6)	Iran	6.46 -8.15		0.71-0.86	Mahmoudi et al., 2010; Mahmoudi et al., 2011	13
Croatian spotted goat	Croatia	8.1	8.1	0.74	Jelena et al., 2011	20
Ardi goat	Saudi Arabia	6.64		0.63	Aljumaah et al., 2012	
Brazilian goat breed (3)	Brazil	3.5 -7.2	3-11	NA	Araújo et al. 2006	11
Namibian goat breeds (4)	Namibia		4.67 –6.0		Els et al., 2004	18
Kalahari Red goat	South Africa	7.77	7.77	NA	Kotze et al., 2004	18
Tete goat	Mozambique	5.58			Garrine et al., 2010	
Pafuri goat	Mozambique	6.94			Garrine et al., 2010	
45 breeds	Mediterranean regions	5.2-9.1	5-43	NA	Caňón et al., 2006	30
Chinese goat populations (22)	China	5.24 -9.1	4-19	0.62-0.88	Li et al., 2002; Qi et al., 2009	17-20
Tswana goat	Botswana		1.83	0.58	Maletsanake et al., 2013	12
Indigenous goat populations (17)	Ethiopia	5.13 -6.73	2.06-23	NA	Tesfaye, 2004; Hassen et al., 2012a	15

Key:- N = number of populations; MS = number of microsatellite; NA = Not available

On the other hand, literatures state that PIC depicts the suitability of the markers and their primers used in the study for analyzing the genetic variability of a given population. Hence, microsatellite markers having greater than 0.5 PIC value are considered as highly informative and highly polymorphic (Botstein et al., 1980; Marshall et al., 1998). And hence, highly polymorphic markers were employed for most of the goat populations studied (Table 2). In contrast, microsatellites which had PIC value <0.5 were employed for the study, for instance, in Korean goats: PIC = 0.35, Kim et al., 2002; Egyptian and Italian goats of few loci: PIC=0.221, 0.482 and 0.389, Agha et al., 2008 and India goats: 28% of the loci had PIC<0.5, Kumar et al., 2009. These markers could have been screened out during the analysis. In fact, PIC is determined by heterozygosity and number of alleles (Aljumaah et al., 2012) and this makes microsatellite markers the choice for genetic characterization and diversity studies.

2.3.1.2 Genetic differentiation

The simplest parameters for assessing diversity among breeds are the genetic differentiation or fixation indices. Several estimators have been proposed (e.g. F_{ST} and G_{ST}), the most widely used being F_{ST} (Weir and Basten, 1990), which measure the degree of genetic differentiation of subpopulations.

In relative to dominant markers, co-dominant markers (e.g. microsatellite) are commonly used to assess genetic relationships between populations and individuals through the estimation of genetic distances (Sodhi et al., 2005; Tapio et al., 2005). The most commonly used measure of genetic distances is Nei's standard genetic distance (D_S) (Nei, 1972). However, for closely related populations where genetic drift is the main factor of genetic differentiation, as is often the

case in livestock populations particularly in the developing world, the modified Cavalli-Sforza distance (D_A) is recommended (Nei et al., 1983).

In relative to other reports, Cañón et al. (2006) obtained lower average values of F_{ST} for the four goat populations clusters (East Mediterranean: F_{ST}=0.033, Central Mediterranean: F_{ST}=0.040, West Mediterranean: F_{ST}=0.051 and Central-north European: F_{ST}=0.069) than F_{ST} values of 0.14 recorded for Asian goats (F_{ST}=0.17; Barker et al., 2001) and for Swiss goat populations (Saitbekova et al., 1999), F_{ST}=0.10 for a set of Chinese goat populations (Li et al., 2002). Similar low estimate of mean differentiation among populations (F_{ST} = 0.0717) was also reported for the goat populations in Brazil (Araújo et al., 2006) indicating presence of mixing among population and the most variability occurs within a population. According to Tesfaye (2004) and Hassen et al (2012a), F_{ST} values for each pair of Ethiopian goat populations varied from 0.001 to 0.050. In the same study, the average F_{ST} value over all microsatellite loci was 0.026. This might be because of that gene flow among most breeds has probably been restricted by geographical isolation rather than adherence to pedigree; i.e. a geographical restriction of genetic contacts of population may cause geographical clines or maintain clines that predate breed formation (Cañón et al., 2006).

Weir (1996) and Kalinowski (2002) had suggested the highest genetic distance (F_{ST}) to be higher than 0.25, moderate to be between 0.05 and 0.25 and the lowest estimate below 0.05. In general, the genetic distance between populations obtained by many of the scholars (Li et al., 2002; Tesfaye, 2004; Araŭjo et al., 2006; Hoda et al., 2011; Hassen et al., 2012a) is almost negligible (<0.05) and/or moderate (0.05<F_{ST}<0.25) values. Frankham et al. (2002) also explained that a

fixation index (F_{ST}) of about 0.15 is considered to be an indication of significant differentiation among populations.

Level of inbreeding (F_{IS}) is among the fixation indices that measures reduction of heterozygosity of an individual as a result of non-random mating within its subpopulation (Widmer and Lexer, 2001). It is an average increase of homozygous loci by F_{IS} by decreasing the heterozygous loci by the same proportion (Dorji et al., 2012; Maletsanake et al., 2013). F_{IS} is estimated for populations which show significant deviation from the HWE and is significant for significant HWE estimation (Widmer and Lexer, 2001; Ojango et al., 2011).

In connection to this, moderate and high level of inbreeding coefficients were reported by various scholars for different goat populations: $0.16 \leq F_{IS} \leq 0.26$ for Indian goats (Kumar et al., 2005; Gour et al., 2006; Aggarwal et al., 2007; Dixit et al., 2009), F_{IS} =0.18 for Ardi goat of Saudi Arabia (Aljumaah et al., 2012), F_{IS} =0.12 for Tswana goat in Botswana (Maletsanake et al., 2013). For the later goat population in particular, the F_{IS} estimate ranged from -0.2340 (INRA006) indicating low levels of inbreeding at that marker locus to 0.8772 (MCM527) depicting high levels of inbreeding. The large proportion of loci that deviate from HWE strengthened the highest F_{IS} estimations in those goat populations studied. This could be because of those loci being under within major histocompatibility complex (Schwaiger et al., 1993), under strong natural selection pressure (Hedrick and Kim, 2000) presence of null or non-amplified alleles, sampling structure effect, selection against heterozygotes or inbreeding (Araújo et al., 2006). However, tolerable mean value of F_{IS} (F_{IS} = 0.03) was obtained for 12 Chinese indigenous goat populations which employed 17 microsatellites (Li et al., 2002).

2.3.2 Genetic diversity of goat populations using single nucleotide polymorphism

Single nucleotide polymorphism (SNP) is a co-dominant marker which is widely used for the study of population genetic diversity and structure. Especially, the development of Chromatin Immunoprecipitation (CHIP) array for each livestock species has made the study of various molecular genetic studies to be addressed at genome wide level very easily. From ruminant livestock species, the study of genetic diversity and structure of *C. hircus* is the least addressed ruminant species using SNP markers. Before the development of the *capra* CHIP panel in 2012 (Tosser-Klopp et al., 2014), very few scholars (Cappuccio et al., 2006; Pariset et al., 2009a; Hykaj et al., 2013) used the SNP markers by sequencing a specific regions of the DNA for a particular study of *C. hircus*. As a result, those findings were more biased to marker evaluation than the status of the populations. Cappuccio et al. (2006) used 27 SNPs; Pariset et al. (2009a) and Hykaj et al. (2013) employed 26 SNPs each.

However, recently, following the release of the CHIP panel, few scientific reports have been documented on domestic goats at genome wide level since 2013. These are Carillier et al. (2013) on French dairy goat, Mucha et al. (2014) on UK dairy goat, Solomon (2014) on Nigerian and Ethiopian goats, Zidi et al. (2014) on Florida dairy goats, Benjelloun et al (2015) on Morocco goats and Nicoloso et al (2015) on Italian goats. Except Solomon (2014) and Nicoloso et al. (2015) who reported on genetic diversity of domestic goats, all were focused on the study on genomic selection and adaptation. Therefore, the following sections summarize those efforts carried out based on few numbers of SNP markers and the CHIP panel.

2.3.2.1 Estimation of heterozygosities

The mean number of alleles (MNA) per population, and observed (H_O) and expected (H_E) heterozygosities are the most common parameters for assessing within-breed diversity. The average H_E and H_O values of eight European goats, based on 27 SNPs, were 0.358 and 0.290, respectively (Cappuccio et al., 2006). Similarly, Pariset et al. (2009a) reported the average H_E and H_O values 0.300 and 0.272 for 16 Albanian, Greek and Italian goat populations using 26 SNPs. Hykaj et al. (2013) used 26 SNPs and obtained 0.316 and 0.282 mean values of H_E and H_O, respectively for six Albanian goat populations.

However, with high density SNP CHIP array (50K SNP CHIP), Kijas et al (2013) reported H_E = 0.3316 and H_O =0.3482 for five Australian and New Zealand goat population, Nicoloso et al. (2015) revealed H_E = 0.3316 and H_O = 0.3482 for 16 Italian goats. These estimates are by far lower than the findings (H_E = 0.3834±0.1274 and H_O = 0.376±0.1474) of Solomon (2014) on Nigerian and Ethiopian goats reported with the same SNP CHIP panel. Similarly, Cappuccio et al. (2006) obtained PIC values ranging from 0.046 (FABP4) to 0.459 (MEG3). However, relatively high range values (0.0058 to 0.4327; μ=0.2545) of PIC were reported for 26 SNP markers, not for the populations, in Albanian indigenous goats (Hykaj et al., 2013).

On the other hand, according to Pariset et al. (2009a), the range of major allele's frequencies was from 0.508 (for the locus IL4) to 0.992 (for the locus FABP4). On the same report, except for IL2-1 and FABP4 which showed frequencies of the rare alleles of 0.017 and 0.008, respectively, all other SNPs (24 SNPs) have a frequency of the rare allele greater than 5%. This finding was similar with Cappuccio et al. (2006) who observed the same loci on a different European breeds

and with Hykaj et al. (2013) who obtained frequencies of major alleles with a range of 0.524 (mel-g_1) to 0.997 (FABP4) for Albanian local goats.

Except one of the two SNPs in the interleukin-2 gene, all 27 SNPs have a frequency of the rare allele higher than 5% over all breeds and are suitable for genetic analysis (Cappuccio et al., 2006). In the same SNPs panel, except four loci, all loci had frequencies of rare alleles higher than 5% (Hykaj et al., 2013).

2.3.2.2 Genetic differentiation

Pariset et al. (2009a) revealed an overall 0.063 estimate of genetic differentiation (F_{ST}) for the six Albanian, two Greek and eighteen Italian goat breeds. This implies that 0.063 of allelic variation was accounted across breeds and 0.937 within breeds. Similar estimation, F_{ST} = 0.063, was reported for Nigerian and Ethiopia goats (Solomon, 2014) with the 50k SNP CHIP panel. However, very little differentiation (F_{ST} = 0.034) was reported for Albanian goats (Hykaj et al., 2013) and this could be because of high level of admixture. In contrary, higher estimation of population differentiation (F_{ST} = 0.081) was reported for eight goat populations in Europe (Cappuccio et al., 2006). Based on high density SNP CHIPs array, minimal to Moderate level of pairwise distances (F_{ST} = 0.013 to 0.164) were reported for Italian goats (Nicoloso et al., 2015).

2.3.3 Genetic diversity and relationship of indigenous goat populations in Ethiopia

Despite the large goat population found in the country, very few jobs have been done on the study of genetic diversity of Ethiopian goats. Tesfaye (2004) and Hassen et al. (2012a) compiled the genetic diversity of 11 (Abergalle, Arsi-Bale, Afar, Central Highland, Gumuz, Hararghe

Highland, Kaffa, Long-eared Somali, North west Highland, Small-eared Somali and Woyto-Guji) and six (Abergelle, Gumuz, Agew, Bati, Begia-Medir and Central Abergelle) populations using microsatellite markers, respectively.

Hence, the findings on allelic distribution (MNA=4.88: Tesfaye, 2004; MNA=6.73: Hassen et al., 2012a) and heterozygosity (H_E=0.58, H_O=0.55: Tesfaye, 2004; H_E=0.64, H_O=0.62: Hassen et al., 2012a) were very low compared MNA estimations (8.1 - 9.7) for Indian domestic goats (Pramod et al., 2008). Despite the narrow range observed in Tesfaye's report (Tesfaye, 2004), the maximum number of private alleles in the goat populations was only two and the frequencies in all the cases were less than 0.05 implying the divergence between populations is very narrow.

Smallest value of MNA (MNA = 4.8) was reported for Central Abergelle goat population (Hassen et al., 2012a) which might be due to the smallest sample size (N=8) employed. The sample size employed for this goat population is quite far from FAO (2011) recommendation for SSR marker based analysis. According to Kalinowski (2005), large samples are expected to have more alleles than small samples; however, the degree of influence of small sample size is weak as compared with the size of number of markers to be used (Landguth et al., 2012).

On the other hand, only three out of 15 microsatellite loci (Tesfaye, 2004) and only two out of 15 microsatellite loci (Hassen et al., 2012a) used in the analysis showed higher observed heterozygosity values than expected heterozygosity values probably implying the existence of sampling bias (Dorji et al., 2012). In addition, three and four microsatellite loci had shown H_E and H_O estimates of less than 0.5, respectively (Tesfaye, 2004). Similarly, Hassen et al. (2012a)

used two and three loci which showed H_E and H_O less than 0.5, respectively. However, such loci having H_E and H_O values less than 0.5 are not appropriate for heterozygosity evaluation and could be dropped out (Davila et al., 2009; Dorji et al., 2012) or could require to be prudent in selecting microsatellite loci.

Moreover, the number of alleles found per locus is the other perspective to evaluate efficiency of the loci. Barker (1994) and Nassiry et al. (2009) recommended that the number of alleles to be found per locus for remarkable genetic diversity of a population should be equal or greater than four. However, in studies conducted by Tesfaye (2004), few of the microsatellite loci showed MNA per locus lower than four.

On the other hand, F_{ST} values for each pair of populations varied from 0.001 to 0.040 for eleven Ethiopia goat populations; whereas, the genetic distance among Ethiopian goat populations follows the pattern of Isolation-by-distance (Tesfaye, 2004). According to Wright (1943), isolation-by-distance disrupts the gene flow. The rate of gene flow between populations seems inversely related to the geographic distance separating populations and an isolation-by-distance pattern might be evolved. However, Hassen et al (2012a) reported contrasting observation related to isolation-by-distance. Overall, the genetic distance among Ethiopian goat populations in both studies is negligible.

2.4 Fine mapping of kisspeptin (KISS1) gene

Growth and reproduction are two crucial economic traits in animal husbandry, and they are coordinated during normal puberty and the adulthood (Zhang et al., 2011). Growth is a complex

process that involves the regulated coordination of a wide diversity of neuroendocrine pathways (Zhang et al, 2008a). Nutritionally-induced changes in follicular development are mediated by metabolic hormones (Prunier and Quesnel, 2000; Armstrong et al., 2003). Particularly growth hormone (GH) of mammals plays an important role in the control of reproduction involving cell division, ovarian folliculogenesis, oogenesis and secretory activity (Hull and Harvey, 2002; Ola et al., 2008). By acting through specific receptors within the ovary, GH is expedient in controlling proliferation and apoptosis, oocyte maturation, and the expression and synthesis of receptors to hormones and related substances (Hull and Harvey, 2000; Sirotkin et al., 2003). Silva et al. (2009) also stated that the effect of GH on ovary function is mainly through inducing the development of small antral follicles in the gonadotrophin-dependent stages and stimulating oocyte maturation.

Genetic studies have indicated that the litter size and ovulation rate can be genetically determined by the action of genes (Deldar-Tajangookeh et al., 2009). However, it is stated that not only for fitness traits (e.g. litter size), the identification of candidate genes that are responsible for variation in continuous traits or quantitative traits (e.g. growth traits) has been a challenge in modern genetics (An et al., 2013). As effect, to date, little has been divulged on the major genes associated with litter size in goats. The few studies are the inhibin alpha-subunit gene (INHA) (Hua et al., 2008; Wu et al., 2009), the gonadotrophin releasing hormone receptor gene (GnRH) (An et al., 2009), the bone morphogenetic protein receptor-IB gene (BMPRIB) in the prolific Indian Black Bengal goat (Polley et al., 2009), and the bone morphogenetic protein 15 gene (BMP15) in Jining Grey goats (Chu et al., 2007). However, some of the studies (Bai et al., 2005; Gupta et al., 2007) described only the polymorphism and overlooked the association

analyses. In general, till then there are very limited works conducted in relation to association studies particularly in livestock in general and in goat in particular.

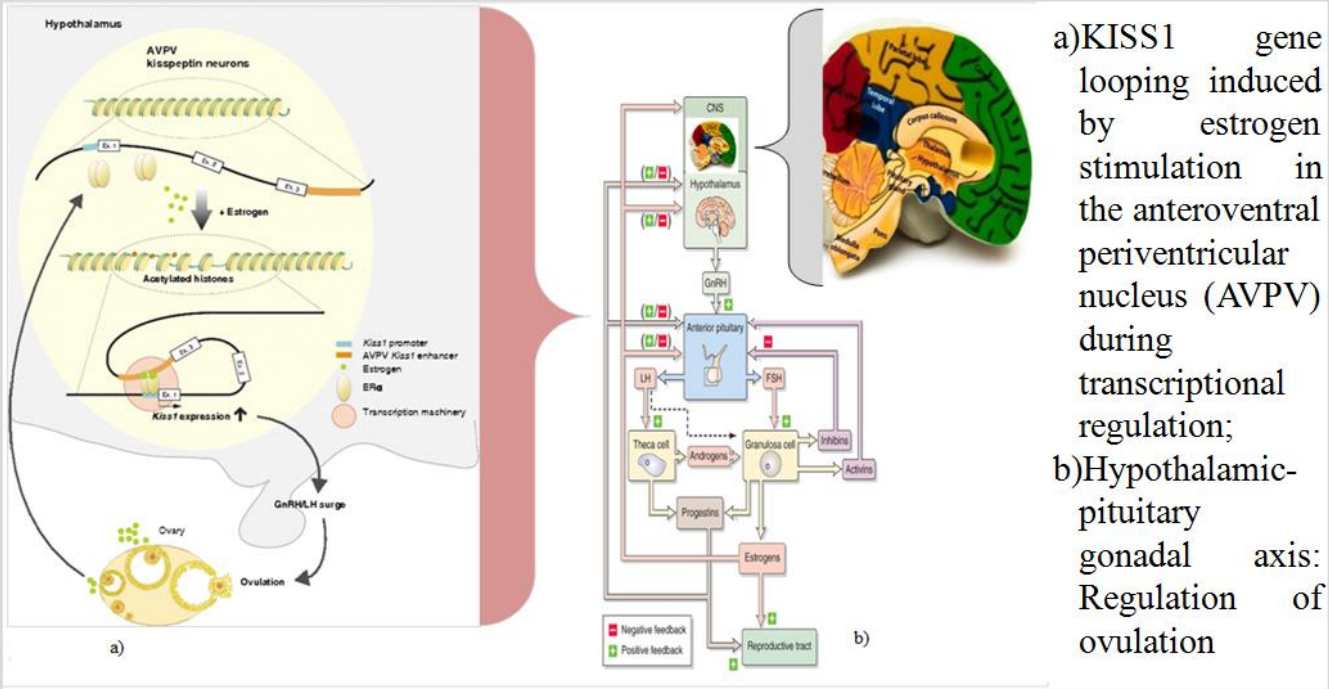
2.4.1 KISS1 gene

2.4.1.1 Roles and regulation of KISS1 gene

Kisspeptin (formerly known as metastin) is a protein encoded by the KISS1 gene (Gottsch et al., 2009b) which is located on the long arm of chromosome 1 (1q32) (Messager et al., 2005). It is expressed in the hypothalamus region of the forebrain (Gottsch et al., 2009b) and is a G-protein coupled receptor ligand (GPR54) (Messager et al., 2005). The name "KISS1" gene and its product, kisspeptin, might have been from where it was discovered in Hershey, Pennsylvania, the home of the chocolate "Hershey's Kisses" (Gottsch et al., 2009a). KISS1 was originally identified as a human metastasis suppressor gene that has the ability to suppress melanoma and breast cancer metastasis (Lee et al., 1996). It recently became clear that kisspeptin-GPR54 signaling has an important role in initiating secretion of gonadotropin-releasing hormone (GnRH) at puberty (Dungan et al., 2006; Smith et al., 2006). Kirilov et al. (2013) also stated that signaling between kisspeptin and its receptor, G-protein-coupled receptor 54 (GPR54) is now recognized as being essential for normal fertility. In this line, beside to the *pulse* mode, that is effected by the secretion of the reproductive neuropeptide gonadotropin releasing hormone (GnRH) and is essential for reproductive events in both sexes (like spermatogenesis, follicular development, and sex steroid synthesis), the *surge* mode of GnRH induces ovulation in females (Figure 3 a and b) (Okamura et al., 2013). Various scholars also stated that the central or peripheral administration of kisspeptin stimulates GnRH-dependent luteinizing hormone (LH) and follicle-stimulating hormone (FSH) secretion in various mammalian species from rodents to humans and kisspeptin administer reproductive functions of animals (Gottsch et al., 2004;

Navarro et al., 2005; Shahab et al., 2005). The gene products kisspeptins, belong to a larger family of peptides known as RFamides which all share a common arginine-phenylalanine-NH_2 motif at their C-terminus (Kirilov et al., 2013).

According to Gottsch et al. (2009b), neurons that express KISS1 play a crucial role in the regulation of pituitary luteinizing hormone secretion and reproduction and these neurons are the direct targets for the action of estradiol-17β (E2), which acts via the estrogen receptor α isoform (ERα) to regulate KISS1 expression. Kisspeptin/neurokinin B/dynorphin (KNDy) neurons located in the hypothalamic arcuate nucleus might play a central role in the generation of GnRH pulses in goats, and perhaps other mammalian species (Okamura et al., 2013). Gottsch et al. (2009b) also stated that in the arcuate nucleus (Arc) where the dynorphin gene (Dyn) is expressed in KISS1 bearing neurons, E2 inhibits the expression of KISS1 mRNA; however, E2 induces the expression of KISS1 in the anteroventral periventricular nucleus (AVPV) (Figure 3a). ERα signals through multiple pathways, which can be categorized either as classical, involving the estrogen response element (ERE), or nonclassical, involving ERE-independent mechanisms.



Source: Adapted from a) Tomikawa et al. (2012) and b) Boron and Boulpaep (2011)

Figure 3. Proposed model for KISS1 gene expression and the role of its product in the regulation of female reproductive physiology

2.4.1.2 Polymorphism on KISS1 gene and its association to multiple births

Various studies confirmed, with varied mutation sites, that KISS1 gene is a candidate gene having an important role for the reproduction of human and animals. For instance, two (c.374C>T and c.422C>G) mutations were identified in human KISS1 gene, and the c.374C>T variant was associated with higher kisspeptin resistance to degradation in comparison with the wild type, suggesting a role for this mutation in the precocious puberty phenotype (Silveira et al., 2010**).** Huijbregts et al. (2012) also detected three SNPs (c.638insT, c.641C>G and c.645G>CA) in the 3′UTR of human KISS1 gene, and the c.645G>CA mutation was associated with central precocious puberty.

Similarly, in Jining Grey goat KISS1 gene, there are two mutations (G3433A and C3688A) in exon 3, three mutations (G296C, G454T and T505A) in intron 1 and an 18 bp deletion/insertion (1960–1977) in intron 2 and no mutations in exon 2 (Cao et al., 2010). Feng et al. (2009) had also detected polymorphism in exon 2 of goat KISS1 gene. Similarly, ten polymorphisms were detected in KISS1 gene of three goat breeds (Xinong Saanen, Guanzhong and Boer goat breeds) (g.1147T>C, g.1417G>A, g.1428_1429delG, g.2124C>T, g.2270C>T, g.2489T>C, g.2510G>A, g.2540C>T, g.3864_3 865delCA and g.3885_3886 insACCCC), and the g.384G>A mutation was detected in Xinong Saanen and Guanzhong goat breeds (An et al., 2013). Cao et al. (2010) indicated an association between allele C of the 296 locus and allele deletion of at 1960–1977 locus in KISS1 gene and high litter size in Jining Grey goats. On the other hand, Hou et al. (2011) identified T2643C and 8bp base deletions (2677AGTTCCCC) in the intron 2 of goat KISS1 gene and the T2643C had significant effects on litter size ($P < 0.05$).

In goat, it is reported that the mean litter size tended to increase in later parities (An et al., 2013); hence, individuals with SC1 (AATTAATT) had higher litter size than those with SC4 (AATTGACT) and SC10 (GGTTAATT) in the second parity of Xinong Saanen breed. In addition, individuals with SC1 (AATTAATT) had higher litter size than those with SC10 (GGTTAATT) in average parity of Xinong Saanen breed. The litter size at second kidding is often a valuable index to determine whether a goat is prolific (Yuqin et al., 2011). Therefore, SC1 (AATTAATT) can be used in marker-assisted selection to select the individuals with higher litter size (An et al., 2013).

In summary, with respect to the global coverage of the study of origin of goat, Africa and Middle East partly have been poorly addressed compared to other parts of the world. This calls for more detailed investigation in these regions with the view to elaborate genetic lineages and patterns of dispersions. Sampling bias could be among the possible reasons that led other haplogroups (B, C, D, F) to have been underestimated. On the other hand, studies conducted by autozomal markers indicated that there is high within population genetic variations and very narrow population differentiation among the goat populations studied. However, technical gaps observed in most of the studies conducted by SSR markers may not help to suggest firm conclusions. Besides, very little works have been conducted globally on the study of genetic diversity, structure and admixture analyses of domestic goat and polymorphism analysis of targeted genes using high density SNP CHIP array and sequencing. As a consequence, genomic and marker assisted selection have been insignificantly employed in the domestic goat breeding programs.

CHAPTER III

3 Analysis of mitochondrial DNA control region (*D*-loop) of Ethiopian indigenous goat populations

3.1 Introduction

Ethiopia is a home to quite large number of goat populations which are kept in various production systems. However, there is lack of clarity on the classification and characterization of the indigenous goat populations to support the current and future breeding programs. This calls for the need to objectively characterize the gene pool in a systematic manner, and at the same time to elucidate the possible origin and routes of introduction of the goat into the country.

Molecular studies on domestic goats revealed six mtDNA *D*-loop lineages (A, B (B1, B2), C, D, F, G) with a weak phylogeographic structure (Luikart et al 2001; Chen et al 2005; Naderi et al 2007). This finding was interpreted to be the result of multiple independent domestications in the Fertile Crescent (Luikart et al 2001). An alternative argument suggested that such diversity in mtDNA lineages was also compatible with a single geographic center of domestication followed by a phase of human management of wild semi-domesticated variants comprising several mtDNA lineages before geographic dispersion and subsequent localized extinction of some lineages. All the six lineages are present in the wild ancestor, the Bezoar (Naderi et al., 2007 and 2008) suggesting the domestication of *C. hircus* occurred in a wide geographic area across South-west Asia, the home tract of the Bezoar. These results agree with archeological evidence which revealed that goats were domesticated in an area between the Zagros mountain and the Fertile Crescent around 10,500 Years ago (Zeder and Hesse, 2000; Zeder, 2008). The analysis of complete mtDNA genomes (Nomura et al., 2013; Doro et al 2014) revealed congruence in

clustering patterns of mitogenomes with those generated with the *D*-loop region and that the process of goat domestication was much more complex than envisaged.

Globally, lineage A has the widest geographic distribution (Pereira et al., 2005) and its most likely origin has been proposed to be Eastern Anatolia where it is common in wild populations (Zeder and Hesse, 2000; Naderi et al., 2008). Haplogroup B is confined to eastern and southern Asia, including Mongolia, Laos, Malaysia, Pakistan and India; and haplogroup C is present in low frequencies in Mongolia, Switzerland, Slovenia, Pakistan and India. Haplogroup D is rare and is observed only in Pakistan and Indian local goats. Haplogroup F is limited to Sicily; whereas, haplogroup G is reported only in Turkey, Saudi Arabia, Iran, Kenya and Egypt (Naderi et al., 2007; Kibegwa et al., 2015).

Ethiopia was considered to be main gate of livestock entry to Africa. As a result large livestock population size, with diversified genetic background, is found in the region. Social anthropology studies revealed that social and natural pressures have contributed to the wider but mixed coverage of livestock species including goats. For instance, the recurrent droughts occurred since the 15th century and even earlier had contributed for the physical movement of the people together with their animals (Girma, 1988), and this could have also influenced the genetic admixture of the livestock populations. However, there is dearth of information about the origin, genetic diversity and demographic expansion of the indigenous goats in particular, especially using maternal DNA. Therefore, this study was aimed to investigate the genetic diversity, origin and population expansion of the native goat populations in Ethiopia.

3.2 Materials and methods

3.2.1 Sampling and DNA extraction

A total of 309 samples representing 13 Ethiopian indigenous goat populations were sampled from farmers' flocks and used for the study. The goat populations included in the study with sampling area in brace are Abergelle (Zequla and Abiyadi), Nubian (Micarda), Gondar (Lay Armacheho), Agew (Addis Kidamin area), Gumuz (Pawe), Ambo (Meta-Robi), Kaffa (Tepi and Sheka), Woyto-Guji (Konso), Small eared Somali (Kebri-Beyah), Long eared Somali (Filtu), Afar (Melka-Werer), Hararghe Highland (Hirna) and Arsi-Bale (Arsi-Bekoji). The classification reported by FARM-Africa (1996) was followed to sample the goat populations. Based on the information obtained from farmers, all efforts were made to ensure that all the sampled individuals were minimally related. The blood samples were drawn out from the jugular vein with a volume of 9 ml under aseptic conditions using ethylene diamine tetraacetic acid (EDTA) anticoagulant. The collected samples were brought to the laboratory with ice box and were stored at -20^0_C until it was subjected to DNA extraction. Total genomic DNA was extracted from the samples using the salting-out extraction procedure (Shinde et al., 2008). The DNA quality and concentration were tested by nanodrop, and 1.0 - 1.5% agarose gel electrophoresis was used to evaluate the degradation. *D-loop*

3.2.2 PCR amplification and sequencing

Nested primers were used to amplify and sequence 1063 bp of the control region of mtDNA (Appendix Table B). Touch-down PCR amplification was performed using AccuPower® PCR Premix (Bioneer-Daejeon, Korea) to which 0.2 µM of each primer, 1.5% formamide (Hi-Di^{TM}; Applied Biosystems-USA), 0.005 mg of Bovine Serum Albumin (BSA; Thermoscientific), 50 ng of template DNA and double distilled H_2O were added to make a final reaction volume of 20 µl.

The touch-down PCR cycling profile involved an initial denaturation step at 95^{o}C for 3 min, followed by the first stage of amplification of five cycles involving a denaturation step at 90^{o}C for 10 sec, annealing at 58^{o}C for 40 sec, and extension at 72^{o}C for 30 sec. The second stage involved 30 cycles involving a denaturation step at 90^{o}C for 10 sec, annealing at 53^{o}C for 40 sec and extension at 72^{o}C for 40 sec. A final extension step at 72^{o}C for seven minutes completed the PCR reactions. The PCR products were purified using the QIAquick® 96 PCR Purification Kit (Qiagen, Hilden-Germany) following the manufacturer's instructions. The purified PCR products were sequenced using the Big Dye Terminator v3.1 cycle sequencing chemistry (Applied Biosystems) and the 3130XL automatic capillary sequencer (Applied Biosystems, USA).

3.2.3 Data analysis

All the chromatograms were generated and visualized with the CLC workbench 7.0.4 (CLC Bio-Qiagen). Multiple sequence alignments were done in CLC working bench employing the ClustalW algorithm (Thompson et al., 1994) and edited manually in MEGA 6 (Tamura et al., 2013). Variable sites were scored/called against the *C.hircus* reference sequence Genbank accession number GU223571 (*direct submission*) that was retrieved from the Genbank database. In total, 309 sequences were generated and collapsed into haplotypes using the DnaSP package v5 10.01 (Librados and Rozas, 2009). The level of genetic diversity represented as the number of haplotypes, haplotype diversity, nucleotide diversity and mean number of nucleotide differences between haplotypes were determined for each population using Arlequin 3.0 (Excoffier et al., 2005).

To visualize the genetic relationship between individuals and populations, phylogenetic tree was constructed for all the haplotypes using the Neighbour-Joining (NJ) algorithm as implemented in

MEGA6. To evaluate the level of confidence that can be associated with each bifurcation, the NJ tree was reconstructed following 1000 bootstrap replications. To complement the NJ tree while obtaining further insights, and in greater detail, into the genetic relationships between the haplotypes, the median-joining (MJ) network of haplotypes was constructed using the Network v4.6 software. For this analysis, all the mutations and character states were equally weighted. A sequence from the GenBank (Accession No.: GU223571) was used to align the whole *D-loop* region. To visualize the variations in Ethiopian goats in the perspective of the global *Caprine* variation, analysis of phylogenetic tree and phylogentic network were performed using hypervariable region I (HV1: 481bp) incorporating 22 reference haplotypes for neighbour-joining (NJ) tree construction and 229 reference haplotypes for median joining network retrieved from the GenBank (Figure 4 and 5). The 229 reference sequences which represent six globally defined haplogroups were retrieved from the GenBank and used for haplogrouping (Appendix Table A). From the wild goats, *Capra aegagrus* (accession no: AJ317864-AJ317867) and *Capra cylindricornis* (accession number: AJ317868-AJ317870) were also incorporated into the study.

To evaluate the partitioning of genetic diversity and variation amongst populations and groups of populations, the analysis of molecular variation (AMOVA) was performed in Arlequin. For this analysis various hierarchical population clusters/groups were tested as follows: i) across the global dataset without any clusters, ii) between four groups of populations proposed by FARM-Africa (1996), and iii) between any population groups or clusters revealed by the NJ and MJ network analysis.

To shed light on the demographic dynamics of Ethiopian indigenous goat populations, mismatch distribution pattern (Excoffier and Schneider, 1999) analysis was also carried out in Arlequin based on the expected and observed mismatch heterozygosities. For this analysis, mismatch distribution patterns were generated for each of the 13 populations and for the overall dataset of Ethiopian goats, and for each of the population clusters revealed by the NJ and MJ analysis. The goodness of fit of the observed pattern to that expected under a demographic equilibrium was tested using the sum of squares deviation (SSD) of the goodness of fit statistic and the raggedness index (Harpending et al., 1993; Harpending, 1994). Mismatch distribution analysis was complemented by Tajima's *D* (Tajima, 1989b) and Fu's F_S (Fu, 1997) statistics which are coalescent based estimators of neutrality. The three tests of demographic dynamics were all performed following 1000 bootstrap replications in Arlequin. Time of introduction of goat to Ethiopian highlands reported based on archaeological data (Clutton-Brock, 2000) was used as initial mutation rates to examine the molecular dating, and evaluated using Network v4.6 software.

3.3 Results

3.3.1 Mitochondrial DNA sequence variation and genetic diversity

From the analysis of 309 mtDNA *D*-loop sequences, a total of 174 variable sites were observed across 1063 bp length fragment that generated 231 haplotypes. All the 13 populations were defined by a high level of genetic diversity. The number of haplotypes ranged between 12 (in Agew population) and 30 (in Afar population) (Table 3). The Kaffa population showed the lowest level of haplotype diversity (*H*d = 0.95000±0.037) while the highest level was observed in Small east Somali (*H*d = 1.00000±0.020), Hararghe Highland (*H*d = 1.00000±0.014) and

Woyto-Guji (*H*d = 1.00000±0.019) goat populations. Similarly, the nucleotide diversity ranged from 0.01430±0.0019 in the Afar to 0.01796±0.0010 in the Abergelle population. Haplotype diversity is known to be influenced by sample size. However, no positive correlation was observed between sample size and haplotype diversity in this study.

Table 3 Genetic diversity of mtDNA *D*-loop haplotypes

Population	**N**	***S***	***H***	***H*d ± SD**	**π ± SD**	**K**	**No. of haplotypes (%)**	
							Haplogroup A	**Haplogroup G**
Short Eared Somali	17	66	17	1.00000±0.020	0.01576±0.0024	16.69118	14(82.35)	3(17.65)
Long Eared Somali	19	65	17	0.98830±0.021	0.01594±0.0024	16.86550	14(82.35)	3(17.65)
Nubian	37	88	25	0.97297±0.013	0.01548±0.0018	16.37838	21(84.00)	4(16.00)
Hararghe Highland	22	65	22	1.00000±0.014	0.01649±0.0019	17.44589	16(72.73)	6(27.27)
Abergelle	35	78	29	0.98824±0.010	0.01796±0.0010	19.00504	22(75.86)	7(24.14)
Arsi-Bale	20	69	18	0.98947±0.019	0.01309±0.0023	13.84737	16(88.89)	2(11.11)
Ambo	16	68	13	0.96667±0.036	0.01701±0.0028	17.99167	10(76.92)	3(23.08)
Afar	33	80	30	0.99432±0.009	0.01430±0.0019	15.12500	24(80.00)	6(20.00)
Agew	15	56	12	0.97143±0.033	0.01435±0.0023	15.18095	11(91.67)	1(8.33)
Gumuz	25	66	19	0.96667±0.024	0.01628±0.0018	17.22000	15(78..95)	4(21.05)
Gondar	27	70	22	0.98860±0.015	0.01760±0.0012	18.62108	14(63.64)	8(36.36)
Kaffa	25	62	20	0.95000±0.037	0.00931±0.0017	09.85334	18(90.00)	2(10.00)
Woyto-Guji	18	63	18	1.00000±0.019	0.01640±0.0021	17.35294	14(77.78)	4(22.22)
Overall	**309**	-	-	**0.99668±0.001**	**0.01583**±0.0005	**17.42637**	**209(79.77)**	**53(20.23)**
Haplogroup A	248	154	185	0.99566±0.025	0.00978±0.0014	10.34396	-	-
Haplogroup G	61	69	46	0.98306±0.018	0.00568±0.0015	6.00576	-	-

Key: N = Sample size, *S* = No. of Polymorphic sites, *H* = Number of haplotypes, *H*d = Haplotype diversity, π = Nucleotide diversity, *K* = Average number of nucleotide differences; SD=Standard deviation

3.3.2 Population phylogenetic analysis

The 231 haplotypes observed in Ethiopian goats were used to construct a phylogenetic tree (based on the NJ algorithm) and a phylogenetic network to assess the relationships between the haplotypes. The NJ tree revealed well resolved three clusters (Figure 4): one cluster for the reference haplotypes except for haplogroup A and G, and two clusters for Ethiopian goats representing globally defined haplogroup A and G (Figure 5). Haplogroup A was the most predominant and included 185 haplotypes (80.1% of the total number of haplotypes observed) while haplogroup G was formed from 46 haplotypes (19.9%). None of these haplogroups was exclusively observed in a single population, geographic region or production system (Figure 6).

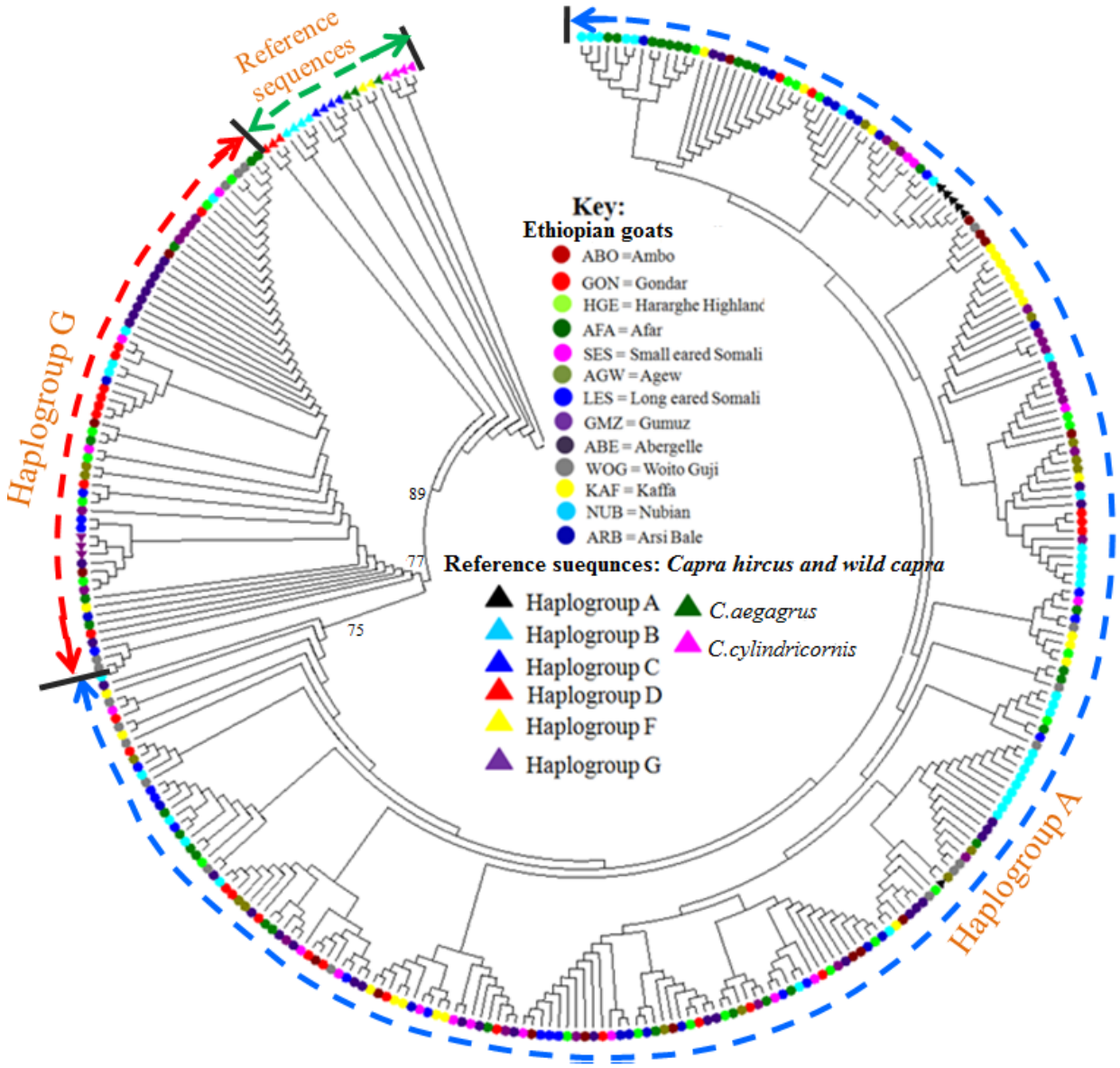


Figure 4. Neighbour Joining tree of 13 Ethiopian goat populations, six reference haplogroups and two wild *Capra*

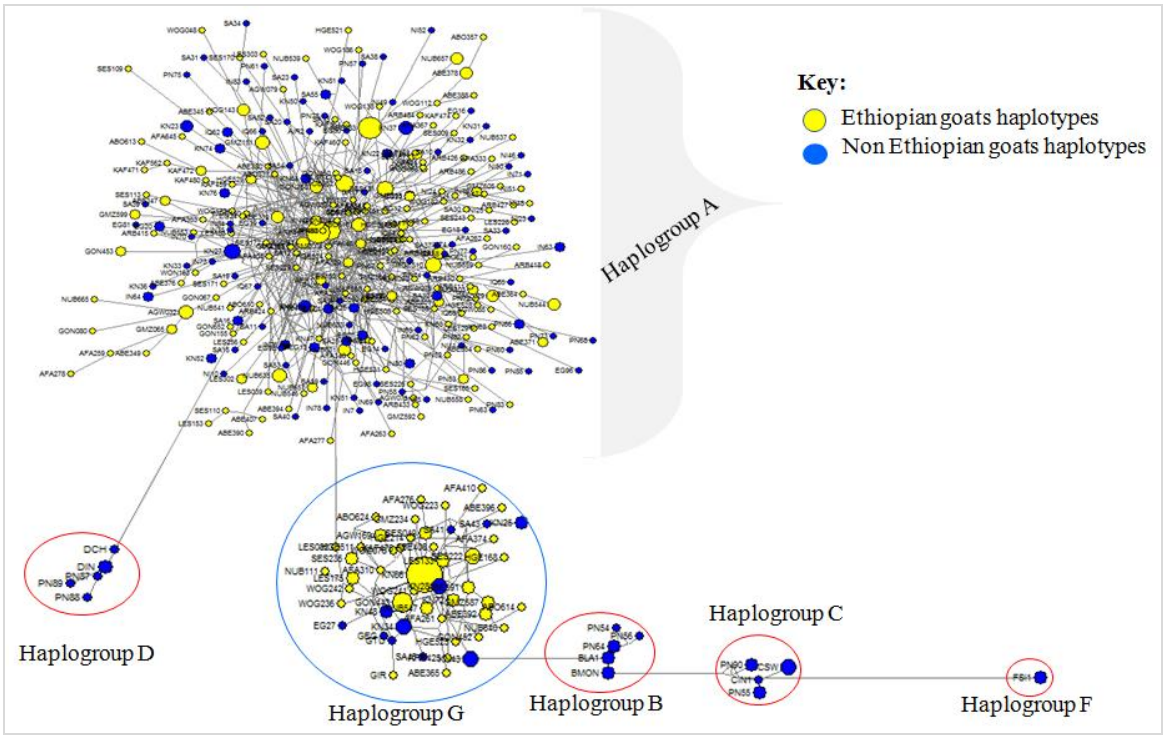


Figure 5. Network tree of the goat populations studied: base on 309 Ethiopian and 229 non-Ethiopian goats HV1 sequence data

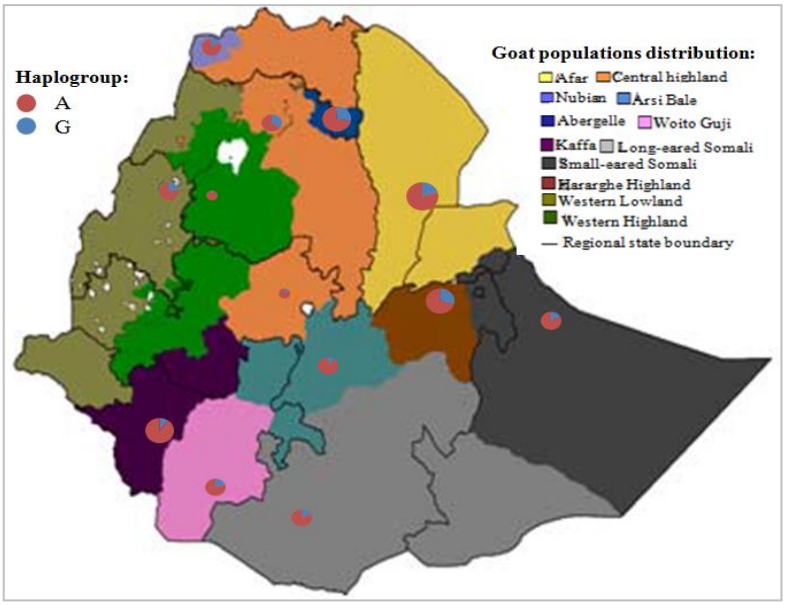


Figure 6 Haplogroup distribution depicted from the haplotypes detected with respect to the populations' distribution

3.3.3 Population differentiation

The global AMOVA incorporating all the 13 populations independent of any hierarchical clustering indicated that 97% of the total genetic variation present in Ethiopian indigenous goats was explained by genetic differences between individuals within populations (Table 4). Only 2.63% of the variation was attributable to genetic differences between populations. Performing the analysis taking into account two hierarchical clusters that were defined based on the clustering patterns observed on the NJ and MJ network revealed that 59.11% of the genetic variation was explained by differences between individuals within haplogroups while 40.89% of the variation was due to genetic differences between the two haplogroups (Table 4).

Table 4 Analysis of Molecular Variance (AMOVA) based on haplogroup and population groupings

Grouping	Source of variation	df	Sum of squares	Variance components	Percentage of variation
All populations	Apopulations	12	159.994	0.22045 Va	2.63
	Within populations	294	2400.906	8.16637 Vb	97.37
	Total	306	2560.906	8.38682	
Agro-ecology	AG	2	38.003	0.06155Va	0.73
	APWG	10	121.991	0.17777Vb	2.11
	Within populations	294	2400.912	8.16637Vc	97.16
	Total	306	2560.906		
Production system	AG	1	16.743	0.02784Va	0.33
	APWG	11	143.251	0.20811Vb	2.48
	Within populations	294	2400.912	8.16637Vc	97.19
	Total	306	2560.606	8.40232	
Goat family	AG	3	47.331	0.03074Va	0.37
	APWG	9	112.757	0.19674Vb	2.34
	Within populations	294	2401.805	8.17103Vc	97.29
	Total	306	2561.893	8.39851	
Based on haplogroups	Among haplogroups	1	409.069	4.88127Va	40.89
	Within haplogroups	305	2151.837	7.05520Vb	59.11
	Total	306	2560.906	11.93648	
	*F*st = 0.40894†; *F*st=0.27-0.28††				

Key: AG= Among groups; APWG= Among populations within groups; AIWP=Among indiv. within population; †=*F*st for among haplogroups; †† = *F*st among the rest of grouping

3.3.4 Population demographic dynamics

Sequence mismatch distribution patterns were assessed for each population, the global dataset incorporating all the 13 populations and for each of the two mtDNA haplogroups revealed by the NJ and MJ network as a proxy to elucidating the demographic history of the Ethiopian indigenous goat populations. Each population was characterized by a bimodal mismatch distribution pattern (Figure 7). For each population, the observed pattern did not differ significantly from the one expected for expanding populations with the exception of Abergelle population. The variations around the curves were also not significant with the exception of Agew population (Table 5). Similar results were also observed for the global dataset and for each of the two haplogroups, respectively. These results were supported by both the Tajima's *D* and Fu's F_S statistics, all of which were negative and significant except for the Abergelle and Gondar populations for Tajima's *D* which was negative but not significant for these two populations. These results taken together reveal a signal of expansion (demographic or spatial) in Ethiopian indigenous goats which may have occurred either prior to or after being introduced into the country. However, the sharp peaks suggest the expansion happened recently.

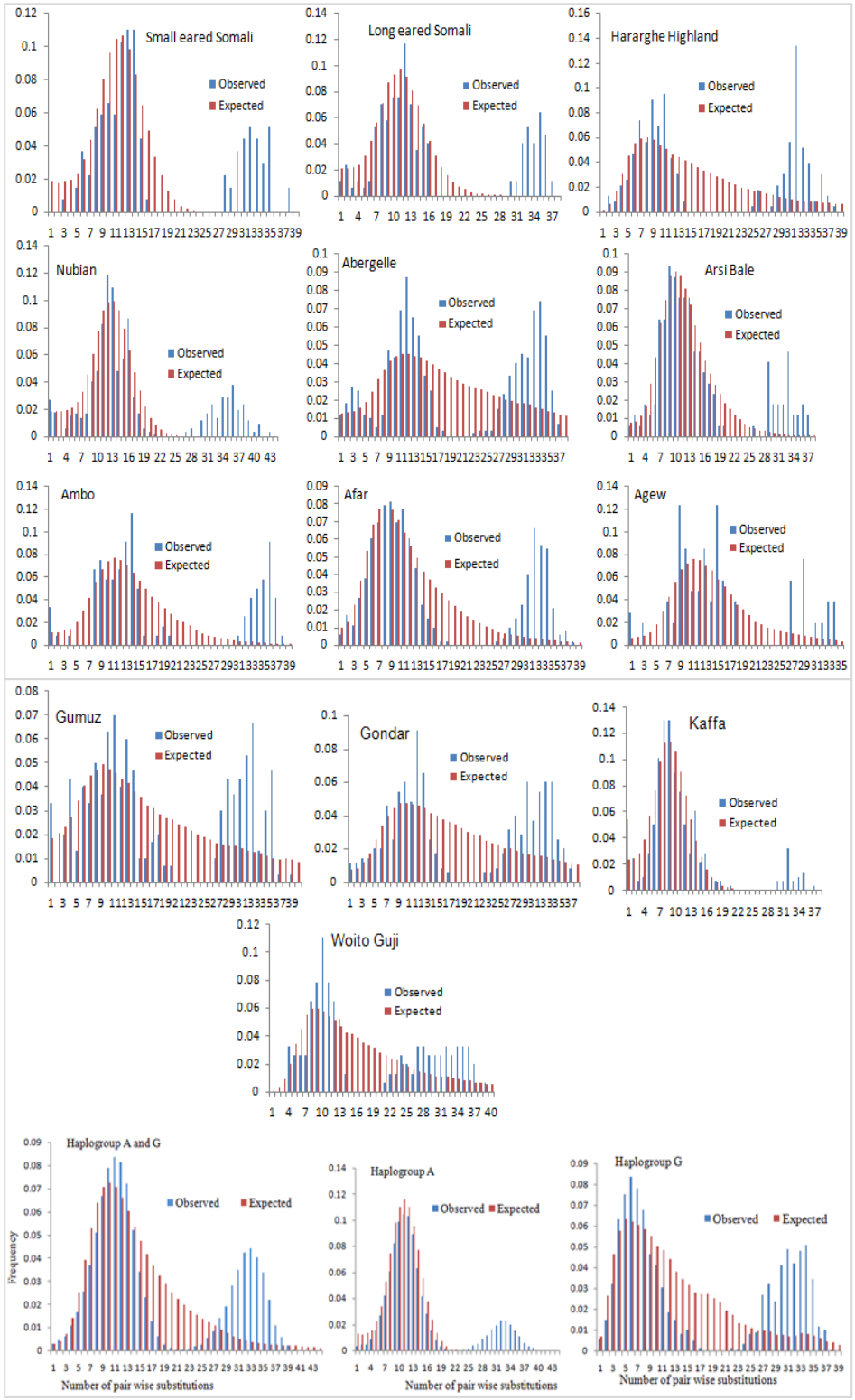


Figure 7 Mismatch distributions for mtDNA haplogroups of Ethiopian indigenous goats: for the overall dataset and the respective haplogroups

Table 5 Population demographic and neutrality test in Ethiopian goat populations

Population/haplogroup	N	S	SSD	Raggedness index "*r*"	Tajima's *D*	Fu's *F*s
Small eared Somali	17	66	0.03 ^{ns}	0.02 ^{ns}	-0.61 ^{ns}	-6.308**
Long eared Somali	19	67	0.02 ^{ns}	0.01 ^{ns}	-0.67 ^{ns}	-3.8^{ns}
Hararghe Highland	22	65	0.04 ^{ns}	0.03 ^{ns}	-0.09 ^{ns}	-9.77***
Nubian	37	88	0.01 ^{ns}	0.02 ^{ns}	0.82 ^{ns}	-3.95^{ns}
Abergelle	35	78	0.02*	0.01 ^{ns}	0.01 ^{ns}	-12.315***
Arsi-Bale	19	69	0.008 ^{ns}	0.009 ^{ns}	-1.165 ^{ns}	-6.218**
Ambo	16	68	0.031^{ns}	0.019 ^{ns}	-0.522 ^{ns}	-0.730^{ns}
Afar	33	80	0.017 ^{ns}	0.005 ^{ns}	-0.870 ^{ns}	-15.690***
Agew	15	56	0.03 ^{ns}	0.06 *	-0.51 ^{ns}	-1.86^{ns}
Gumuz	25	66	0.02 ^{ns}	0.02 ^{ns}	-0.06 ^{ns}	-3.06^{ns}
Gondar	27	70	0.02 ^{ns}	0.01 ^{ns}	0.10 ^{ns}	-6.39*
Kaffa	24	62	0.008 ^{ns}	0.014 ^{ns}	1.574 *	-5.547*
Woyto-Guji	18	63	0.02 ^{ns}	0.01 ^{ns}	-0.22 ^{ns}	-6.8**
Haplogroup A	258	164	0.01^{ns}	0.01^{ns}	-1.50*	-23.81**
Haplogroup G	49	89	0.02^{ns}	0.01^{ns}	-0.74^{ns}	-23.93**
All	309	174	0.013^{ns}	0.01^{ns}	-1.21^{ns}	-23.64**

Key: SSD=Sum of square deviation; ns=non-significant; *= significant at $p<0.05$; **=significant at $p<0.01$.

3.3.5 Molecular dating

Based on the archaeological data goats spread rapidly from the Near East into the central Sahara and Ethiopian highlands between 6,500 BP and 5,000 BP, together with sheep (Clutton-Brock, 2000). Using these times as initial mutation rates, both haplogroups detected in Ethiopian goats were believed to be separated 58,500 and 45,000YA. These figures need to be interpreted in caution since the closest nodes (haplotypes), for instance 10 mutations between haplogroup A and G, were used for examination. Similarly, haplogroup D was separated from haplogroup A in 110,500±26,800 YA and 19 mutation steps separate the two haplogroups; whereas, both haplogroup F (with 43 mutation steps) and haplogroup B (17 mutation steps) were separated from haplogroup G in 286,000±42,124 and 65,000±16,783 YA, respectively.

3.4 Discussion

There is historic evidence that Ethiopia is the main gate, sometimes mentioned reservoir, of livestock genetic resource into Africa from Middle East and South-East Asia (Hanotte et al., 2002). In this study, the control region of mtDNA was sequenced to investigate the genetic diversity and origin of Ethiopian goat populations. Two maternal haplogroups (haplogroup A and G) and high genetic diversity were detected from 231 haplotypes, 309 animals and 13 goat populations studied from which the multiple maternal haplogroups of domestic goat globally identified and agreed (Luikart et al., 2001; Sultana et al., 2003; Joshi et al., 2004; Chen et al., 2005; Naderi et al., 2007; Amills et al., 2008). The MJ network analysis replicated these results but revealed that the two haplogroups detected in the current study were separated by 10 mutation steps. This result provided further support for the existence of the two haplogroups in Ethiopia and their genetic distinctiveness. A total of 137 median vectors which exceed the number obtained in other studies was also observed (Luikart et al., 2001; Sultana et al., 2003; Joshi et al., 2004; Chen et al., 2005; Naderi et al., 2007 and 2008; Amills et al., 2008). The observed median vectors could represent haplotypes that were present in Ethiopia but were not sampled, they could represent haplotypes that were present in the original *Caprine* gene pool but were not introduced into the country or they were introduced into the country but became extinct either upon arrival or after some time. This could be supported by large number of nodes (n=192) and edges/branches (n=335) which are observed in the phylogenetic network construction of the same study goat populations using SNP chip data indicated in section 4.3.5. The nodes represent sub-populations and the edges/branches represent population sub-division (Huson and Bryant, 2006). Haplogroup A generally had a higher level of genetic diversity compared to haplogroup G, and there is 40.89% of variation among the haplogroups. This

percent of variation among the haplogroups is highly lower than 75.78% of among haplogroup variation recently observed in Anatolian goats (Akis et al., 2014).

In addition, weak phylogeographic structure but very high mtDNA diversity was observed; at least one and eleven haplotype(s) obtained in haplogroup A and haplogroup G, respectively from each goat population studied. The very high genetic diversity obtained could be partly explained by presence of high mutation rate in the control (*D*-loop) region (Naderi et al., 2007). In addition, same mtDNA haplotypes were detected in both haplogroups and in different goat populations (Appendix Table C). For instance, haplotype 8th (ET8) was detected in Arsi-Bale, Ambo, Hararghe Highland and Abergelle (lowland goat) goat populations. Except Abergelle goat, all are highland goats. Abergelle goat sampling area is >1200km (ground distance) far from where Arsi-Bale goat was sampled. Similarly, haplotype 31th (ET31) was observed in Ambo, Abergelle, Nubian, Afar, Gumuz and Agew goat populations. Agew and Ambo goats are highland goat populations; whereas, the remaining are lowland goats. Nubian goats were sampled >1400km (ground distance) far from Afar goat sampling area. Similar observation was reported for the goat populations in South-west Asian goats (Naderi et al., 2008). The high rate of migration (*N*m=19.62) obtained that led the distinct goat populations to share some proportion of genetic background they each other and homogenized their genetic architecture can strengthen this idea (Table 10). However, such mixing of haplotypes is very unusual in natural populations except in animals with high dispersal abilities (example, birds) (Naderi et al., 2008).

The haplotype diversity (0.9967) obtained in the current study is similar with estimates of haplotype diversity for Iberian (0.996) and European (0.994) goats; however, slightly higher than

estimates of Sicily goats (0.806-0.969), South and Central American (0.963) and Atlantic (0.965) goat populations (Amills et al., 2008). Similarly, 0.9884 of haplotype diversity estimate together with 221 unique haplotypes was reported for large scale mtDNA analysis of bezoar sequences (Naderi et al., 2008). Very large number of variable sites (N=336 sites) over 558 bp of alignment sequences of HVI region was also reported (Naderi et al., 2007).

The average haplotype diversity estimation obtained in the haplogroups (Table 3) is very high and comparable with previous report. According to Naderi et al. (2007), the haplotype diversity was 0.9992 for haplogroup A, 0.9000 for haplogroup B, 0.8402 for sub-haplogroup B1, 0.8151 for sub-haplogroup B2, 0.9714 for haplogroup C, 0.9487 for haplogroup D, 1.0000 for haplogroup F and 0.9544 for haplogroup G were reported. On the other hand, lowest estimates (0.95000±0.037; 0.00931±0.0017) of haplotype and nucleotide diversities, respectively, were observed in Kaffa goat population compared to the rest of Ethiopian goat populations studied. Moreover, Kaffa goat showed relatively higher differentiation than the rest of Ethiopian goats. The phylogenetic network, F_{ST} and populations admixture analyses of the SNP CHIPs indicated in the next chapter (Table 9) support this observation. This could be explained by level of gene flow towards Kaffa area is minimal and/or the Kaffa goat habitat could be unfavourable for other Ethiopian indigenous goat populations to adapt the local environment.

Similarly, the average number of haplotypes per population observed in this study was 20.31 with the range of 12 haplotypes for Agew to 30 haplotypes for Afar goat populations. This range varies from 3 to 25 in eighteen Chinese goat populations with haplotype diversity that ranges from 0.7121 to 0.9804 (Chen et al., 2005). In three Morocco goat populations, 64 polymorphic

sites and 40 haplotypes, in which all of them belong to haplogroup A, were detected (Benjelloun et al. 2015). Colli et al. (2015) also reported 229 polymorphic sites. In Sicily goat populations, 33 haplotypes were reported. Similarly, 54, 28, 71 and 53 haplotypes were identified for SCA, Atlantic, Iberian and European goat populations, respectively (Amills et al., 2008).

In the current study, more that 97% of the variation is explained by within populations' variation (Table 4). This estimation is highly higher than the within population variation (83%) for Indian goats (Joshi et al., 2004), 69% for SCA goats (Amills et al., 2008), 78.7% for European, African and Asian goats (Luikart et al., 2001) and 77% of within goat genetic variation at global level that included 54 countries (Naderi et al., 2007). In the latter report, 11% and 12% of variations explained among breeds within geographic regions and among geographic regions indicating weak geographic structure concurrently because of the widest coverage of global distribution of haplogroup A.

In global context, there is also low phylogeographic structure in domestic goat (Naderi et al., 2007). This weak phyologeographic structure was reported for 18 Chinese goat populations and in other ruminant livestock (Mannen et al., 2004; Freeman et al., 2004; Chen et al., 2005; Meadows et al., 2005 and 2007). However, the extent is higher in domestic goats than other ruminant livestock that may be because of high mobility of goats in relation to human migration and commercial trade due to their versatility in feeding habits and ability to live under extreme conditions (Clutton-Brock, 1999; Naderi et al., 2007).

Based on archaeological evidences, one of the domestic goat expansion routes to south of Egypt was through Nile delta, to Ethiopia (Clutton-Brock, 2000). In Egypt only haplogroup A and G were detected using mtDNA analysis (Naderi et al., 2007). These reports strengthen that Ethiopian goats are possibly descendant from Egypt. In addition, the second possible route of introduction to Ethiopia could be, still from Middle East, via Yemen of terrestrial route or from Pertia via the Sea route to east part of the country (Figure 8). This observation is strengthened by the SNP CHIP data analysis indicated in the next chapter. However, there is no any molecular clue about the other route of introduction directly from South-east Asia following Indian Ocean and Red Sea to Ethiopia. It would have been confidently proven if haplogroup B, which is dominantly found in Asia, could be observed in Ethiopia in this study. Haplogroup B was detected in South African countries in limited proportion (Naderi et al., 2007) following the Indian Ocean route. An individual which belonged to haplogroup B was also identified in Greece (Naderi et al., 2007).

Haplogroup D was only detected in Pakistan, India and China (Sultana et al., 2003; Joshi et al., 2004; Liu et al., 2007; Naderi et al., 2007). However, there is no any signature of this haplogroup detected in Ethiopia. Moreover, haplogroup G is not detected elsewhere in the world except in Iran, Turkey, Saudi Arabia, Kenya and Egypt (Naderi et al., 2007; Kibegwa et al., 2015). In addition, none of the six haplogroups, other than Haplogroup A and G, could be directly extended from either of the median vectors detected in the Network diagram (Figure 5). Literatures explain that there was political and strong trade ties between East Africa, Ethiopia in particular, and the Indian and Persian regions started two millennium ago (http://www.mea.gov.in/portal/foreignrelation/ethiopia-february-2012.pdf) which could possibly

contribute for introduction of domestic goat to East African region. Therefore, given small sample size employed for this study, this route of introduction was limited only for goats from haplogroup A and G, which is most unlikely, or the goats were arrived in Ethiopia only through the Nile Delta and via Yemen routes eastward.

On the other hand, previous reports identified the closest wild ancestor of *Capra hircus* is *Capra aegagrus* (Harris, 1962; Zeuner, 1963). However, in this study, Ethiopian indigenous goats appeared more close to *Capra cylindericornis* than *Capra aegagrus*. In addition, an animal from Afar goat population of Ethiopia appeared among sequences of *Capra cylindericornis* and reference sequences of haplogroup F, which were detected in Sicily (Sardina et al., 2006), (Appendix Figure A). However, the information obtained in this study is not sufficient to clarify these exceptional observations at this stage, and require further investigation.

Based on the information from mtDNA, it is noted that signatures of population expansion can be detected through frequency distribution of the number of pairwise differences between haplotypes and thus statistics based on the mismatch distribution (Rogers and Harpending, 1992). The bimodal distributions observed in the graph of population expansion, in the current study, indicate there were two major events of expansion in Ethiopian goat populations in sometimes ago (Figure 7) (Rogers and Harpending, 1992). Moreover, significant and the largest (referring the magnitude) negative *F*s values obtained in our study confirm presence of large and sudden populations' expansion (Chen et al., 2005). The authors reported equivalent *F*s estimate (*F*s = -23.57<0.01) for Chinese goats. Significant large negative F_S value indicates presence of rapid population expansion (Fu, 1997; Josh et al., 2004; Hasan et al., 2008). Comparatively,

minimal signature of expansion was detected in the second event of expansion in Kaffa goat; whereas, unique L-shaped mismatch distribution (in reverse direction) observed in Ambo and Abergelle goat populations indicates recent demographic expansion (Barluenga et al., 2006; Diegisser et al., 2006; Hasan et al., 2008). The socio-cultural and economic interactions among human societies could mainly contribute for the demographic expansion of the goat populations. However, Akis et al. (2014) observed multimodal mismuch curves in Anatolian Black and Angora populations, and unimodal curve in Kilis population.

The past historic, social as well as antropogenetic evidences in Ethiopia indicate that human population movement together with their animals from South to North parts of the country, and vice versa, were frequently observed for many centuries. Movement of Wolayta tribe from the south and central Ethiopia to Axum (extreme north) in 13th century, Tigray tribe to Central and Southern Ethiopia during the 10th century, massive movement of Oromo tribe up to North Ethiopia during the 15th and 16th centuries and Amhara tribe to South, East and Central Ethiopia at various times are some of examples of massive physical movement of tribes in Ethiopia (Yilma, 1967; Mpofu, 2002; Habitamu, 2014). In line with this, recent Y-chromosome, mtDNA (*D*-loop) and high density SNP array (1 Mb SNP CHIPs panel) data revealed that there is no significant variation among major tribes in Ethiopia (Christopher, 2011; Pagani et al., 2012). Similar trend was also observed in genetic diversity and structure of Ethiopian indigenous cattle (Sisay, 1996; Dadi et al., 2008; Edea et al., 2013).

Therefore, the existing genetic structure of livestock species in Ethiopia in general could have been highly influenced by human movement. In addition, the goat populations could possibly be

mixed before arrival and distributed throughout the country though male mediated gene flow was not tested in this investigation. It is noted that haplogroups could be mixed, even impossible to exclude each wild ancestor, before domestication and translocation of the goat worldwide (Naderi et al., 2007).

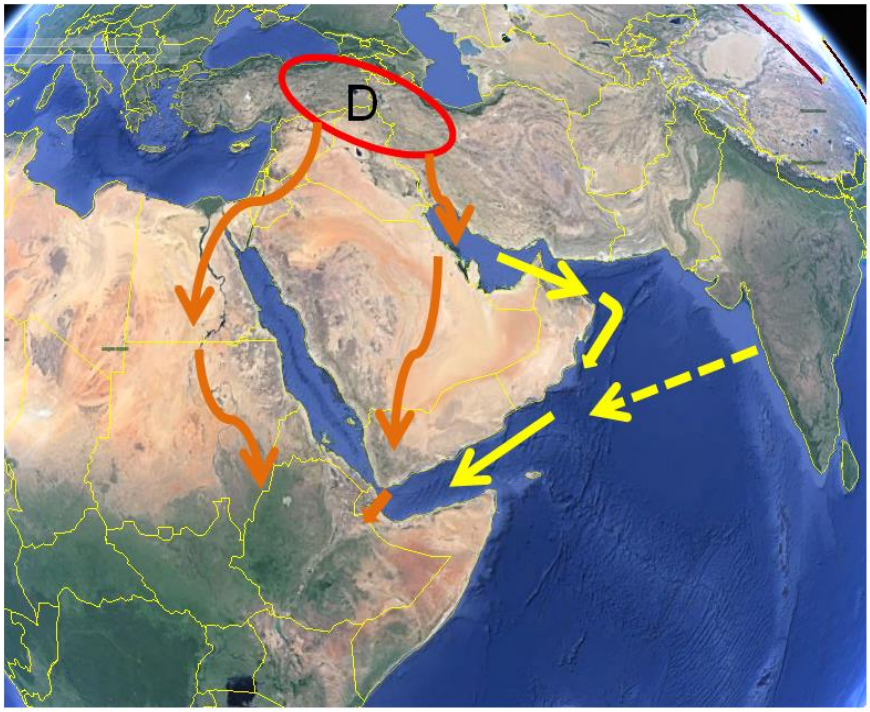


Figure 8 Possible routs of introduction of domestic goat to Ethiopia from center of domestication

CHAPTER IV

4 Genome-wide genetic diversity and structure of Ethiopian indigenous goat populations

4.1 Introduction

Ethiopia is regarded as a reservoir of livestock genetic diversity following being the gate way in to Africa (Devendra and McLeroy, 1982). The country is endowed with a large population of goats in which most of them are indigenous and are reared in diverse production and agro-ecological systems (FARM-Africa, 1996).

Despite the presence of a large gene pool, the classification of the indigenous goat populations in Ethiopia is ambiguous. Based on physical descriptions and microsatellite genotyped data, they were classified in to 13 (FARM-Africa, 1996) and eight clusters (Tesfaye, 2004), respectively. The 13 goat populations were further regrouped into four family groups based on phenotypic description and geographic location and into two groups based on the production system, and three groups based on the agro-ecological location of the goats (FARM-Africa, 1996) (Table 6). With the probable limited coverage of characterization studies, indiscriminate intercrossing between local ecotypes, and socio-economic exchange of livestock amongst different communities are all factors that may have influenced the genetic profile of indigenous goats (Workneh, 1992; Capote et al., 2004; Amills et al., 2008). To exploit objectively, the genetic potential and diversity of indigenous livestock, a proper classification and understanding of the genetic diversity of indigenous livestock populations is very vital (FAO, 2007b; Mahmoudi et al., 2011).

Understanding genetic diversity and structure forms the basis for achieving genetic improvement, formulating strategies for the sustainable utilization of farm animal genetic resources through genome-wide association analysis, genomic selection and the dissection of quantitative traits (Kijas et al., 2009; Uzzaman et al., 2014). It is also important in mapping Mendelian traits, in investigating patterns of linkage disequilibrium, and in evaluating the consequence of selection and genome-wide selection as a method to accelerate genetic gain in livestock (Kijas et al., 2009).

Very limited efforts have been made to identify Ethiopian goat types including understanding the inherent genetic diversity and structure using state of the art molecular tools such as SNP genotyping and full-genome sequence analysis. Like indigenous cattle (Edea et al., 2013) and sheep populations (Gizaw et al 2007), the analysis of genetic diversity and structure in Ethiopian indigenous goats has also been limited and those efforts showed there is high genetic diversity in Ethiopian goats but weak population structure (Chenyambuga et al., 2004; Tesfaye, 2004; Hassen et al., 2012a; Solomon, 2014). Except Solomon (2014), all did the analysis with microsatellite markers in which this marker type only spans on the neutral portions of the genome and the markers are less dense and hence less informative. A critical assessment of Ethiopian goats has not been undertaken using SNP genotype data. In here, the genetic diversity and structure of 14 Ethiopian indigenous goat populations were analyzed using SNP genotypes generated using the *Caprine* 50k SNP CHIP. This is the first comprehensive study targeting indigenous goat populations in Ethiopia and Africa based on high density SNP genotyping except a work on three goat populations in Morocco. In addition, the comparative analysis of Ethiopian and Chinese goats (Cashmere and its Ibex-hybrid population) included in the study

gives clue about regional population structure and effect of introgression of domestic goat with its wild (*Ibex*) ancestry. The later may also draw the attention to include the wild genome in the breeding programs in the future. In general, this study was aimed to evaluate genetic diversity and structure of Ethiopian goat populations.

4.2 Materials and methods

4.2.1 Sample collection and DNA extraction

A total of 468 samples from 14 indigenous goat populations in Ethiopia (n = 378), Chinese Ibex-Cashmere hybrid (n = 30), Mongolian Cashmere (n = 30) and European Ibex (n = 30) were genotyped for the study. The indigenous goat populations indicated in section 3.2.1 together with Barka goat which was sampled in Shire and Shiraro areas were included in the study. Physical description of the goat populations is presented by FARM-Africa (1996) and Mekuriaw et al. (2016). The classification of Ethiopian indigenous goats that was proposed by FARM-Africa (1996) and Gizaw (2009) was followed while sampling the 14 Ethiopian indigenous populations (Table 6; Figure 6). Representative picture of each goat population is depicted in Figure 9. Samplng material (i.e. blood), genomic DNA extraction protocol and DNA quality assessment are indicated in section 3.2.1.

Table 6 Summary of sampled goat populations and their classification (FARM-Africa, 1996)

No.	Population	Family	Production system	Agro-ecology
1.	Small eared Somali	Somali	Pastoral	Arid and Semi-arid
2.	Long eared Somali	Somali	Pastoral	Arid and Semi-arid
3.	Hararghe Highland	Somali	Mixed livestock	Humid
4.	Afar	Rift valley	Pastoral	Arid and Semi-arid
5.	Abergelle	Rift valley	Mixed livestock	Sub-humid
6.	Woyto-Guji	Rift valley	Pastoral	Arid and Semi-arid
7.	Arsi-Bale goat	Rift valley	Mixed livestock	Humid
8.	Nubian	Nubian	Mixed livestock	Arid and Semi-arid
9.	Gondar	Small East African	Mixed livestock	Humid
10.	Gumuz	Small East African	Mixed livestock	Sub-humid
11.	Agew	Small East African	Mixed livestock	Humid
12.	Ambo	Small East African	Mixed livestock	Humid
13.	Kaffa	Small East African	Mixed livestock	Humid
14.	Barka	Nubian	Mixed livestock	Arid and Semi-arid

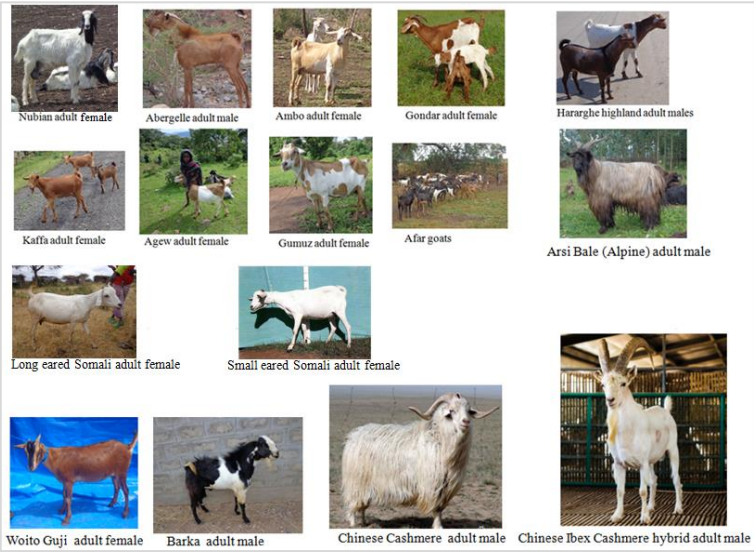


Figure 9. Representative pictures from each goat population

4.2.2 Genotyping, quality management and marker selection

The DNA samples were genotyped using the *Caprine* 50K SNP BeadCHIP (Tosser-Klopp et al., 2014; Illumina Inc., San Diego, CA). Genotyping of the Ethiopian goats was performed by DNA Land Marks (www.dnalandmarks.com), Canada and the Chinese goats by Sangon Biotech (www.sangon.com) Shanghai, China. Genotypes with call rates of ≥ 95%, Hardy-Weinberg equilibrium cut off of $P < 0.000001$ and minor allele frequency (MAF) ≥ 0.05 were used in the analysis. These quality control criteria resulted in 49,795 SNPs from 53,347 SNPs present in the CHIP for downstream analysis. The large number of SNPs that were retained for the final analysis indicates most likely the little influence of ascertainment bias of the panel in the data set as also observed by Nicoloso et al. (2015). Only autosomal SNPs, a total of 47474 loci (46520 polymorphic) were used for analysis. SNPs on the X chromosome were included for the SNPs dynamics analysis. For admixture analysis, only SNPs that were at HWE were employed.

4.2.3 Data analysis

Population genetic diversity and differentiation: Arlequin ver 3.5.1.2 (Excoffier and Lischer, 2010) was also used to calculate F_{IS}, H_O, H_E, effective migration rate per generation (Nm:

Slatkin, 1995) and to test for deviations from Hardy-Weinberg equilibrium. Same package was used to calculate Reynolds' (Reynolds et al., 1983) and F_{ST} (Weir and Cockerham, 1984) pair-wise genetic distances between populations. Minor allele frequency (MAF) was calculated using gPLINK package Ver.2.050 (Purcell et al., 2007). To test whether the classification of Ethiopian goats proposed by FARM-Africa (1996) could be supported by genetic data, the partition of genetic variation among different groups between the Ethiopian goat populations was evaluated based on the analysis of molecular variance (AMOVA) as implemented in Arlequin. The regional goat population differentiation (among Ethiopian and Chinese goats) was also carried out using the same package.

Phylogenetic and population structure analysis: To investigate the genetic relationships between populations, using allele frequency differences between populations, the principal components analysis (PCA) was carried out with SNPRelate package of R. Furthermore, pair wise population differentiation (F_{ST}) and Reynolds' genetic distances were used to reconstruct the Neighbour-net network and neighbour joining (NJ) phylogenetic tree using SplitsTree ver. 4.10 (Huson and Bryant, 2006) and MEGA6 (Tamura et al., 2013), respectively. To complement the PCA analysis, population structure was investigated using ADMIXTURE ver. 1.23 (Alexander et al., 2009). The number of hypothetical pseudo-populations (K) tested was ranged between $2 \leq K \leq 15$. Cross-validation (CV) error rates were computed for each K using a 5-fold cross-validation procedure. These were then used to evaluate the most optimal partitioning of the population genetic structure.

4.3 Result

4.3.1 Population genetic diversity

In the current study, the average minor allele frequency (MAF) value was 0.289 and ranged from 0.276 in Cashmere to 0.298 in Nubian (Table 7). Similarly, the observed (H_O) and expected (H_E) heterozygosity ranged from 0.351 (Kaffa) to 0.408 (Barka) and 0.366 (Cashmere) to 0.407 (Barka), respectively with an average values of 0.375 ± 0.1 (H_O) and 0.383 (H_E) across the 16 study populations. On the other hand, both regression coefficients of heterozygosity and geographic distances are negative showing a decrease in genetic diversity as someone moves far away from the entry point (Figure 10). This suggests that goats could have entered through the east as well as the north part of the country. This observation is strengthened by the results of the correlation analysis (east entry point: Kendall's *tau* = -0.436, calculated *P* = 0.019; Spearman's *rho* = -0.526, calculated, *P* = 0.026; north entry point: Kendall's *tau* = -0.106, calculated *P* = 0.306; Spearman's *rho* = -0.155, calculated, *P* = 0.298) between heterozygosity and geographic distances.

The F_{IS} value, a proxy of the population level of inbreeding, ranged from -0.020 (Ibex-Cashmere hybrid) to 0.073 (Nubian goat). Amongst the Ethiopian populations, only Barka had a negative F_{IS} value in addition to the Ibex-Cashmere hybrid and the Cashmere breed indicating a deficiency of heterozygotes in the three populations. The overall average F_{IS} value for the Ethiopian goats was 0.018 indicating a relatively low level of inbreeding. The highest value of F_{IS} was 0.045 in the Kaffa population while the lowest was 0.000 in Gondar population indicating it to be completely outbred.

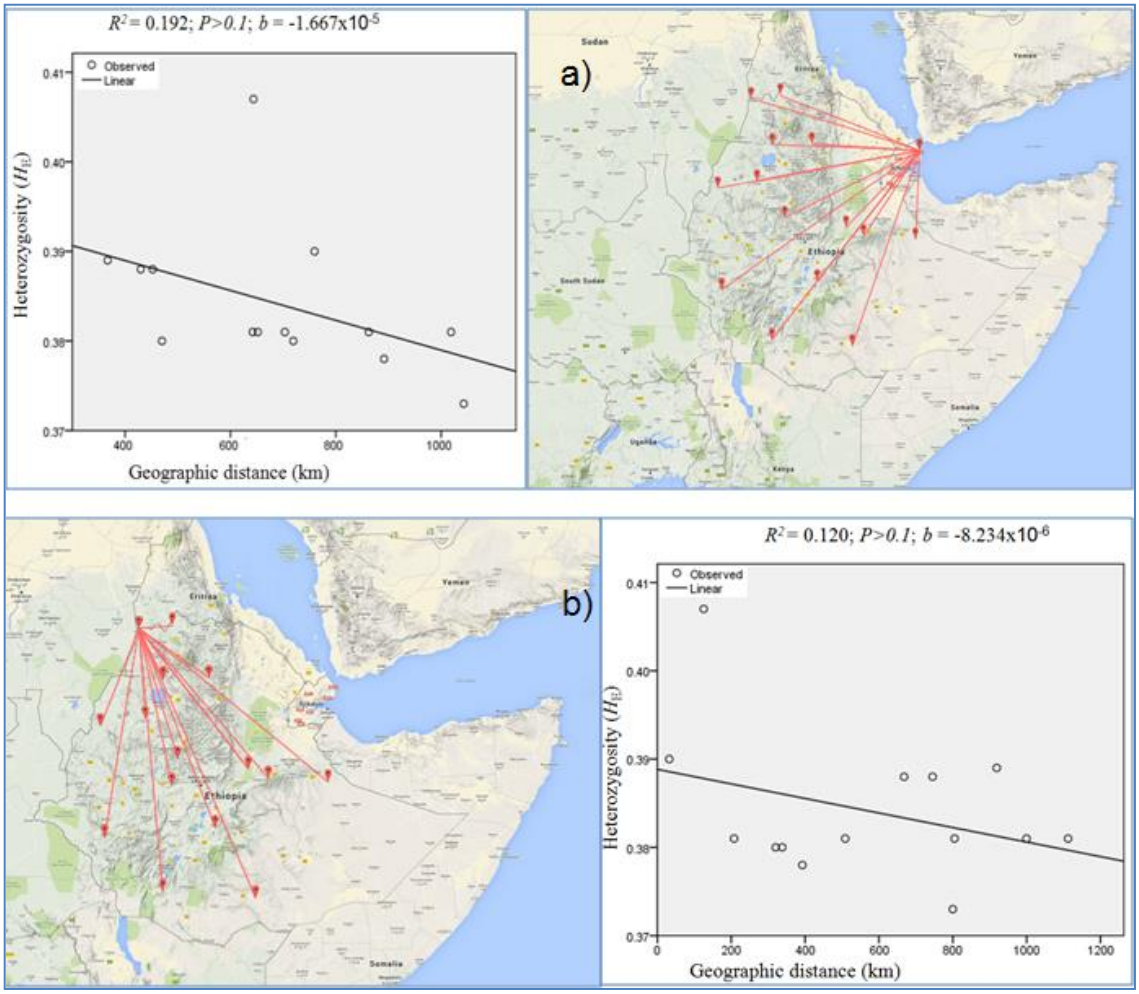


Figure 10 Regression analysis of heterozygosity vs. geographic distances from possible entry points

4.3.2 Level of SNPs polymorphism

The overall average value of monomorphic and polymorphic loci were 1749.19 ± 627.37 (3.5%) and 48000 ± 627.52 (96.5%), respectively in all the goat populations studied but excluding the European Ibex (Table 7). Both the highest number of monomorphic SNPs (49,111 constituting 98.6%) and the lowest (684 SNPs constituting 1.4% of the total number of SNPs and of which 13.5% deviated from HWE) polymorphic SNPs were detected in the European Ibex populations. A subsample of 95 Ibex individuals from three populations (Albris and Rheinwald, n = 48; Cape Moine, n = 24; Weisshorn, n = 23) were re-analyzed, and similar results of monomorphic and

polymorphic loci were obtained (the figures not indicated). This led to exclude the European Ibex populations from the downstream analysis. In the Ethiopian and Chinese goat populations, the proportions of monomorphic and polymorphic loci ranged from 1.6% - 6.7% and 93.3% - 98.4%, respectively.

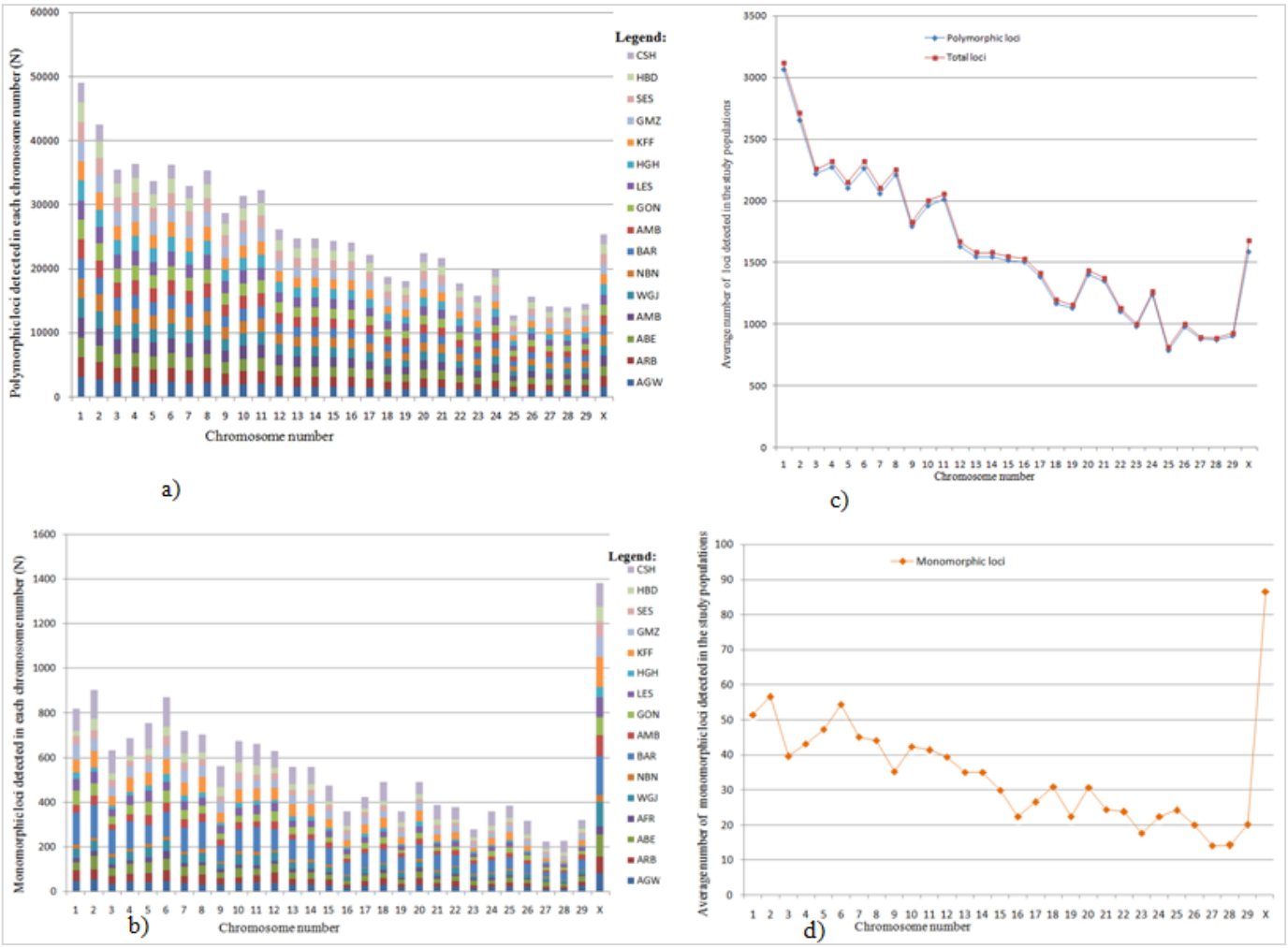
Except for Afar, Nubian and the Chinese Ibex-Cashmere hybrid, comparatively higher levels of monomorphic loci were detected in the remaining 14 goat populations studied. This suggests a drift effect in the populations analyzed. The highest proportion of monomorphic loci was observed in Barka goat (6.7%) followed by Kaffa goat (4.8%).

Table 7 Genetic variability within goat populations

Population	N	H_O	H_E	F_{IS}	MAF(μ)	Monomorphic loci (%)	Polymorphic loci (%)	% SNPs not in HWE (P≤0.05)
Agew	28	0.373	0.380	0.018	0.287	1885(3.8)	47911(96.2)	1366(2.9)
Arsi-Bale	29	0.367	0.381	0.034	0.288	1717(3.5)	48079(96.6)	1557(3.2)
Abergelle	30	0.373	0.380	0.011	0.286	1834(3.7)	47961(96.3)	0169(0.4)
Afar	33	0.378	0.388	0.022	0.295	1059(2.1)	48737(97.9)	1564(3.2)
Woyto-Guji	25	0.373	0.381	0.008	0.287	1787(3.6)	48008(96.4)	0653(1.4)
Nubian	34	0.359	0.390	0.073	0.298	0878(1.8)	48917(98.2)	3327(6.8)
Barka	8	0.408	0.407	-0.013	0.296	3339(6.7)	46456(93.3)	0461(1.0)
Ambo	30	0.371	0.381	0.011	0.286	1688(3.4)	48107(96.6)	2445(5.1)
Gondar	27	0.378	0.381	0.000	0.288	1812(3.6)	47984(96.4)	1446(3.0)
LES	27	0.378	0.381	0.002	0.287	1774(3.6)	48022(96.4)	1670(3.5)
HGH	29	0.381	0.388	0.016	0.295	1183(2.4)	48613(97.6)	1437(3.0)
Kaffa	30	0.351	0.373	0.045	0.281	2357(4.8)	47438(95.6)	2846(6.0)
Gumuz	27	0.371	0.378	0.013	0.288	2022(4.1)	47774(95.9)	1582(3.3)
SES	20	0.379	0.389	0.017	0.293	1514(3.0)	48282(97.0)	1356(2.8)
Hybrid*	30	0.384	0.382	-0.020	0.293	0788(1.6)	49008(98.4)	2434(5.0)
Cashmere	30	0.373	0.366	-0.017	0.276	2350(4.2)	47446(95.3)	0204(0.4)
Mean±SD		**0.375±0.1**	**0.383±0.00**	**0.014±0.02**	**0.289±0.01**	**1749.19±627.37 (3.5±10)**	**48000±627.52 (96.5±10)**	**1532.31±903.14 (0.032±0.02)**
European Ibex	30	0.329	0.315	-		49111(98.6)	684 (1.4)	92(13.5)

Key: *= Hybrid of Chinese Ibex and Xi-Jiang Cashmere goat; LES=Long eared Somali goat, SEL=Small eared Somali goat; HGH=Hararghe Highland goat

Genome-wide, the highest (3062.6) and lowest (787.8) average number of polymorphic loci were detected in Chromosome 1 and 25, (Appendix Table D; Figure 11) respectively. Similar trends were reported for nine Canadian goat populations (Brito et al., 2015). This can be attributed to the sizes of the two chromosomes, chromosome 1 (154.929 Mb) being the longest and chromosome 25 (41.478 Mb) being the shortest (Brito et al., 2015). However, the average highest number of monomorphic loci was detected in X-chromosome (86.6) followed by Chromosome 2 (56.6) and chromosome 1 (51.4). In all the goat populations studied ~3000 polymorphic loci were on the first chromosome. On the other hand, the 96.27% of polymorphic loci observed in the Cashmere were shared with the Chinese Cashmere-Ibex hybrid which supports the introgression of the Ibex with domestic goats (Appendix Table D and E). On the other hand, the highest proportions of loci which deviated from HWE were observed in the Nubian (6.8%) and Kaffa (6%) goats, respectively. Only 0.4% of polymorphic loci deviated from HWE in Abergelle and Cashmere goat populations.



a) polymorphic loci; b) monomorphic loci; c) average number of total and polymorphic loci; d) average number of monomorphic loci

Figure 11 Graph of loci detected across chromosomes

4.3.3 Genetic differentiation and structure

AMOVA revealed that 11.92% of the total genetic variation was explained by the genetic differences between the Chinese and Ethiopian goat populations (Table 8). This is higher than the value of 5.8% that was reported between three geographical groupings of intercontinental populations of sheep (Kijas et al., 2009). AMOVA performed among three groups of Ethiopian

goat populations as proposed by FARM-Africa (1996) revealed a variation of less than 1% with more than 94% being the result of differentiation between individuals within the respective goat populations. This suggests that the three criteria proposed for classification of Ethiopian goats into three groups seems not to be in agreement with the genetic structure of the populations. Similarly, the variation among populations within groups was 2% suggesting that the populations and groups may not be reproductively isolated. The overall average F_{ST} value among Ethiopian goat populations was 0.026, whereas the values of F_{ST} estimates among the Ethiopian goats with the Cashmere-Ibex hybrid and the Cashmere respectively were 0.15 and 0.17 (Table 9). Similarly, the Reynolds' genetic distance among Ethiopian goat populations, Ethiopian goat with *Ibex*-Cashmere Hybrid and the Cashmere goats were 0.0265, 0.1615 and 0.193, respectively.

Table 8 Analysis of MOlecular variance (AMOVA): based on different group set

Groups	**Variance components (%)**			
	AG	APWG	AIWP	WP
Ethiopian and Chinese goat populations	11.92	2.83	1.45	83.80
Ethiopian goats grouped in two production systems	0.64	2.31	2.22	94.82
Ethiopian goats grouped in three agro-ecologies	0.56	2.22	2.23	94.99
Ethiopian goats grouped in four goat families	1.02	2.15	2.20	94.63

Key: AG = Among groups; APWG = Among populations within groups; AIWP = Among indiv. within population; WP = within populations; The fixation indices are significant ($P<0.001$) in all the group set

Among the Ethiopian goat populations, the estimates of F_{ST} and Reynold's genetic distance ranged between 0.34% to 4.93%, and 1.1% to 5.1%, respectively. The lowest F_{ST} estimates (< 0.1%) were observed between the Small eared Somali with Hararghe Highland, and between the

Afar and Long eared Somali goat. The Kaffa goat showed relatively higher level of genetic differentiation (F_{ST} = 3.12% to 4.93%; Reynold's genetic distance = 3.2% to 5.1%) from other Ethiopian goat populations. Gondar and Ambo populations were previously grouped together as Central Highland goat (FARM-Africa, 1996). In the same study, the Nubian and Barka were grouped together as the Nubian goat family. However, in the current study, the lowest values of F_{ST} and Reynolds' genetic distances (0.008 each) were observed between Gondar and Abergelle compared to between Ambo and Gondar goat populations. Overall, genetic distance increases as geographic distance increases (Figure 12).

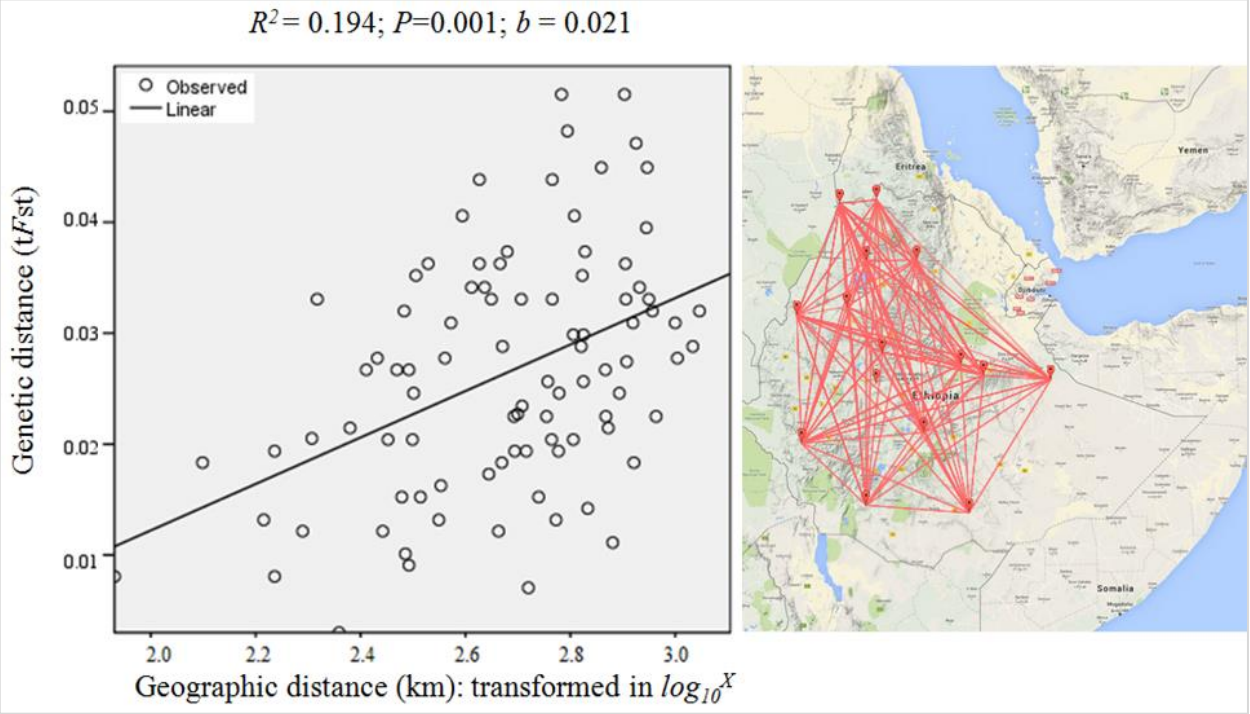


Figure 12 Regression analysis of genetic distance vs. geographic distance

Table 9 Populations differentiation: pair-wise (F_{ST}) (below diagonal) and Reynolds' (above diagonal) genetic distance

Population	AGW	ARB	ABE	AFA	WOG	NBN	BAR	ABO	GON	LES	HGE	KAF	GMZ	SES	Hybrid	CASH
AGW		0.022	0.020	0.034	0.029	0.036	0.027	0.012	0.013	0.036	0.020	0.037	0.019	0.026	0.171	0.198
ARB	0.022		0.022	0.026	0.021	0.033	0.025	0.015	0.019	0.026	0.012	0.036	0.033	0.018	0.166	0.193
ABE	0.020	0.022		0.030	0.030	0.035	0.021	0.015	0.008	0.034	0.017	0.044	0.033	0.024	0.169	0.196
AFA	0.033	0.026	0.030		0.026	0.026	0.020	0.028	0.029	0.020	0.008	0.051	0.043	0.009	0.141	0.165
WOG	0.028	0.020	0.030	0.025		0.031	0.028	0.023	0.027	0.016	0.013	0.032	0.036	0.011	0.168	0.195
NBN	0.035	0.032	0.034	0.025	0.030		0.018	0.033	0.033	0.031	0.021	0.051	0.040	0.022	0.138	0.162
BAR	0.027	0.024	0.021	0.020	0.027	0.018		0.022	0.020	0.028	0.014	0.047	0.036	0.018	0.145	0.175
ABO	0.012	0.015	0.015	0.027	0.0229	0.032	0.0222		0.010	0.030	0.013	0.033	0.024	0.020	0.167	0.193
GON	0.013	0.019	0.008	0.028	0.0267	0.0320	0.0201	0.010		0.032	0.015	0.039	0.026	0.023	0.167	0.194
LES	0.035	0.026	0.033	0.019	0.016	0.031	0.028	0.029	0.031		0.012	0.047	0.044	0.007	0.164	0.190
HGE	0.019	0.012	0.017	0.008	0.013	0.021	0.014	0.013	0.015	0.012		0.035	0.030	0.003	0.147	0.172
KAF	0.036	0.035	0.043	0.049	0.031	0.049	0.045	0.033	0.039	0.046	0.034		0.043	0.039	0.188	0.216
GMZ	0.019	0.032	0.032	0.042	0.036	0.039	0.035	0.024	0.026	0.043	0.029	0.042		0.033	0.178	0.206
SES	0.026	0.018	0.024	0.009	0.011	0.022	0.018	0.020	0.022	0.007	0.003	0.038	0.032		0.152	0.178
Hybrid	0.157	0.153	0.155	0.131	0.155	0.129	0.135	0.154	0.154	0.151	0.137	0.172	0.163	0.141		0.104
CASH	0.180	0.175	0.178	0.152	0.177	0.150	0.161	0.176	0.176	0.173	0.158	0.194	0.186	0.163	0.098	

Key: AGW=Agew, ARB=Arsi-Bale, ABE=Abergelle, AFA=Afar, WOG=Woyto-Guji, NBN=Nubian, BAR=Barka, ABO=Ambo, GON=Gondar, LES=Long eared Somali, HGE=Hararghe highland, KAF= Kaffa, GMZ= Gumuz, SES= Small eared Somali, Hybrid= Ibex-Cashmere hybrid, CASH= Cashmere

4.3.4 Principal component and cluster analyses

Figure 13 shows the PCA plots for all the goat populations analyzed. PC 1 and 2 differentiated the Cashmere goat from the Cashmere-Ibex hybrid population while PC 1 and 3 differentiated the Kaffa goat from the rest of the Ethiopian goat populations, which were poorly differentiated.

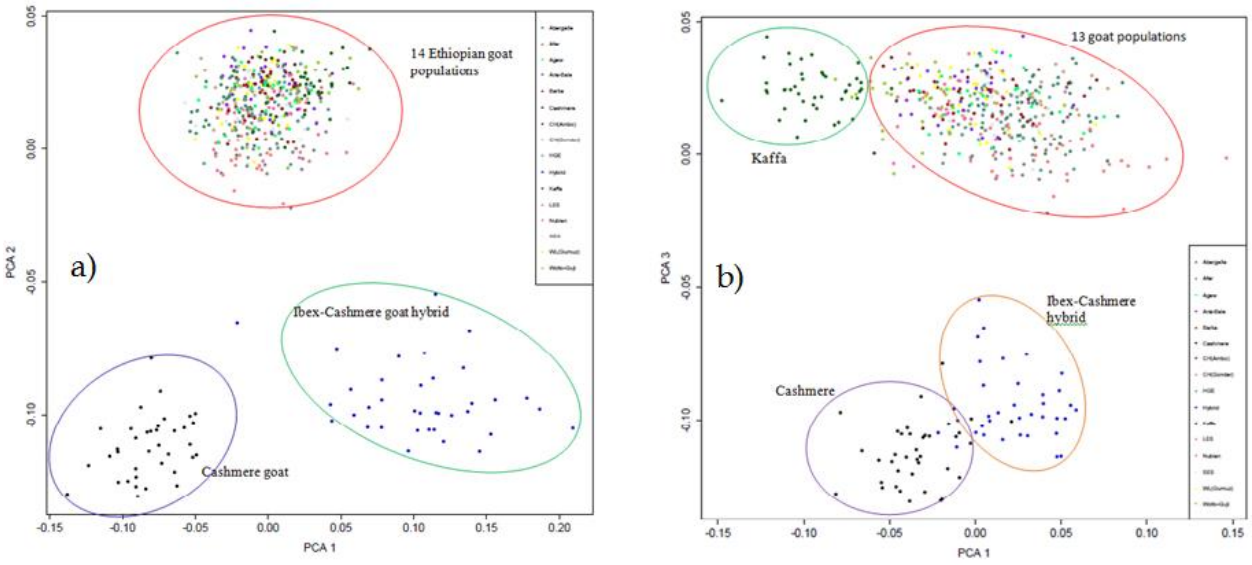


Figure 13 Principal component analysis of the goat populations studied: a) PCA 1 and 2; b) PCA 1 and 3

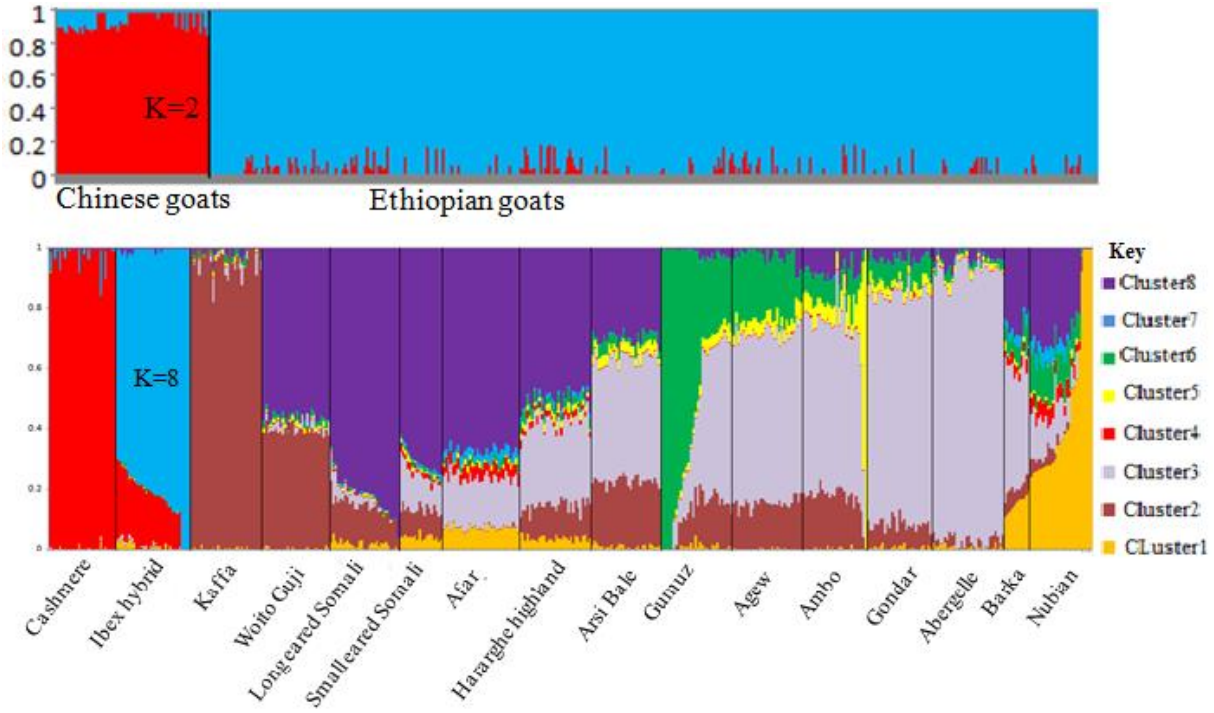


Figure 14 Bayesian clustering performed with ADMIXTURE software

On the other hand, a graphical representation of admixture analysis is indicated in Figure 14. A total of 15 hypothetical population clusters (*K*) were tested in the analysis. The lowest value of cross-validation error was attained at $K = 8$ (Appendix Figure B) indicating this to be the optimal number of genetic clusters explaining the variation in the populations. The Chinese and Ethiopian goats were differentiated at $K = 2$. Appendix Table F illustrates the proportion of genetic clusters in each population. Approximately 44.8 % of Nubian and 13.8% of Barka belong to cluster 1. The highest proportion (40.65%) of genetic background of Barka goat is found in cluster 3, which is predominated by genetic background of Abergelle goat (90.14%). The closest ancestor of Barka was proposed to be the Nubian goat (FARM-Africa, 1996), but this seems not to be supported by the admixture results. The rest of the populations share less than 7% of genetic background in cluster 1; and 95.39% of Kaffa goat genetic background belongs to cluster

2 and this genetic background is shared by other Ethiopian goat populations at different proportions except Afar goat. Geographical proximity may cause for the highest share of same genetic background by Woyto-Guji (39.12%) with Kaffa goat. Similar rationale may work between Woyto-Guji (56.23%) and Long eared Somali (81.85%) goats at cluster 8. Cluster 3, where Abergelle goat is predominantly nominated, is shared by most of Ethiopian goat populations.

Similarly, previous studies indicated that both Ambo and Gondar goat populations were grouped under Central Highland goat. However, the current study revealed that Gondar goat (its 77.04% genetic background belongs to cluster 3) is more close to Abergelle than with Ambo. This finding is strengthened by the lowest estimates of F_{ST} and Reynolds' distances obtained in this study. For the latter population, 12.73% of genetic background belongs to cluster 5 which is not observed in any other Ethiopian goat populations as high proportion as Ambo goat.

On the other hand, most proportion (58.58%) of genetic background of Gumuz goat is found in cluster 6. Agew goat shares 22.52% genetic background in cluster 6. Home tracts of the two goat populations are very tied up that may facilitate ease of flock exchange between farmers. The goat populations found in South, South--east and East parts of Ethiopia share highest proportion of similar genetic background in cluster 8 (Long eared Somali: 81.85%, Small eared Somali: 70.72%, Woyto-Guji: 56.23%, Afar: 66.95% and Hararghe Highland: 49.81%) of genetic background in cluster 8. This implies that livestock movement and/or exchange is very high in the regions where these goat populations are found. The highest level of population migration per generation (Nm =24) strengthen this idea (Table 10). The goat populations found in the

regions, except Hararghe Highland, are lowland goats and the regions are characterized as arid and dry environment. Being lowland goats encourage the animal exchange among the regions since individual animals can easily adapt the new environment. However, from the same region, Arsi-Bale goat population shares the least proportion of genetic background (29.50%) in cluster 8 that deviates the premises of influence of close geographical proximity. The reason could be Arsi-Bale goat is Afroalpine-Subafroalpine goat which lives up to 4000 m.a.s.l. This might limit the animal exchange with the lowland goats. Instead, 40.96% of its genetic background is comprised in cluster3 where genetic backgrounds of North West-Central Highland and Tekeze valley goat types (Ambo, Agew, Abergelle and Gondar) are constituted. Overall, Kaffa, Abergelle, Ibex-Cashmere hybrid and Cashmere goat populations differentiated clearly. This goes in line with the PCA, as described above. Indeed, the 15.97% of genetic background of Ibex-Cashmere hybrid is constituted in cluster 4 where 97.41% of Cashmere goat genetic background is clustered. The lowest proportion of the Ibex-Cashmere hybrid observed in this cluster could be because of presence of backcrossing.

Table 10 Effective rate of migration among the study goat populations: Based on Slatkin (1995) linearized F_{ST} derivation

Population	AGW	ARB	ABE	AFR	WGJ	NBN	BAR	AMB	GON	LES	HGH	KFF	GMZ	SES	HBD	CSH
Agew(AGW)	0.00															
Arsi-Bale (ARB)	23.17	0.00														
Abergelle (ABE)	25.11	23.12	0.00													
Afar (AFR)	14.52	19.21	16.48	0.00												
Woyto-Guji (WGJ)	17.12	24.37	16.32	19.27	0.00											
Nubian (NBN)	13.77	15.34	14.25	19.62	16.12	0.00										
Barka (BAR)	18.61	21.12	24.27	24.85	18.16	27.73	0.00									
Ambo (AMB)	41.17	33.72	32.69	17.88	21.13	15.09	22.37	0.00								
Gondar (GON)	37.29	27.13	64.00	17.04	18.17	15.13	25.53	49.77	0.00							
LES	14.18	19.29	14.66	25.51	30.48	15.84	17.46	16.63	15.41	0.00						
HGH	25.62	42.55	28.62	65.04	37.42	23.60	36.63	38.48	32.87	41.32	0.00					
Kaffa (KFF)	13.15	13.68	11.04	9.69	15.58	9.65	10.53	14.73	12.43	10.51	14.07	0.00				
Gumuz (GMZ)	26.59	15.19	15.17	11.33	13.47	12.23	13.85	20.37	19.08	14.59	16.70	11.48	0.00			
SES	19.07	28.80	20.51	57.13	44.15	22.29	28.87	24.27	22.02	72.77	168.69	12.56	15.00	0.00		
HBD	2.68	2.78	2.72	3.32	2.72	3.38	3.21	2.75	2.75	2.80	3.16	2.41	2.57	3.06	0.00	
Cashmere (CSH)	2.27	2.34	2.31	2.79	2.31	2.83	2.61	2.34	2.33	2.38	2.65	2.07	2.18	2.56	4.76	0.00

Key: LES= Long Eared Somali; HGH = Hararghe Highland; SES=Small Eared Somali; HBD=Chinese Ibex-Xi Jiang goat (Cashmere) hybrid

4.3.5 Phylogenetic tree/ network analysis

The Neighbour-Net network constructed using the population pairwise F_{ST} values and the NJ tree reconstructed using the Reynolds' genetic distance (Figures 15 and 16) revealed nine phylogenetic groups. Group 1: Ibex-Cashmere hybrid; group 2: Mongolian Cashmere; group 3: Small eared Somali, Long eared Somali, Woyto-Guji and Afar (Afar and Small eared Somali goats are separated by five population sub-divisions); group 4: Ambo and Agew (separated only by three population sub-divisions); group 5: Abergelle and Gondar (separated only by four population sub-divisions); group 6: Gumuz; group 7: Nubian and Barka; group 8: Kaffa. Hararghe Highland goat (the only goat population sound on the internal node) was basal to group 3 and Arsi-Bale emerged out at the side of the same group independently, and these two goat populations are considered as mixed type and last group. The fit indices for the split network (Fit = 98.17 for F_{ST} distance; fit = 97.76 for Reynolds' genetic distance) indicate high robustness of the network and tree-likeness of the data (Robles-Sikisaka et al., 2012).

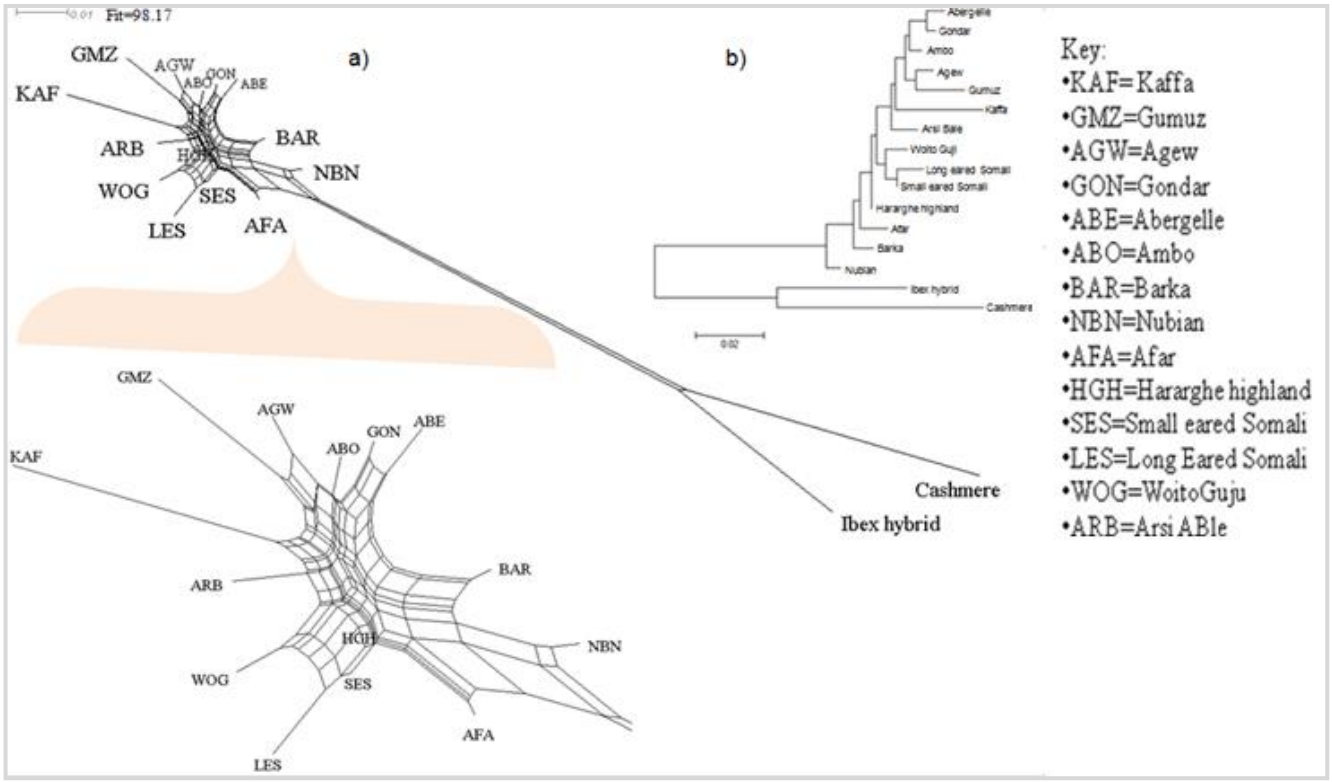


Figure 15 Phylogenetic network (a) and phylogenetic NJ-tree (b): based on F_{ST} distance

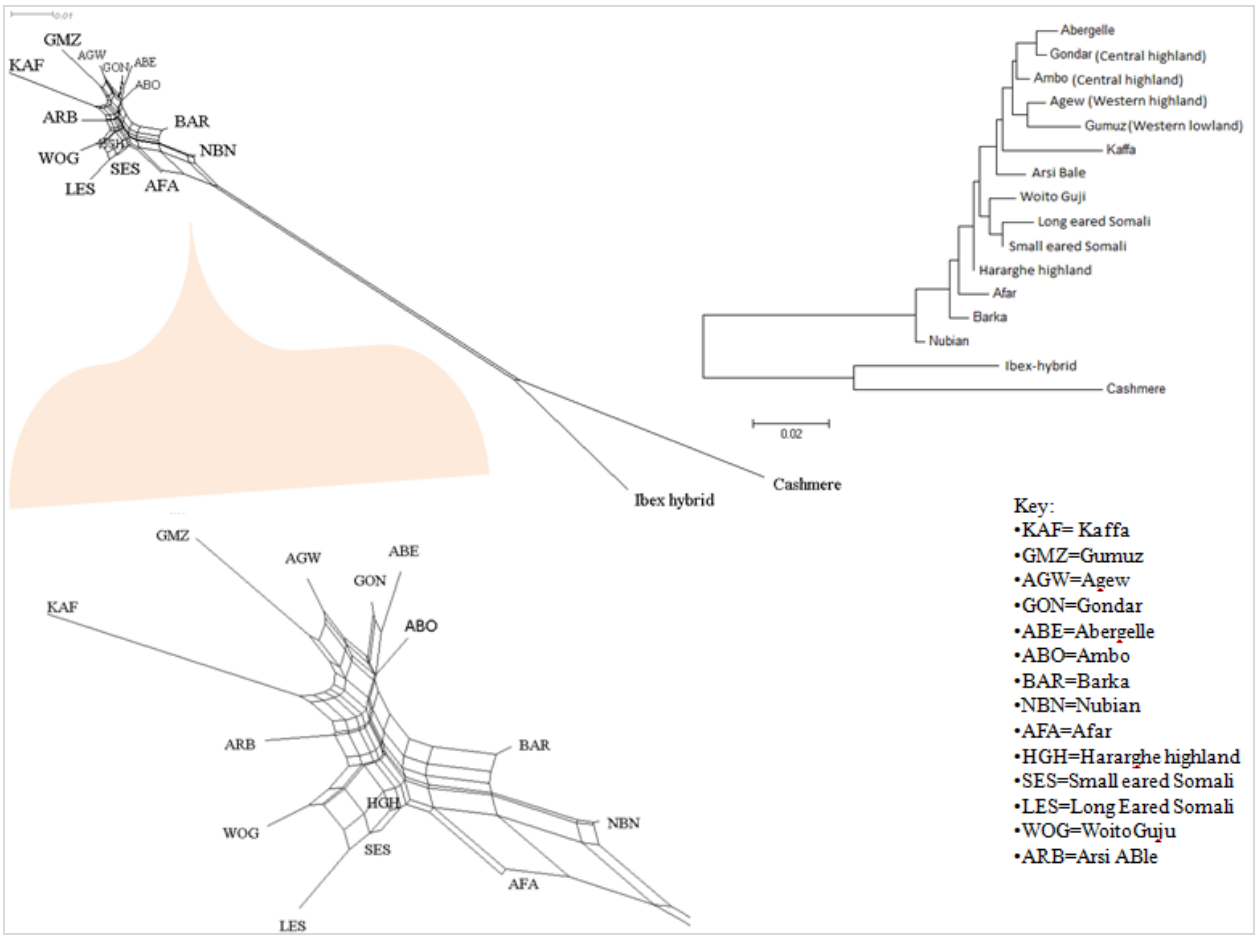


Figure 16 Phylogenetic network (a) and phylogenetic NJ-tree (b): based on Reynolds' distance

4.4 Discussion

Various molecular tools and markers have been employed to study genetic diversity and structure in livestock species. Here, using the 50k SNP CHIP genotype data was generated and used to carry out an extensive analysis of the genetic diversity and structure of Ethiopian indigenous goats. For comparison analysis, two Chinese goat populations were included as reference or out-group populations.

4.4.1 Nature of SNPs' polymorphism and genetic diversity

From the total 53347 SNPs genotyped for each goat population studied, 93.3% - 98.4%, of the SNPs were polymorphic suggesting the presence of high level of genetic polymorphism in the study populations despite the fact that the design of the *Caprine* 50k SNP CHIP panel did not include any feral goat population as well as the wild ones (Tosser-Klopp et al., 2014). The estimates are comparable to those obtained for Italian goats (96.76 – 99.7%; Nicoloso et al., 2015), and the Ethiopian Gumuz goat (Solomon, 2014) but comparable to the estimate observed in Australian goat (> 97%; Kijas et al., 2013) all of them were analyzed with the 50k SNP CHIP. However, the range of estimates generated in the current study were higher than those observed in *Bos taurus* cattle (79%; Dadi et al., 2012) and in indigenous Ethiopian cattle populations (83.36%) analyzed using the 8k CHIP (Edea et al., 2012) but are comparable with the values reported for Hereford cattle (95%) based on the Bovine 50k SNP Beadchip analysis (Matukumalli et al., 2009).

From the panel of 53,347 SNPs present in the CHIP, 46,520 autosomal SNPs that passed the quality control thresholds were used for downstream analysis. Given variable quality control

parameters and thresholds employed in the respective literatures, this number of SNPs is almost equivalent to the one obtained in the South African Angora goats (46, 983; Lashmar et al., 2015), in the French dairy goat (46,959; Carillier et al., 2013), and in UK dairy goats (47,306; Mucha et al. 2014) but it is slightly higher than the 45,268 SNPs used in Canadian goat populations that were genotyped using with the same CHIP (Brito et al., 2015). The number is however lower than the 49,156 SNPs used in Florida dairy goats where MAF was not mentioned as one of the quality control parameter (Zidi et al., 2014) and 51,136 SNPs which passed the quality control for Italian goat populations (Nicoloso et al., 2015). The comparative number of SNPs obtained in the current study after quality assessment shows the utility of the SNP CHIP in studying feral goat populations, as also indicated by Huson et al. (2014). It also indicates a high level of genetic diversity despite the weak population differentiation between Ethiopian indigenous goat populations considered in this analysis. The expected (H_E) and observed heterozygosity (H_O) obtained in the current study were comparable with those reported for Abergelle and Gumuz goats in Ethiopia (Solomon, 2014) and Italian goats (Nicoloso et al., 2015), but higher than the average heterozygosity estimates of New Zealand and Australian goats (Kijas et al., 2013).

4.4.2 Population differentiation and admixture, and identified goat types

In the current study, the analysis of molecular variance revealed that most of the variation in Ethiopian goats was due to between individual variability (94%). This is comparable to the value of 93.7% (Solomon, 2014) reported for two Ethiopian goat populations with the same SNP chip; but slightly higher than the within population variation (89.83%) of Italian goats (Nicoloso et al., 2015). The variation between individuals within populations reduced to 83.80% when the Chinese goats were included in the analysis of the current study. The highest within population

variation could be explained by presence of uncontrol mating together with absence of selection and breeding strategy (Gizaw et al., 2008).

Formerly, morphological and phenotypic information were used to classify the indigenous goat populations in Ethiopia (FARM-Africa 1996). However, grouping based on this traditional approach is not generally supported with molecular marker based classification and grouping. With different proportions, the influence of Kaffa goat genetic background, the only goat population adaptive to the humid and forest area in Ethiopia, was observed in all Ethiopian goat populations except in Afar goat. In addition, low estimate of *N*m was observed among Kaffa goat with the rest of Ethiopian goats. The Kaffa goat home tract, which is highly tsetse infested area, is highly humid to be adapted by other Ethiopian goat types. The place is similar with the area where West African dwarf goat, which is trypanotolerant (Chiejina et al., 2015), is found. In connection to this, decades ago, it was reported that the Nilotic dwarf goat of southern Sudan (the present South-Sudan), which is a member of the dwarf goats of Central and West Africa, extends its breeding area the north, west and south goats of Ethiopia could have fallen under the influence of this goat (Epstein, 1971). However, additional evidences are required to arrive at firm conclusion and confirm that Kaffa goat could be extended from West Africa region. Moreover, farmers, in Kaffa goat home tract, explained that the goats are tolerant to tsetse (personal communication) like their Sheko cattle which duly invites researchers to confirm this perception.

Barka goat was classified under Nubian goat family and believed as it is descendant of Nubian goat (FARM-Africa, 1996); however, the admixture analysis showed it shares highest genetic background from Abergelle and Gondar goat populations (Figure 14) and the gene flow

estimation with Nubian goat was found to be equivalent (even lower) with some other Ethiopian goat populations, like Gondar, Hararghe Highland and Small eared Somali (Table 10). Phylogeny trees of the F_{ST} and Reynolds' distances (Figure 15 and 16) support this observation and strengthen the assertion that Barka goat is getting intimidated in Ethiopia. In addition to the diverse populations' migration events, the long lasting civil war held since the last century in the Barka goat area could be mentioned among the contributing factors for the threat (Zerabruk *et al*., 2007; Getinet and Adebabay, 2015). However, considering two genetic backgrounds (cluster 1 and 8 of the admixture result) and the few number of nodes (population sub-divisions) observed in the phylogeny trees of the F_{ST} and Reynolds' distances Barka goat is grouped together with Nubian goat than with Abergelle and goat populations. Nei et al (1977) remarked the influence of small sample size on the estimation of rate of migration (*m*) which could have an implication on Barka goat population in the current study. However, this remark was forwarded based on analysis of protein polymorphism that assumes patterns of migration among populations are stable over time to infer patterns of gene flow (Slatkin, 1981). Besides, stable pattern of livestock migration, following the human population dynamics, is unlikely to occur in the context of Ethiopia.

On the other hand, in both pair wise (F_{ST}) and Reynolds' distance estimation methods, highest estimates were observed among Ethiopian and Chinese goat populations (Table 9). These estimations are higher than the intercontinental total mtDNA variation (10%) of domestic goats (Luikart et al., 2006). The Chinese Ibex-Cashmere hybrid is closer to Ethiopian populations than Cashmere goat does. This might be due to Ethiopian goat populations are autochthonous in which they have not been developed targeting any trait. Whereas, the Cashmere is allochthonous

population developed for Cashmere production. Similar finding was reported in sheep recently (Tesfaye, 2015). Autochthonous populations are deemed to be representative of the wild state and sometimes be close to the wild populations (Lv et al., 2014) like obtained in this study. Both the PCA and admixture analyses revealed that the Ethiopian and Chinese goat populations have been differentiated perfectly, as supported by output of the AMOVA. However, very minimal differentiation, which does not show clear structure, has been observed among Ethiopian indigenous goats. Similar observation was reported for Moroccan goats (Benjelloun et al., 2015). In contrast, other recent studies indicated clear differentiations among the goat populations studied with the same SNP CHIP panel (Kijas et al., 2013; Brito et al., 2015; Nicoloso et al., 2015).

Arsi-Bale and Hararghe Highland goats seem composite populations. Particularly, the latter is observed at the internal nodes and the former is emerged out alone from the nodes which represent sub-populations. Moreover, the admixture analysis showed that both goat populations have significant proportions of three genetic backgrounds, and are considered as mixed goat type. Based on the admixture analysis, the influence of Gumuz goat population, the only goat population represented in cluster 6,was observed in Agew, Ambo, Gondar, Barka and Nubian goat populations at various proportions. In the same cluster, some individuals from Gumuz goat were clearly appeared and the network graphs in both pair wise (F_{ST}) and Reynolds' distance estimations differentiated this goat population from others. Hence, the goat population is considered as Wet Lowland goat type considered the wet lowland area it is found. In the former report, Abergelle goat was classified under the rift valley goat family and was reported as it is descendant of the Rift Valley goat types from South East Asia (FARM-Africa, 1996). However,

the admixture analysis in the current study does not support this, little or no genetic backgrounds shared from previous classification of Rift Valley goats; rather it seems its current genetic structure is shaped through continuous breeding among highland goats and developed its unique drought tolerance character through time. Despite the agro-ecology variation, Abergelle and Gondar goat populations share 90.14% and 77.04% similar genetic background from cluster 3 and, besides, both populations are separated only by four population sub-divisions. The closest geographical proximity which they share same watershed (Tekeze Valley watershed) could have contributed to have similar genetic background and are considered as Tekeze Valley goat type.

From the eight clusters (Ks) represented by different colors (Figure 14), the yellow color (i.e. cluster 5) have not been represented by any of the populations except by one or two individuals of Ambo, which could be because of sampling bias. This goat populations share 71% similar genetic background at cluster 2 and 3 with Agew goat. The few number of nodes (population sub-divisions) observed in network graphs support this result. Both goat populations are highland goats with similar agro-ecology and altitudes. Geographically, they are close each other and considered are North West-Central Highland goat type. Similarly, Afar, Long eared Somali, Small eared Somali and Woyto-Guji goat populations share 56.23-81.85% similar genetic background at cluster 8. All goat populations are found in dry and lowland areas and considered are Dry Lowland goat type.

Overall, inconsistent with former classifications (agro-ecologies, production systems and goat families) high level of population admixture that could be because of inflated level of gene flow per generation were observed in this study. Moreover, all the current goat populations are formed from continuous interbreeding of several populations in recent times. The multiple internal nodes

(N=192) and short length edges (335) from 16 goat populations (Figure 15a and 16a) strengthen this argument. The multiple median vectors (N=137) and the NJ tree (Figure 4) obtained on control region of mtDNA in the former chapter support this observation. Animals having more diverse estimates of breed composition (less than 75% of their genes coming from a single breed) might be because of recent admixture event (Brito et al, 2015) which could be essential for identification of certain QTL present in only one breed (Larmer et al. 2014). Similarly, exchange of high level of genetic background was observed among Agew and Gumuz goat populations (Figure 14).

In this study, the contribution of geographical isolation-by-distance was inconsistently observed; rather, evidences of the anthropological history of the country have been reflected in the current genetic architecture of Ethiopian indigenous livestock populations. For instance, the gene flows among Agew goat with Hararghe Highland and Arsi-Bale goats (geographically far apart) were very high. Similar results were also observed in Ethiopian indigenous cattle population studied using protein polymorphism (Sisay, 1996) and microsatellite markers (Dadi et al., 2008). In addition, Mpofu (2002) also reported as the cattle found around Lake Tana were named by Tana land Boran (today's Fogera cattle) by the tribe who moved from South-east to North-west parts of Ethiopia. Therefore, it is most unlikely to believe that the people moved only with their cattle rather they moved together with their other livestock species including goats. The anthropological and anthropogenetic studies also state that during 16^{th} century, different tribes moved from the South-East to the North-West Ethiopia (Yilma, 1967; and Habitamu, 2014).

From the anthropogenetic perspective, the recent mtDNA and Y-chromosome analyses (Christopher, 2011) and 1M SNP CHIP array (Pagani et al., 2012) showed that ethnic groups in

Ethiopia nearly significantly differentiated from each other; however, the pattern of similarities indicates some recent gene flow between northern ethnic groups and some groups in the south. Therefore, the wave of human population movement experienced for many centuries together with the past and present weak livestock management schemes and geographical proximity have favoured gene flow among livestock species including goats (Hassen et al., 2012a; Edea et al., 2013; Solomon, 2014). However, it does not necessarily mean that geographical isolation-by-distance did not influence the current genetic background of the goat populations. The regression analysis between genetic pairwise difference and geographic distance resulted negative regression coefficients indicating a decrease in genetic diversity as one moves far away from the possible entry point via northern as well as eastern Ethiopia (Figure 10 and 12).

Indeed, there are huge observed morphological and phenotypic differences among Ethiopian goat populations (FARM-Africa, 1996; Hassen et al, 2012b; Grum et al 2013; Hulunm, 2014; Netsanet, 2014; Alubel, 2015; Hussein, 2015). However, these could be explained by developmental homeostasis mechanisms that generate variation in body shape corresponding to an optimal size for fitness (Brown et al., 1973) and the underlying genetic structure could be shaped by natural selection. In addition, coat color variation is highly associated with ecological variation (Brehem et al., 2001; Gizaw et al., 2007) and strongly responds to natural selection (Thorpe et al., 1996).

CHAPTER V

5 Analysis of kisspeptin (KISS1) gene polymorphism and its association with multiple births

5.1 Introduction

Reproduction traits are among crucial economic traits in animal husbandry, and are coordinated during normal puberty and the adulthood (Zhang et al., 2011). Among reproduction traits, genetic studies have indicated that the litter size and ovulation rate can be genetically determined by the action of genes (Deldar-Tajangookeh et al., 2009). However, it is stated that not only for fitness traits (e.g. litter size), the identification of candidate genes that are responsible for variation in continuous traits (e.g. growth traits) has been a challenge in modern genetics (An et al., 2013). As effect, to date, little has been divulged on the major genes associated with, for instance, litter size in goats. Efforts conducted related to the litter size in goat are the inhibin alpha-subunit gene (INHA) (Hua et al., 2008; Wu et al., 2009), the gonadotrophin releasing hormone receptor gene (GnRH) (An et al., 2009), the bone morphogenetic protein receptor-IB gene (BMPRIB) in the prolific Indian Black Bengal goat (Polley et al., 2009), the bone morphogenetic protein 15 gene (BMP15) in Jining Grey goats (Chu et al., 2007) and the kisspeptin (KISS1) gene in Xinong Saanen, Guanzhong and Boer goat populations in China (An et al., 2013).

Some studies were focused only on polymorphism evaluation (Bai et al., 2005; Gupta et al., 2007) and did not include effects of the gene expression on the respective traits. According to An et al. (2013), KISS1 gene that encodes kisspeptin highly contributes for multiple births in goat. It is expressed in the hypothalamus region of the forebrain (Gottsch et al., 2009b). Signal of

kisspeptin and its receptor G-protein coupled receptor ligand (GPR54) has an important role in initiating secretion of gonadotropin-releasing hormone (GnRH) (Dungan et al., 2006; Smith et al., 2006), and is now recognized as being essential for normal fertility by regulating the reproductive system (De Roux et al., 2003; Funes et al., 2003; Kirilov et al., 2013).

The central or peripheral administration of kisspeptin stimulates GnRH-dependent luteinizing hormone (LH) and follicle-stimulating hormone (FSH) secretion in various mammalian species from rodents to humans and it also administers reproductive functions of animals (Gottsch et al., 2004; Navarro et al., 2005; Shahab et al., 2005). However, with these all regulatory functions of the KISS1 gene and its receptor, failure or loss of function or deletion of, for instance, GPR54 causes to hypogonadotropic hypogonadism (De Roux et al., 2003; Chu et al., 2012). It is a deficiency of the pituitary secretion of FSH and lLH which cause impairment of pubertal maturation and reproductive function.

In general, despite the limited efforts done in small ruminant livestock, the expression and regulation of KISS1 gene plays magnificent role in multiple births in goat. This study might serve as an additional input for further evaluation and utilization of the gene in marker assisted selection breeding program. Therefore, this study was initiated to assess the polymorphic nature and role of KISS1 gene on multiple births in Gondar and Woyto-Guji indigenous goat populations in Ethiopia.

5.2 Materials and methods

5.2.1 The study goat populations

Two indigenous goat populations, Gondar and Woyto-Guji, were selected for the study. The two populations are found in the extreme north and south corners of the country, 1350km (ground distance) far apart. The agro-ecology and the production systems of both populations are quite different. Woyto-Guji goat population is found in arid environment. The production system is categorized predominantly by agro-pastoral production system with some crop production in practice (Workneh, 1992; Netsanet, 2014). Whereas, Gondar goat population is found in mid-and high-altitude areas. The area is characterized by mixed-crop livestock production system.

The blood samples were drawn out from the jugular vein with a volume of 9 ml under aseptic conditions using ethylene diamine tetraacetic acid (EDTA) anticoagulant. The collected samples were brought to the laboratory with ice box and were stored at -20^{0}_{C} until it was subjected to DNA extraction using salting out protocol (Shinde et al., 2008). The DNA quality and concentration were tested by nanodrop, and 1.0 - 1.5% agarose gel electrophoresis was used to evaluate the degradation.

5.2.2 Primers employed, target regions and PCR conditions

For PCR amplification, three pairs of primers were newly designed from sequences of *Ovis aris (*Acc:HGNC:6341*)* and *Capra hircus (Acc.* GU142847.1) KISS1 genes obtained in the data base, DNA Data Bank of Japan (DDBJ) and were evaluated by Oligo Analyzer 3.1: Integrated DNA technology package (available at: https://goo.gl/IfGA2F) (Appendix Table G). Exon1 has a length of 1,210 bp and exon2 has 325 bp. For both exons touch down PCR program was employed with the respective PCR conditions (Appendix Figure C). The purified PCR products

were sequenced using the Big Dye Terminator v3.1 cycle sequencing chemistry (Applied Biosystems) and the 3130XL automatic capillary sequencer (Applied Biosystems, USA). The sequences were aligned with complete *cds* of Jining Grey goat population kisspeptin (KISS1) gene available in the GenBank (accession no. GU142847).

5.2.3 Data management, statistical analysis and packages employed for analysis

All the chromatograms were generated and visualized with the CLC workbench 7.0.4 (CLC Bio-Qiagen). Multiple sequence alignments were done in CLC working bench employing the ClustalW algorithm (Thompson et al., 1994) and edited manually in MEGA6 (Tamura et al., 2013). Variable sites were scored/called against Jining Grey goat population sequenced by Cao et al. (2010). In total, 173 sequences (58 in Gondar and 115 in Woyto-Guji) for exon 1 and 242 sequences (117 in Gondar goat and 133 Woyto-Guji goat) were generated for exon1 and exon2, respectively. Estimates of genetic diversity measures, like heterozygosity at loci level, were computed for KISS1 gene for both goat populations with Arlequin ver.3.0 (Excoffier et al., 2005). The codon bias index (CBI) and GC content were evaluated using DnaSP 5.0 software (Rozas et al., 2003).

The association study had been evaluated by using segregating sites (SNPs). Literatures (Shifman et al., 2002; Beaty et al., 2005; Balding, 2006; Gong et al., 2007; Pei et al., 2009) suggest that haplotype analysis is more powerful than using markers for association study; however, haplotype analysis has not been employed for this study due to the following basic reasons: i) number of haplotypes was small. This is because the few number of SNP markers obtained in the target region. ii) The size of sequenced region, which led to have shortest physical distance among segregating sites, was short. Short physical distances among

polymorphic regions are not suggested for association study (Pritchard and Przeworski, 2001). iii) Estimates of LD correlation coefficient (R^2) obtained for both goat populations (0.083 for Woyto-Guji and 0.081 for Gondar population) were very low. iv) Evans and Cardon (2005) also explained that low *D´* estimation between populations suggest individual estimates of pairwise *D´* are likely to be of limited use in guiding association mapping. Therefore, the lowest estimates obtained (-0.181 for Woyto-Guji and -0.114 for Gondar population) had also led to stick to the marker based association analysis. v) SNPs which have no/lack of strong linkage disequilibrium are not likely appropriate for genetic association studies; i.e. in regions of high LD, a reduced set of haplotype tag SNPs may be selected to detect efficient associations between variations in that gene or region and a trait of interest (Beaty et al., 2005; Gong et al., 2007).

The phenotype data were associated with information of the polymorphic sites using SAS ver.9.1. General linear model (GLM) was used to analyze the association of genotype with phenotypic performances of the populations studied. i) Exon1: $Y_{ijklmn} = \mu+P_i+B_j+G1_k+G2_l+G3_m+e_{ijklmn}$; where, P = Parity at i^{th} (i= 1, 2, 3, 4, 5) parity of the doe gave birth; B=Population at j^{th} population (j = Gondar, Woyto-Guji); G1= Site1 at k^{th} genotype (k = GG, GC); G2= Site2 at l^{th} genotype (l = CC, CT, TT); G3 = Site3 at m^{th} genotype (m = CC, TC, TT). ii) Exon2: $Y_{ijkl}=\mu+P_i + B_j +G_k+ e_{ijkl}$. Parities after the fifth parity have been merged as the fifth parity because of fewer observations and non significant influence of parties after the fifth parity. Genotypes other than indicated in the model have been excluded from the model since they showed non-significant effects and did not contribute for fitness of the model.

5.3 Result

5.3.1 Detection of single nucleotide and codon usage

In this study, complete substitutions with respect to the reference sequence (Gene bank: acc. GU142847) were observed at five sites (at g.3436 T >C, g.3592C>A, g.3688A>C, g.3878 A >C and g.4023A>C), of which four of them were *transversions*. In addition, twelve polymorphic sites in exon1 and three in exon2 were detected (Table 11) in which ten of them (at 945, 950, 3354, 3533, 3649, 3696, 3808, 3811, 3963 and 3989 loci) were *transition* mutations and the remaining five were *transversion* mutations. On the other hand, most of the detected polymorphic sites were observed in more than 1% of the sequences that qualifies the SNP definition. Moreover, from the usable nucleotides, the GC contents were 64.09% for exon1 and 62.47% for exon2. The codon bias index, which is species-specific deviation from uniform codon usage in the coding regions of genomic sequences, was estimated to be 0.301. Five of the polymorphic sites detected were caused to be degenerative codons and most of them were very few in number for association study though they qualified the SNP definition. Hence, only four loci (g.950T>C; g.3416, g.3811 C>T and g.3963T>C) that constitute 10 genotypes were remained for the association analysis (Table 13). The genotypes are TT and TC at g.950T>C, CC and GC at g.3416G>C, CC, TC and TT genotypes at g.3811C>T and TT, TC and CC genotypes at g.3963T>C.

Table 11 Detected SNPs and IUPAC symbols (N)

Locus(→)	Nucleotide in the reference sequence	IUPAC symbol (N)	Nucleotide substituted (N)	Nucleotide similar with the reference (N)	Total
g.0895G>C	G	-	C(239)	C(3)-	242
g.0945C>T	C	-	T(3)	T(239)	242
g.0950T>C	T	Y(6)	-	C(236)	242
g.3354A>G	A	R(1)	G(172)	-	173
g.3416G>C	G	S(8)	C	G(165)	173
g.3436T>C	T	-	C(173) †	-	173
g.3533A>G	A	R(1)	-	A(172)	173
g.3592C>A	C	M(173)	A†	-	173
g.3649G>A	G	R(51)	A(41)	G(81)	173
g.3688A>C	A	M(173)	C†	-	173
g.3696C>T	C	Y(1)	-	C(172)	173
g.3770G>C	G	S(1)	C	G(172)	173
g.3783T>A	T	W(2)	A	T(171)	173
g.3808G>A	G	R(53)	A(58)	G(62)	173
g.3811C>T	C	Y(22)	T(3)	C(148)	173
g.3878A>C	A	M(173)	C†	-	173
g.3927C>G	C	S(2)	G	C(171)	173
g.3963T>C	T	Y(38)	C(4)	T(131)	173
g.3989G>A	G	R(54)	A(39)	G(60)	173
g.4023A>C	A	M(173)	C†	-	173

Key: †=Complete substitution in all sequences of the study populations; IUPAC=International Unit for Pure and Applied Chemistry

5.3.2 Estimation of heterozygosities

In exon1, the overall average expected heterozygosity (H_E) estimation of all the polymorphic loci was 0.18863±0.21. The estimation of Gondar goat population (H_E = 0.25287±0.20) is almost equivalent with H_E of Woyto-Guji goat population (H_E = 0.25182±0.22). Some SNPs showed modest levels of heterozygosity, whereas, this estimation in other SNPs in exon1 (at locus 3649, H_E =0.50155and at locus 3989, H_E = 0.50007) approached and slightly above 50%, theoretical maximum heterozygosity (Tokarska et al., 2009). All the H_E estimations were greater than 1%

(Table 12). However, most of the polymorphic loci were obtained in low H_E frequencies and lack strong linkage disequilibrium estimates (data not indicated).

Table 12 Expected heterozygosity (H_E) estimates of exon1 and exon2 of KISS1 gene of Woyto-Guji and Gondar goat population in polymorphic loci

Exon 1				Exon 2			
Locus (→)	**Over all H_E (n=173)**	**H_E for Woyto-Guji (n=115)**	**H_E for Gondar (n=58)**	**Locus (→)**	**Over all H_E (n=173)**	**H_E for Woyto-Guji (n=133)**	**H_E for Gondar (n=117)**
g.3354A>G	0.01156	-	0.03448	g.895G>C	0.02381	0.04443	-
g.3416G>C	0.08872	0.06773	0.13067	g.945C>T	0.02381	0.04443	0.01352
g.3533A>G	0.01156	0.01739	-	g.950T>C	0.04704	0.05878	0.03389
g.3649G>A	0.50155	0.50252	0.50817				
g.3696C>T	0.01156	-	0.03448				
g.3770G>C	0.01156	0.01739	-				
g.3783T>A	0.02299	0.03448	-				
g.3808G>A	0.46256	0.46773	0.45977				
g.3811C>T	0.24869	0.25416	0.24198				
g.3927C>G	0.02299	-	0.06776				
g.3963T>C	0.36981	0.40519	0.29038				
g.3989G>A	0.50007	0.49977	0.50817				
Average	**0.18863±0.21**	**0.25182±0.22**	**0.25287±0.20**		**0.03155±0.01**		

Table 13 Amino acid changes observed in polymorphic sites

Locus	Amino acid in the reference sequence		Amino acid in the sequences studied		Amino acid position	N	Remark
	Amino acid	Codon†	Amino acid	Codon			
g.0895G>C	Arginine	CGG	Glycine	GGG	R0298G	3	
g.0945C>T	Tryptophan	UGG	Cysteine	UGU	W0315C	3	
g.0950T>C	Stop codon	UAG	Tryptophan	UGG	*0317W	6	**
g.3354A>G	Cysteine	UGC	Cysteine	UGU	C1118C	1	
g.3416G>C	Alanine	GCC	Glycine	GGC	A1139G	8	**
g.3436T>C	Serine	AGC	Glycine	GGC	S1146G	173	
g.3533A>G	Methionine	AUG	Threonine	ACG	M1179T	1	
g.3592C>A	Alanine	GCC	Serine	UCC	A1198S	173	
g.3649G>A	Cysteine	UGU	Arginine	CGU	C1217R	94	
g.3688A>C	Stop codon	UAG	Glutamic acid	GAG	*1230E	173	
g.3696C>T	Threonine	ACA	Threonine	ACG	T1232T	1	
g.3770G>C	Serine	UCG	Tryptophan	UGG	S1257W	1	
g.3783T>A	Proline	CCA	Proline	CCU	P1261P	2	
g.3808G>A	Histidine	CAC	Tyrosine	UAC	H1270Y	111	
g.3811C>T	Glutamic acid	GAA	Lysine	AAA	E1271K	25	**
g.3878A>C	Valine	GUG	Glycine	GGG	V1293G	173	
g.3927C>G	Threonine	ACG	Threonine	ACC	T1309T	2	
g.3963T>C	Isoleucine	AUA	Methionine	AUG	I1321M	42	**
g.3989G>A	Serine	UCA	Leucine	UUA	S1330L	93	
g.4023A>C	Valine	GUU	Valine	GUG	V1341V	173	

Key: N=number of individuals the mutations observed at the respective locus; †=codon with respect to the reference sequence; **= contributed for litter size; Note that thymine should be change to Uracil in the codon columns

5.3.3 Amino acid substitutions and their association to litter size

In exon1, from the five sites where complete nucleotide substitutions were detected the following amino acid changes were observed on sequences of the populations studied: serine to glycine at locus g.3436T>C, alanine to serine at locus g.3592C>A, stop codon to glutamine at locus g.3688A>C and valine to glycine at locus g.3878A>C (Table 13). However, no amino acid changes were observed at loci g.3354A>G, g.3696C>T, g.3783T>A, g.3927C>G and g.4023A>C because of the degenerative codons (Table 13). Similarly, the remaining mutations at g.0895G>C,

g.0945C>T, g.3533A>G, g.3649G>A, g.3770G>C, g.3808G>A and g.3989G>A did not contribute for the fitness of the model suggesting these mutations do not influence the litter size although they caused amino acid changes. Therefore, mutations only at g.950T>C (mutation at stop codon), g.3416G>C, g.3811C>T and g.3963T>C were considered for the association study (Table 14). Hence, the analysis of least square mean and standard error (LSM±SE) revealed that there was a 24% increment of litter size in GC genotype individual animals than GG genotypes at locus g.3416G>C in the goat populations studied. Similarly, 18% (TT and TC), 22% (CC and TT) and 31% (TC and CC) litter size differences were observed at loci g.950T>C, g.3811C>T and g.3963T>C, respectively in the animals studied. On the same analysis, the highest (1.614±0.14) and lowest (1.307±0.14) litter size were obtained in the 3rd and 1st parties in exon 1 and 1.696±0.11 and 1.219±0.10 in exon 2, respectively. Similarly, higher estimate of litter size was observed in Gondar goat than in Woyto-Guji goat population.

Table 14. Least square mean standard error (LSM±SE) estimation of fecundity trait

Exon1			Exon2		
Factors	**N**	**LSM±SE**	**Factors**	**N**	**LSM±SE**
Overall mean	177	1.403±0.12	**Overall mean**	242	1.435±0.09
Parity		***	**Parity**		***
1	40	1.307± 0.14^{c}	1	50	1.219±0.10^{e}
2	34	1.328± 0.15^{c}	2	47	1.300±0.11^{d}
3	39	1.614±0.14^{a}	3	47	1.696±0.11^{a}
4	32	1.448±0.14^{b}	4	49	1.537±0.11^{b}
5	32	1.319±0.14^{c}	5	49	1.425±0.11^{c}
Population		***	**Population**		***
Gondar	115	1.681±0.13^{a}	Gondar	111	1.688±0.10^{a}
Woyto-Guji	62	1.126±0.13^{b}	Woyto-Guji	131	1.182±0.09^{b}
Genotype at g.3416G>C		***	**Genotype at g.950T>C**		***
GG	169	1.284±0.09^{b}	TT	236	1.349±0.03^{b}
GC	8	1.523±0.19^{a}	TC	6	1.520±0.18^{a}
Genotype at g.3811C>T		***			
CC	151	1.539±0.12^{a}			
CT	22	1.355±0.17^{b}			
TT	4	1.316±0.25^{c}			
Genotype at g.3963T>C		***			
TT	134	1.350±0.14^{b}			
TC	38	1.581±0.14^{a}			
CC	5	1.279±0.23^{c}			

Key: ***=Significant at 0.1% (P<0.001)

5.4 Discussion

5.4.1 Analysis of polymorphisms and heterozygosity

In the current study, most of the identified SNPs qualified the SNP definition. The bi-allelic form of variation at a specific location in the genome could be a SNP if it is found in more than 1% of the population (Brookes, 1999). The author also stated that SNPs are abundant forms of genome variation, distinguished from rare variation by a requirement for the least abundant allele to have a frequency of 1% or more population (Brookes, 1999). However, in rare cases, there are tri- or tetra-allelic forms for SNPs at a specific location (Kim and Misra, 2007) though this case was not observed in the current study. In line with this, 39 tri-allelic polymorphisms were detected in sheep (Kijas et al., 2009).

Most of the polymorphisms were observed on the first exon, of which three of them were detected in exon2. In contrast, no mutations were observed in the latter exon in girls (Luan et al., 2007) and in goat (Feng et al., 2009 and Cao et al., 2010). Instead, six polymorphisms (G296C, G454T and T505A in intron 1; G3433A and C3688A in exon 3 and a 18 bp deletion/insertion in 1960–1977 site in intron 2) were identified on the same gene in five goat breeds (Cao et al., 2010). Similarly, An et al (2013) reported ten polymorphisms (g.1147T>C, g.1417G>A, g.1428-1429delG, g.2124C>T, g.2270C>T, g.2489T>C, g.2510G>A, g.2540C>T, g.3864_3865delCA and g.3885_3886insACCCC) on the KISS1 gene of three Chinese goat populations. In contrast to the current study, mutation at g.384G>A was limited only to Saanen and Guanzhong goat populations (An et al., 2013). Mutation at T2643C and 8 bp base deletions (2677AGTTCCCC) in the intron2 of the KISS1 gene in goat were reported (Hou et al., 2011). Introns have regulatory function and do not have direct involvement in the regulation of transcription of highly expressed

genes (Behzadi et al., 2015); however, systematic differences in motif distributions suggest that introns play a role in the rate of their transcription (Zhang et al. 2008b). Moreover, SNPs in the non-coding regions are required for evolutionary genetic studies by serving as genetic or physical markers (Kim and Misra, 2007).

In the present study, the *transition* to *tranversion* ratio, in both exon1 and exon2 was 2:1. In former report, G to C *transversion* at site 296, G to T *transversion* at site 454, T to A *transversion* at site 505, G to A *transition* at site 3433 and C to A *transversion* at site 3688 were reported in goats (Cao et al, 2010). Similarly, 2.44:1 SNPs *transition* to *transversion* ratio was detected in the genome wide sequences analysis of three Moroccan goat populations (Benjelloun et al., 2015). *Transition* type substitutions may occur more frequently than *transversion* type substitutions (Kimura, 1980). Moreover, from the usable nucleotides, highest proportions of GC contents were obtained indicating important influence of the codons in the target gene expressions (Bernardi et al., 1985; Ikemura, 1985) in the current study. On the other hand, the average CBI (0.301) obtained is equivalent with the value of the index (0.302) in the complete *D-loop* analysis of Ethiopian indigenous goat populations indicated in the above section. This estimate shows presence of considerable natural selection pressure that shaped structure of the populations studied (Sharp and Li, 1987). It is also noted that natural selection favours higher expression and enhanced codon usage optimization in short genes (Fox and Erill, 2010). Frequencies of amino acids can be modified by selection; this is due to variability of tRNA abundances of functional similar amino acids which could require different metabolic costs for their production (Novoa et al., 2012; Qian et al., 2012). Previous reports indicated that pattern of codon usage in very highly expressed genes can reveal the alternative synonymous codons which

are most efficient for translation (Sharp and Li, 1987), and the pattern of amino acid usage is influenced by the genomic base composition (Lobry and Gautier, 1994; Fryxell, 2008).

On the other hand, lower estimates of H_E were detected in most of polymorphic loci in both goat populations (Table 12). This shows presence of excess rare alleles indicating recent mutations. Recent mutant is most likely to be present in a small number of individuals, a model in recent mutation often results in an excess of the number of rare alleles (*i.e.,* alleles at low frequencies) and negative F_S value (Fu, 1997). Negative *F*s values were obtained on the haplotype analysis of the KISS1 gene in the current study. Similarly, as indicated in the above section, similar negative values were observed on mtDNA analysis of Ethiopian indigenous goat populations that include Gondar and Woyto-Guji.

5.4.2 Amino acid changes and their contribution for multiple births

Efforts on animal genetic improvement consist of identifying, mapping and analyzing polymorphisms of genes involved in various metabolic pathways that facilitate growth and reproduction of the animals and delivery of the required nutrients for the respective tissues (Schwerin et al. 1995). Behzadi et al. (2015) stated that molecular genetic selection on individual genes is a promising method to genetically improve economically important traits in livestock. A wide range of genetics disciplines stand to benefit greatly from the study and use of SNPs (Brookes, 1999). Despite the agricultural and biological importance of goats, breeding and genetics studies have been hindered by the lack of a reference genome sequence (Dong et al. 2013). In particular, SNPs which occur in coding regions can change structure of protein which

ultimately alters function of the gene though they occur much less frequently in these regions (Li et al., 1991; Nickerson et al., 1998; Kim and Misra, 2007).

With this, in the current study, most of the SNPs identified in the KISS1 gene are either cause to be degenerative codon which did not show amino acid changes or have no contribution for litter size. In addition, most of them have no/lack strong linkage disequilibrium (data indicated in the next section). SNPs which have no/lack of strong linkage disequilibrium are not likely appropriate for genetic association studies since a reduced set of haplotype tags cannot be detected for efficient association between a gene and the trait of interest (Beaty et al., 2005; Gong et al., 2007). Four SNPs influence litter size of the goats studied (Table 14). However, in the analysis of variance, the relative small estimate of the fitness model (R^2 ~35%) (Appendix Table H) might be because of the fact that multiple birth can also be influenced by other genes, like INHA, GDF9, BMPR1B, BMP15 genes and also controlled by growth hormones (Chu et al., 2007; Hua et al., 2008; Polley et al., 2009; Wu et al., 2009).

In addition, growth hormone (GH) of mammals plays an important role in involving cell division, ovarian folliculogenesis, oogenesis and secretory activity (Hull and Harvey, 2002; Ola et al., 2008). By acting through specific receptors within the ovary, GH is expedient in controlling proliferation and apoptosis, oocyte maturation, and the expression and synthesis of receptors to hormones and related substances (Hull and Harvey, 2000; Sirotkin et al., 2003). Silva et al. (2009) also stated that the effect of GH on ovarian function is mainly through inducing the development of small antral follicles in the gonadotrophin-dependent stages and stimulating oocyte maturation. On the other hand, in the absence of the genotypes indicated in the model, it was fitted only to 20%. A 15% increment of the fitness model implies the

significant association of the candidate gene (the KISS1 gene) on multiple births compared to other genes, growth hormones and other phenotypic fixed effects. This shows that kisspeptin together with its receptor (GPR54) may stimulate the release of LH and FSH in female goats. Former reports also indicated that KISS1 gene is a key regulator and catalyst for the puberty onset and is a fundamental gatekeeper of sexual maturation in mammals (Hashizume et al., 2010; Cao et al., 2010 and 2011; Chu et al., 2012).

As a result, the genotypes identified on the coding regions of the KISS1 gene in the current study had shown remarkably significant ($P<0.001$) contribution (18% - 31% increment of litter size) on fecundity trait. However, this result is far lower than the finding reported for the CC genotype of Jining Grey goat does (litter size difference estimated to be 0.80 at locus 296) (Cao et al., 2010). In addition, Cao et al (*ibid*) reported that G3433A caused one amino acid change (Ala, A, GCC) to (Thr, T, ACC) at residue 86 (A86T) though non-significant influence of the genotypes (CC, CA, AA) was observed at locus 3688. In the current study, complete substitution of A>C (monomorphism) was detected at the latter locus, and genotype TC of locus g.3963T>C is the highest contributor. This could be because of the fact that the mutation occurred at this locus caused to shift to the initiation codon (AUG) and methionine is expressed. Presence of this codon may help more amino acid to be expressed in the hypothalamus region of the brain. However, genotype CC of the same locus was the least contributor for multiple births.

In other locus (T2643C), significant effect ($P < 0.05$) on litter size was reported in goats (Hou et al., 2011). In sheep, KISS1 mRNA expressing cells are found in the arcuate nucleus (ARC) and dorsallateral preoptic area and both appear to mediate the positive feedback effect of estradiol to

generate the preovulatory GnRH/LH surge (Smith et al., 2011). The luteinizing hormone (LH) surge has been associated with an increase in the LH response to kisspeptin in humans and sheep (Dhillo et al, 2007; Smith et al., 2009), indicating the surge may be generated by increased kisspeptin output and sensitivity. In human, amino acid substitutions were observed at P110T and P81R in KISS1 gene (Luan et al., 2007).

It is also possible to deduce that in the ANOVA table (Appendix Table H), population as source of variation took the highest share of variation among the fitness model suggesting the genotype component plays a vital role in regulating the reproductive cycles of female animals. Of course, joint consideration of multiple traits can provide additional information compared to information contained in individual traits as suggested by Pei et al. (2009). Overall, these all findings indicate that KISS1 gene is an excellent candidate gene for reproductive traits in goats and other livestock species.

CHAPTER VI

6 Haplotype information and linkage disequilibrium analysis of detected SNPs in kisspeptin (KISS1) gene

6.1 Introduction

Apart from the study of genomic DNA (nuclear as well as mitochondrial DNAs), the study of association of alleles, which are non random, plays fundamental roles in evolutionary and history of demographic expansion of population genetics (Fields, 2014). This nonrandom association of alleles at two or more loci that structure the genome is called linkage disequilibrium (LD) (Slatkin, 2008). LD plays a pivotal role in genomic selection, mapping quantitative trait loci (QTL), estimates for effective population size, marker assisted selection and association study (Nachman, 2002; Khatkar et al., 2008; Zhu et al., 2013).

At genome wide scale, LD can serve to uncover the population history, population characteristics, the breeding system, patterns of gene exchange and geographic subdivision (Zhu et al., 2013); whereas, at the level of genomic region/s it reflects the history of natural selection, gene conversion and mutation (Slatkin, 2008). Linkage disequilibrium is facilitated by genetic and non-genetic factors like, genetic drift, genetic linkage, mutation, selection, population structure, demographic expansion and non random mating (Majo, 2008; Zhu et al., 2013). However, these forces which affect LD in the genomic region depends on rate of recombination (Slatkin, 2008) and the extent of their effect vary from each other. For instance, natural selection affects only one or a small number of loci; by contrast, population subdivision, changes in population size and the exchange of individuals among populations affect LD throughout the genome.

Methods that directly evaluate LD by using haplotype data are more powerful than methods that examine multiple loci without evaluation of haplotype sharing (Service et al., 1999). This is because of true haplotypes are more informative than genotypes (Gong et al., 2007; Pei et al., 2009) and are more powerful than single markers for genetic association analysis due to the highest statistical power haplotype based association test has than tests using single SNPs (Shifman et al., 2002; Pei et al., 2009). Balding (2006) also mentioned LD will remain crucial to the design of association studies until whole-genome re-sequencing becomes routinely available. However, there is the issue of uncertainty of individual haplotype in haplotype based analysis, which can be resolved by haplotype phasing algorithm like haplotype trend regression analysis, which is an efficient genetic association analysis method by indicating the relationship between LD with physical distance (Sokal and Rohlf, 1981; Pei et al., 2009). In addition, the information loss that arises from phasing is small when LD is strong (Balding 2006). The pattern of LD varies across chromosomes and genomic regions (Zhu et al., 2013).

The extent and distribution of LD in livestock is becoming a center of discussion. It is because of the fact that it plays a fundamental role in gene mapping, both as a tool for fine mapping of complex disease genes and in proposed genome wide association studies (Service et al., 1999; Nachman, 2002; Slatkin, 2008). In line with this, number of markers required for a purpose like marker-trait association study and mapping is determined by the extent of LD (Abecasis et al., 2001; Khatkar et al., 2008). Moreover, if alleles at two loci are in LD and they both affect reproductive fitness, the response to selection on one locus might be accelerated or impeded by selection affecting the other (Slatkin, 2008). The ultimate value of SNPs for linkage and

association mapping studies depends, in part, on the distribution of SNP's allele frequencies and inter-marker linkage disequilibrium across populations (Goddard et al., 2000).

Analyses of haplotype and linkage disequilibrium have been carried out in various farm animals like cattle, sheep, pig and chicken. However, there is, if any, limited effort conducted on domestic goats. On the other hand, smaller effective population size and selection practices led LD to be far reached in farm animals than in human (McRae et al., 2002). However, analysis of haplotype and linkage disequilibrium at a segment of the genome has been rarely carried out in livestock species. Instead, genome wide analysis of LD has been extensively done by various scholars. This is because of the fact that the number of haplotypes would be small in a segment of DNA (Shifman et al., 2002; Beaty et al., 2005; Balding, 2006). Moreover, due to the smallest size of sequenced region which leads to have shortest physical distances among segregating sites, short physical distances among polymorphic regions are not suggested for association study (Pritchard and Przeworski, 2001). Therefore, scholars prefer to focus analysis of LD at genome wide level.

However, it does not necessary mean that a segment of DNA; i.e. a target gene, is not useful for association study. There is/are gene/s, called pleiotrophy or polygenes, which controls/control the expression of a phenotypic trait/s. Moreover, literatures confirmed that haplotypes can explain more information about an unobserved causal variant by identifying it uniquely or by identifying related haplotypes which are overrepresented among cases (Beaty et al., 2005). Therefore, polymorphism analysis on specific genes and their associations to targeted traits of interest have been carried out by various scholars. GPR54 and KISS1 genes for litter size in

goats and sheep, major histocompatibility complex (MHC) gene for immunity in goats are some of the genes on which the polymorphism and gene-trait association studies were carried out (Cao et al., 2010 and 2011; Grossen et al., 2014). In the above section (chapter 5), on the same target regions of the KISS1 gene and the SNPs observed in both Gondar and Woyto-Guji goats, the polymorphism analysis and association of the KISS1 gene with twining ability were carried out. As a follow up, this study was initiated to evaluate the haplotype diversity and extent of linkage disequilibrium of detected loci in KISS1 gene of Gondar and Woyto-Guji goat populations.

6.2 Materials and methods

The study animals, agroecologies and production systems where the populations are managed, sampling and sampling procedures, number of animals included in the study from each goat population, DNA extraction and PCR protocols are described under section 5.2.1 and 5.2.2 above.

6.2.1 Data analysis

Estimates of haplotype frequencies, measures and patterns of pairwise LD and neutrality test were computed for both Gondar and Woyto-Guji goat populations. To investigate the variability associated with fine-scale measures of LD, measures of pairwise LD decays (*R*, R^2, *D*, *D'* and |*D'*|) between adjacent markers were calculated for both populations. According to Lewontin (1974), level of linkage disequilibrium between allele A and B found at different loci can be evaluated by coefficient of linkage disequilibrium D_{AB}, which is defined as:

$D_{AB} = pAB - pApB$; *p* refers allelic frequency. This can be normalized as follows:

$D' = {}^{D}/_{D_{max}}$, where; $D_{max} = \{\min\{pApB, (1-pA)(1-pB)\}\ when\ D < 0;$

$D_{max} = \min\{PA(1-pB), (1-pA)pB\ when\ D > 0$

Correlation coefficient of linkage disequilibrium is calculated as: $r^2 = \left(\frac{D}{\sqrt{pA(1-pA)pB(1-pB)}}\right)^2$

Measures of pairwise LD decays, recombination rate, number of haplotypes and haplotype diversity estimates were analyzed by DnaSP 5.0 software (Rozas et al., 2003); whereas, loci linkage and heterozygosity estimation of haplotypes were analyzed by Arlequin ver. 3.0 (Excoffier et al., 2005). Measures of haplotype diversity were evaluated based on estimated haplotype frequencies (Beaty et al., 2005). This measure of gene diversity is analogous to the heterozygosity at a single locus and attains its maximum when haplotypes observed in the

sample occur at equal frequencies. Number of different haplotypes in each population reflects this haplotype diversity. To understand whether the observed number of unique haplotypes was different among populations, the number of expected haplotypes for each population sample was calculated first. To calculate F_{ST}, single-nucleotide polymorphisms (SNPs) contain much less information when taken one at a time (Browning and Wei et al., 2010); and hence F_{ST} values have not been tested for this study. Moreover, the preliminary analysis indicated negative result, which is unexpected and that might be because of the least information each SNP contain when it is taken at a time and the overall smaller numbers of SNPs obtained in the target region. In relation to this, it is suggested to calculate averages over windows of markers or even over the whole genome (Weir et al. 2005), which is far from the objective, approach and size of target region of this study. The neutrality tests used for the study were Tajima's *D* (Tajima, 1989b) and Fu's *F*s (Fu, 1997), which both assume infinite-site model. In addition the following models were employed evaluate the neutrality test further: *ZnS* (Kelly, 1997), *Za* (Rozas et al., 2001) and Fay and Wu's *H* (Fay and Wu, 2000).

6.3 Result

6.3.1 Assessment of haplotype diversity

A total of 29 haplotypes in exon 1 and three in exon 2 were obtained in both populations, of which only 10 of them in exon1 and two in exon2 were shared haplotypes by both populations (Table 15). The remaining haplotypes were not common for both populations. The haplotype frequencies in exon1 range from 0.0087-0.2430 and in exon2 from 0.0171-0.9830. The 2nd, 3rd and 5th haplotypes in exon1 and the 1st haplotype in exon2 registered the highest haplotype frequencies in both goat populations.

The overall gene (haplotype) diversity was 0.8703±0.0137 for exon1 and 0.0703±0.0222 for exon2; whereas the mean nucleotide diversity estimated to be 0.00275±0.001572 for exon1 and 0.00029±0.00002 for exon2 of both populations. The expected heterozygosity (H_E) estimations were not consistently higher than observed heterozygosity (H_O) estimates (Appendix Figure D).

Table 15. Haplotype frequency of KISS1 gene of the goat populations studied

Haplotype	Exon1		Exon2	
	Woyto-Guji (n=116)	Gondar (n=62)	Woyto-Guji (n=133)	Gondar (n=117)
1	0.0435	0.0172	0.9470	0.9830
2	0.2430	0.2410	0.0301	0.0171
3	0.1910	0.1550	0.0226	
4*	0.0087			
5	0.1480	0.2240		
6	0.0174	0.0172		
7	0.0261	0.0345		
8*	0.0087			
9	0.0609	0.0172		
10	0.0783	0.0690		
11*	0.0522			
12	0.0261	0.0172		
13*	0.0435			
14	0.0087	0.0345		
15*	0.0870			
16*	0.0087			
17*	0.0087			
18**		0.0172		
19* *		0.0172		
20**		0.0172		
21* *		0.0172		
22* *		0.0172		
23* *		0.0172		
24* *		0.0172		
25 **		0.0172		
26 **		0.0172		
27 **		0.0172		
28 *	0.0087			
29 *	0.0087			

Key:- * = Private haplotypes in Woyto-Guji population; ** = private haplotypes in Gondar population

6.3.2 Analysis of linkage disequilibrium and neutrality test

Linkage disequilibrium (LD) is a sensitive indicator of the population genetic forces that structure a genome (Slatkin, 2008). In this study, most estimates of *D'* and *R* were obtained below and close to zero (Appendix Table I). Similarly, the average estimates of R^2, which is the major measure of LD, was very low.

The average R^2 values were 0.083 and 0.081 for Woyto-Guji and Gondar goat populations, respectively (Table 16). Whereas, the mean value of |*D'*| were 0.656 for Woyto-Guji and 0.635 for Gondar goat. However, most relationships of the SNPs and LD measures are concentrated at the maximum value for |*D'*| (|*D'*|=1) and minimum value for R^2 (R^2=0). The average distance among segregating/polymorphic sites is comparable for both goat populations. It was estimated 226.19 bp for Woyto-Guji and 234.86 bp for Gondar goat populations. On the other hand, the |*D'*| regression showed positive relationship whereas the R^2 had negative but weak relationship with respect to the physical distances of polymorphic sites in all categories of the goat populations (Appendix Figure E).

Table 16 Descriptive statistics of measures of linkage disequilibrium

M	Woyto-Guji				Gondar			
	Mean±sd	Range	Min	Max	Mean±sd	Range	Min	Max
D	0.00440±0.06	0.31	-0.14	0.17	0.01500±0.05	0.26	-0.080	0.18
D'	-0.18052±0.74	2.00	-1.00	1.00	-0.11405±0.72	1.85	-1.000	0.85
|*D'*|	0.65633±0.36	0.86	0.14	1.00	0.63462±0.32	0.97	0.033	1.00
R	0.02410±0.29	1.26	-0.57	0.69	0.07643±0.28	1.08	-0.338	0.74
R^2	0.08334±0.15	0.47	0.0006	0.47	0.08111±0.15	0.55	0.00005	0.55
Dis.	226.19±161.86	570.00	3.00	573.00	234.86±170.64	570.00	3.00	573.00

Key: Dis.= distance among segregating sites, M=Measures of LD

Table 17 Correlation analysis of LD: Woyto-Guji (below diagonal) and Gondar (above diagonal)

	D	***D'***	***|D'|***	***R***	***R^2***
D		0.665**	0.013ns	0.952**	0.719**
D'	0.588**		-0.562**	0.795**	0.461*
|D'|	-0.105ns	-0.465*		-0.092ns	0.148ns
R	0.972**	0.728**	-0.146ns		0.759**
R^2	0.204ns	0.260ns	0.224ns	0.272ns	

Key: *=significant at 5% significant level; **= significant at 1% significant level

On the other hand, the correlation analysis indicates that there is modest to highest correlation among most of the LD measures in the goat populations studied (Table 17). For instance, *R* and *D* had shown strong correlation whereas modest correlation was observed between *R* and *D'*. However, R^2 showed non-significant correlations with all the LD measures. Negative correlations with variable power of correlations were observed between |*D'*| and *D'*, and *R* and |*D'*| in both goat populations.

Table 18 Linkage analysis observed among loci in exon1

L/L	Both goat populations simultaneously												Woyto-Guji (below diagonal) and Gondar (above diagonal)									
	3354	3416	3533	3649	3696	3770	3783	3808	3811	3927	3963	3989	3416	3533	3649	3770	3783	3808	3811	3927	3963	3989
3354	*												-		-			-	-	-	-	-
3416	-	*											*		-			-	-	-	-	-
3533	-	-	*										-	*	-			-	-	-	-	-
3649	+	-	-	*									-	-	*			+	-	-	-	+
3696	-	-	-	-	*													-	-	-	-	-
3770	-	-	-	+	-	*							-	-	-	*		-	-	-	-	-
3783	-	-	-	+	-	-	*						-	-	+	-	*	-	-	-	-	-
3808	-	-	-	+	-	-	-	*					-	-	+	-	-	*	-	-	-	+
3811	-	-	-	-	-	-	-	-	*				-	-	-	-	-	-	*	-	+	+
3927	-	-	-	-	-	-	-	-	-	*										*	-	-
3963	-	-	-	-	-	-	-	+	+	-	*		-	-	-	-	-	-	+		*	+
3989	+	-	-	+	-	+	-	+	-	-	+	*	-	-	+	-	-	+	-	-	+	*

Key: L=locus

Among the twelve loci combinations in both goat populations 16.67% of them had significant marker-marker linkage disequilibrium in exon1 of both goat populations (Table 18). Comparatively, highest significant LD accumulation was observed at 3989 locus association in Gondar goat population, and at 3649 and 3963 loci in Woyto-Guji goat population. The detection of significant linkages observed in few loci goes in line with the percentage estimations of linked loci per locus (Table 19). Similarly, association of the nine polymorphic sites in each goat population indicated that there was similar (16.67%) accumulation of linkage disequilibrium (Table 18). There was no LD accumulation detected at loci 3416, 3533 and 3770 with the respective loci combinations in Woyto-Guji goat population and at loci 3354, 3416, 3696 and 3963 in Gondar goat population. On the other hand, the overall estimated recombination rate was 0.0567 in this study. Based on Hudson (1987) test, in exon1, the estimated recombination rate was detected in five adjacent sites: (3416, 3649), (3783, 3808), (3808, 3811), (3811, 3963), (3963, 3989).

Table 19. Percentage of linked loci per locus (α (*P*) = 0.05)

Neutrality model	**Over all**	**Woyto-Guji**	**Gondar**
ZnS	0.0265	0.0504	0.0509
Za	0.0064	0.0636	0.0203
ZZ	-0.0200	0.0132	-0.0306
*F*s values	-17.7960^{**}	-8.0980^{*}	-12.0800^{**}
Tajima's *D*	0.19900^{ns}	0.85931^{ns}	0.46476^{ns}
Fay and Wu's *H*	-2.11640		

The neutrality estimates vary among polymorphism and divergence, and overview of polymorphism in all tests of both populations. The neutrality test for polymorphism and divergence were 0.82933 (Tajima's *D* test), -0.22458 (Fu and Li's *D** test), 0.16986 (Fu and Li's

*F** test) (Fu and Li, 1993) and -7.114 (Fu's *F*s test). Whereas the overview of polymorphism was 0.19900 (Tajima's *D* test), -1.29583 (Fu and Li's *D** test, -0.89153 (Fu and Li's *F**) and -17.796 (Fu's *F*s test) (Table 20).

Table 20 Neutrality test in exon1 across populations

	y\Locus	**3416**	**3533**	**3649**	**3770**	**3783**	**3808**	**3811**	**3963**	**3989**	**No. of loci**
Woyto-Guji	1	0.0	0.0	37.5	0.0	12.5	25.0	12.5	25.0	37.5	9
	2	0.0	*	50.0	*	16.7	33.3	16.7	33.3	50.0	7
	3	0.0	*	40.0	*	*	40.0	20.0	40.0	60.0	6
	4	0.0	*	40.0	*	*	40.0	20.0	40.0	60.0	6
	5	*	*	50.0	*	*	50.0	25.0	50.0	75.0	5
	y\Locus	**3354**	**3416**	**3649**	**3696**	**3808**	**3811**	**3927**	**3963**	**3989**	**No. of loci**
Gondar	1	0.0	0.0	25.0	0.0	25.0	25.0	0.0	25.0	50.0	9
	2	*	0.0	33.3	*	33.3	33.3	0.0	33.3	66.7	7
	3	*	0.0	40.0	*	40.0	40.0	*	40.0	80.0	6
	4	*	0.0	40.0	*	40.0	00.0	*	40.0	80.0	6
	5	*	*	50.0	*	50.0	50.0	*	50.0	100.0	5

6.4 Discussion

6.4.1 Haplotype analysis

In these days, fine-mapping studies and identification of candidate genes are conducted by haplotype association analysis of the SNPs detected in the target regions (Beaty et al., 2005). However, it is mentioned that there has been surprisingly little work done on haplotype based multivariate association analyses (Pei et al. 2009). Haplotype based analysis of the kisspeptin gene was carried out in this study. The result indicated that from the total 29 haplotypes, only 12 of them are common for both goat populations studied. Majority of them are not shared haplotypes resulted from the rare alleles and majority of the rare haplotypes were detected in Woyto-Guji goat population. This might be due to relatively more sample size used in Woyto-Guji goat population compared to Gondar goat. Large sample size in a population more likely includes more rare haplotypes (Beaty et al., 2005). In the shared haplotypes, relatively highest haplotype frequencies were obtained in Woyto-Guji goat population than Gondar except the 1st haplotype of exon2 (Table 15). However, almost all non-shared haplotypes have less than 2% haplotype frequency estimates, and all private haplotypes except the 13th haplotype showed frequencies closer to 1%. However, no haplotype was observed having a frequency of <1% in Gondar goat population. This is contrary to small sample size used for the latter goat population. In line with this, three nonsynonymous mutations showed a frequency <1% in populus nigra cinnamyl alcohol dehydrogenase (CAD4) gene and stated that it would not have been identified by studies using smaller sample size (Marroni et al., 2011).

6.4.2 Linkage disequilibrium

The study of variations in linkage disequilibrium (LD) and in haplotype frequencies within and across populations is highly relevant in the choice of “tagging” SNPs for candidate gene or

whole-genome association studies (Beaty et al., 2005). This is due to the fact that some markers will not be polymorphic in all samples and some haplotypes will be poorly represented or completely absent. In the LD measures, very low estimated *D'* and *R* were obtained in the current study. This might be because of the shortest size of target region that leads to short physical distances among segregating sites (Appendix Table I). The average distances among segregating sites are 226.19 bp for Woyto-Guji and 234.86 bp Gondar goat populations (Table 16). When markers are separated by <1 kb of DNA, *D'* values could be on average <1 (Abecasis et al., 2001) implying that an excess of LD does not appear in short physical distance (e.g. <10kb) (Pritchard and Przeworski, 2001). However, it's agreed that in analysis of whole genome or large size target region, measures of LD decays decrease as physical distance among loci increases. This is due to the fact that the recombination events will make the distribution of alleles at linked loci occur independently of each other (Lin, 2005). Another argument is, the low estimates of *D'* can be explained by H_O, which is lower than H_E (i.e.; $\Theta_\pi < \Theta_K$) (Appendix Figure D) in most of heterozygosity estimates. This could be due to presence of more rare alleles at low frequencies and there might have been recent selective sweep and population expansion (Tajima, 1989a). This idea is strengthened by high level of population migration per generation (*N*m=24) and recent and rapid *bi-modal* demographic expansion events (Table 10; Figure 7).

Similarly, the average R^2 value was very small (R^2 = 0.08334±0.15) (Table 16) suggesting the little power of coefficient of correlation to detect association among the loci (Pritchard and Przeworski, 2001) and were almost similar for both populations. However, there was slightly higher estimate $|D'|$ for Woyto-Guji goat population. This might be explained by the highest flock size farmers owned in Woyto-Guji area than Gondar (Netsanet, 2014; Alubel, 2015) which could provide better selection practice in the latter goat population studied. Population growth

leads to an excess of low-frequency variants (Tajima, 1989a); whereas, population structure tends to increase levels of LD (Pritchard and Przeworski, 2001). Similarly, in other reports, the source of variation of LD measures among populations could be selective sweeps, history of natural selection, gene conversion, mutation, genetic drift and other forces that cause gene-frequency to evolute (Abecasis et al., 2001; Slatkin, 2008). Presence of very far geographical distance among Gondar and Woyto-Guji goat populations might be one possible reason which contributed for the variation observed among themselves. Though the biological reasons have not been known yet, majority of residual variation for the distribution of LD is explained by physical distance among study populations (Abecasis et al., 2001; Pritchard and Przeworski, 2001).

Trends of linkage disequilibrium decays for all SNPs detected are illustrated at Appendix Figure E. Both $|D'|$ and R^2 suffer ceiling and floor effects, respectively. Most of the pairwise comparisons of polymorphic sites are concentrated at maximum value for $|D'|$ ($|D'|=1$) and at minimum value for R^2 ($R^2 = 0$) (Marroni et al., 2011). This could be because of one of the four possible haplotypes is not observed in the sample for the former (Mueller 2004; Marroni et al., 2011) and presence of excess rare alleles for the latter (Hedrick and Kumar, 2001). The abundance of pairwise comparisons of the ceiling and floor effects are irrespective of each other (the combination graph of $|D'|$ and R^2 is not indicated).

On the other hand, very low recombination rate (c=0.0567) was observed in both goat populations indicating the non significant contribution of genetic drift on LD accumulation rather it could be happened by selection and migration or population expansion (Kelly, 1997). The lowest neutral estimate of *ZnS* (*ZnS* = 0.0265), discussed below, strengthened this argument.

When a favourable mutant at the locus under selection sweeps detected in the population, it drags along the neutral locus and therefore the pattern of polymorphism at the neutral locus can be strongly affected by the linkage to the selected locus (Fu, 1997). However, this recombination evaluation in this specific segment of the DNA does not represent the status of the recombination in the whole genome. It is because, the rate of recombination varies across the genomic regions (Payseur and Nachman 2000; Yu et al. 2001).

Most of the SNPs detected have no/lack strong linkage disequilibrium indicating they are not likely appropriate for genetic association studies. In regions of high LD, a reduced set of haplotype tag SNPs may be selected to detect efficient associations between variations in that gene or region and a trait of interest (Beaty et al., 2005; Gong et al., 2007). Another possible reason could be variability in LD is also a function of sample size (Beaty et al., 2005). Sample size of Gondar goat population is by half smaller than that of sample size of Woyto-Guji goat population. In the shared loci, both goat populations have almost similar patterns of pairwise LD accumulation except at two loci combination in each population; this can ease to identify the minimum number of SNPs that tag the most common haplotypes, termed "tagging SNPs". The similar trends of haplotype frequencies of shared haplotypes of both goat populations (Table 15) strengthened this argument. However, according to Evans and Cardon (2005), whenever haplotype frequencies vary considerably across populations, it becomes more difficult to predict which SNPs will identify enough of the existing haplotypes in all subpopulations to ensure adequate coverage, and the chance of spurious findings due to confounding increases in tests of association. Of course, factors such as sample size become important when estimating haplotype frequencies too; but, the key determinant of differences remains underlying level of haplotype diversity and LD across populations (Beaty et al., 2005). The relatively higher estimates of LD in

the study conducted (Table 16), is because the more practice of selection than the effect of genetic drift. The later argument can be strengthened by the relative low estimate of recombination of linked loci obtained and is supported by Slatkin (2008). Genetic drift which can create small amounts of LD interacts with selection (Hill and Robertson, 1966; Slatkin, 2008) and this reduces the response to selection. The low recombination rate has also an implication that the common ancestor in the sequences was created recently which was initially linked to the selectively favoured mutation (Kelly, 1997).

In addition, it is explained that changes in population size, particularly an extreme reduction in size (a population bottleneck), can increase LD (Slatkin, 2008). With respect to this, the population/flock size of Gondar per household is lowest compared to Woyto-Guji. On the other hand, Netsanet (2014) reported that in Woyto-Guji area, the maximum goat holding in her study group was 200 per house hold; however, it was also observed up to 400 heads of goats per house hold during the field work in the current. Whereas, in Gondar, the average goat holding per household was 10.5±7.5 (Alubel, 2015). The smallest flock size per house hold at Gondar compared to Woyto-Guji invites to deduce practice of selection in the husbandry program but not expected to be routinely practiced by farmers. Besides, the selection practices by farmers are from own flock which lead the within flock differentiation to be very narrow. On the contrary, the recurrent drought that occurs in Woyto-Guji area could let to select animals which can withstand the drought challenge and frequency of alleles which fit for drought tolerance increases. Therefore, natural selection could be more prevalent than artificial selection though the average goat holding per house hold in Woyto-Guji is higher than Gondar. This might result comparatively to observe higher LD estimate in Woyto-Guji goat population compared to

Gondar goat. The haplotype frequency differences observed in both goat populations (Table 15) supports presence of selection pressure. Previous reports indicated that the increased inter population haplotype frequency differences are indications of selection pressures (Weir et al., 2005; Voight et al., 2006).

From the total SNPs detected, the association analysis among loci indicated that only four loci showed highest and significant LD accumulation (Table 18 and 19). Strong LD is expected in tightly linked loci (Pritchard and Przeworski, 2001). Variability at linked markers will be higher on chromosomes bearing that allele than other chromosomes whenever an advantageous allele is fixed (Slatkin, 2008). In section 5.3.3 above, significant ($P<0.001$) contribution of mutations at g.950T>C, g.3416G>C, g.3811G>T and g.3963T>C on multiple birth was observed in Gondar and Woyto-Guji goat populations. It is reported that strong positive selection quickly increases the frequency of an advantageous allele (Slatkin, 2008). This results linked loci to remain in strong LD with that allele, which is called genetic hitch-hiking (Maynard and Haigh, 1974). The second primarily route of selection, epistatic selection, might have its own contribution for relatively higher estimates of measures of LD by Woyto-Guji goat population than Gondar. The latter selection type leads to have the association of particular alleles at different loci that provide motivation of historical studies of LD in the study population. The insignificant recombination rate estimate strengthens this argument. It is explained that epistatic selection would have to be very strong to maintain allelic associations at the scale of megabases, in the face of substantial recombination (http://goo.gl/jk4E7v).

In general, significant population variation in the candidate gene was also observed, particularly in the shared haplotypes, in haplotype diversity and in differences in LD implying that some of the SNPs and haplotypes are ‘‘useful’’ for association studies (Beaty et al., 2005). Woyto-Guji goat populations had fairly similar haplotype diversity but slightly higher levels of measures of LD than did Gondar goat population. The rate of recombination (R) is also relatively lower in Woyto-Guji goat population (R=0.0505) than Gondar goat population (R=0.0765) and strengthened the idea that the recombination rate decreases as LD accumulation increases.

6.4.3 Neutrality test

The randomly evolving mutations are called "neutral", while mutations under selection are "non-neutral" (Tajima, 1989b). In the current study, the estimates in polymorphism and divergence were higher than estimates of the polymorphism overview in all neutrality tests. Negative and highly significant *F*s values were obtained in both goat populations studied. According to Fu (Fu, 1997), *F*s test is especially sensitive to population demographic expansion, which generally leads to have large negative *F*s values. However, Gondar goat showed higher significant negative value of *F*s than Woyto-Guji. This could be because of high demographic expansion towards Gondar area. As result, a genome wide SNP CHIP array study indicated in the above section (chapter 4) revealed that Gondar goat has more than three major genetic backgrounds, which could let this goat population to have higher negative values whereas Woyto-Guji has only two genetic backgrounds. However, all the estimates, except the Tajima’s *D* test, were negative values in the later group (Table 20). Large negative value which indicates a one-sided test, for in instance in F_S**,** is an indicator against the neutrality of mutations implying an excess of number of rare alleles and a reduction of the number of common alleles (Fu, 1997). Fu (*ibid)* proved that

in showing the effect of population growth on neutrality test, the *F*s test is the most powerful one; in fact, it is often more than twice as powerful as any other test examined. On the other hand, Watterson's *W* test is the least powerful test. In between are Tajima's test *T*, Fu and Li's tests *D** and *F** and the new test *F'* (-1, 1).

Moreover, negative values of Tajima's *D* in particular, which is non-significant positive value in this study, shows presence of negative selection, population growth and genetic hitchhiking (Tajima, 1989b). Similarly negative value was also observed in Fay and Wu' *H* test (*H*= -2.11640) suggesting genetic hitchhiking (Fay and Wu, 2000). This goes in line with Tajima's *D* estimates of mtDNA of Gondar (*D* = 0.10) and Woyto-Guji (*D* = -0.22) goat populations (Table 5). The coexistence of negative values for both *D* and *H* could be related to demographic history of the population (Marroni et al., 2011) that could be explained by a bottleneck event (Heuertz et al., 2006). On the other hand, as described in the methods section, both goat populations are geographically isolated and individuals from different geographic areas could cause allelic frequencies to be skewed toward rare alleles resulting in the detection of negative Tajima's *D* values due to population structure (Städler et al. 2009). However, the positive Tajima's *D* value detected in the KISS1 gene could be due to high estimation of level of population migration per generation (*N*m =18.17) and this is strengthened by lowest pairwise F_{ST} distance (F_{ST}= 0.0267) between the two goat populations (Table 10).

On the other hand, the lowest *ZnS*, which is a measure of allele frequency equivalency (ranges 0 to 1) across polymorphic sites in the absence of recombination, obtained implies acceptance of the neutral model and encourage to use it as a test (Kelly, 1997). According to Kelly (1997), the values *ZnS* measures declines as asymmetry among loci increases; when natural selection acts on

a polymorphism that is closely linked to neutral sites, allele frequency asymmetries may be reduced. For this reason, lower expected values of *ZnS* may represent a molecular signature of natural selection. The considerable codon bias index (CBI=0.301) obtained could also strengthen the effect of natural selection.

In General, the overall neutrality evaluation of the KISS1 gene shows influence of selection on the goat population studied. However, it is mentioned that neutrality tests are quite sensitive to variations in sample size (Marroni et al., 2011). The reason is small sample sizes lead to a relatively large variance of π and *D* (Lohse and Kelleher, 2009). However, how small sample size is small and how variable sample size differences among study populations need to be defined. For instance, in the current study the average estimate of π (π= 0.00275) and its variance were very low and neutrality test was detected in contrast to highest variation of sample size between the two goat populations included in the study.

CHAPTER VII

7 Conclusions and recommendations

7.1 Conclusions

Given Ethiopia served as the main gate of livestock entry to Africa, very low variation was observed among the Ethiopian goat populations studied. Both mtDNA and SNP CHIPs panel analyses revealed high level of genetic diversity but weak genetic structure among Ethiopian goat populations. From the six haplogroups globally indentified, only haplogroup A and G were detected in Ethiopia. These two haplogroups are compatible with the haplogroups observed in Egypt and Saudi Arabia implying that the goat populations found in Ethiopia are descendants of Egypt's and Saudi Arabian's goat populations. This observation is strengthened by output of the autosomal markers regression analysis of geographic distance and heterozygosity. The latter result suggests that goats arrived in Ethiopia following two routes of introduction, via the north and east direction, and genetic diversity gets decreasing as one moves far away from the entry points.

Based on the mtDNA analysis, there were two recent and rapid major demographic expansions held in Ethiopian goat populations. These could have influenced the current genetic structure of the goats together with other potential contributing factors. In addition, the Ethiopian goat populations studied have only six genetic backgrounds showing the goat populations have passed through continuous intermixing. These all could have resulted because of population migration events following the vast physical movement of human populations in the country due to various reasons. This argument could be strengthened by the haplotypes, which were grouped in both haplogroups, detected in all the goat populations studied.

The current study does not support the former classifications of the indigenous goat populations that based production systems, agro-ecologies, goat families and SSR markers; however, the admixture and phylogenetic network analyses suggest that the 14 Ethiopian goat populations can be re-grouped in to seven goat types.

Kaffa and Abergelle goat populations have relatively maintained pure genetic background compared to other Ethiopian indigenous goat populations. Kaffa goat is reared in highly tsetse infested area, and Abergelle goat is also known with its drought tolerance nature. Hence, their genetic potential might have helped both goat populations to adapt the local environment and maintained pure genetic background. On the other hand, the phylogenetic network indicated that there are only few ancestral populations that separated Chinese goat from Ethiopian goat populations despite the large genetic distance observed. Instead, high number of ancestral populations obtained among Ethiopian goat populations implies the current indigenous goat populations are formation of continuous introgressions or interbreeding.

In the kisspeptin gene analysis, haplotype frequencies, together with patterns of pairwise LD, were used to assess genetic variation in Woyto-Guji and Gondar goat populations. These goat populations showed fairly similar haplotype frequencies and heterozygosities. However, relatively higher LD decays, caused by natural selection as confirmed by neutrality tests, were observed in Woyto-Guji goat population than Gondar. In general, some of the polymorphic loci detected in the target regions showed comparatively highly significant linkages among themselves. In line with this, four polymorphic sites have significant contribution to litter size confirming relevance of the KISS1 gene for fecundity trait.

7.2 Recommendations

Based the outputs of the mtDNA, high density SNP CHIPs and the KISS1 gene analyses, the following recommendations are suggested:

- The findings reported on the origin as well as demographic dynamics of Ethiopian goats were limited to maternal origins; assessing paternal origins (Y-chromosome) and ancient DNA may provide further insights about origin and history of the goat populations.
- Genetic background of Kaffa and Abergelle goat populations appeared relatively pure and requires further work to confirm and validate whether both goat populations are potential candidates for trypanotolerant and drought tolerant populations, respectively.
- In this study, highest level of within populations variation which could be resulted from highest level of gene flow among the goat population was observed. Therefore, strong animal regulatory policy and strategy are imperative parallel to designing effective-informed breeding schemes.
- In the current study, only six genetic backgrounds are revealed using 50K SNP chip; more dense SNP chip panel may differentiate the indigenous goat populations better than the current findings.
- The association analysis of kisspeptin gene indicated that there is 18%-31% increment of litter size because of mutations observed in few of the loci detected. Hence, the KISS1 gene can be suggested for marker assisted selection breeding interventions. However, sequencing the whole length of KISS1 gene and testing both the haplotype and measures of LD decays with more sample size may help to arrive at firm conclusion.

8 FUTURE WORKS

- ✓ Further investigation on genetic qualities of Abergelle, Kaffa, Arsi-Bale goat populations targeting trypanotolerance, drought tolerance and hair fiber contributing genes respectively are duly suggested.
- ✓ Genome wide assessment of signature of selection and linkage disequilibrium on all the indigenous goat populations are paramount.
- ✓ Genome-wide survey of *Capra Walie* (*Walia ibex*) is duly required to see the genetic background and status of the *ibex* in general and its influence on Ethiopian indigenous goat populations. This may also help to design learned and sustained conservation strategy for the *ibex* itself.

9 REFERENCE

Abecasis GR, Noguchi E, Heinzmann A, Traherne JA, Bhattacharya S, Leaves NI, Anderson GG, Zhang Y, Lench NJ, Carey A, Cardon LR, Moffatt MF and Cookson WO, 2001. Extent and distribution of linkage disequilibrium in three genomic regions. Am J. Hum Genet 68:191–197.

Aggarwal RAK, Dixit SP, Verma NK, Ahlawat SPS, Kumar Y, Kumar S, Chander R and Singh KP, 2007. Population genetics analysis of Mehsana goat based on microsatellite markers. Current Science 92:1133–1137.

Agha SH, Pilla F, Galal S, Shaat I, D'Andrea M, Reale S, Abdelsalam AZA and Li MH, 2008. Genetic diversity in Egyptian and Italian goat breeds measured with microsatellite polymorphism. J. Anim. Breed. Genet. 125:194–200.

Akis I, Oztabak K, Mengi A and Un C, 2014. Mitochondrial DNA diversity of Anatolian indigenous domestic goats. J. Anim. Breed. Genet. 131: 487–495.

Akis I, Onar V, Toker N, Belli O, Pazvant G and Oztabak K, 2016. Ancient DNA analysis of Anatolian goat remains excavated from a Urartian Castle in eastern Turkey. Int. J. Osteoarchaeol. 26: 246–254

Alexander DH, November J, Lange K, 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res.19:1655–64.

Ali BA, 2003. Genetics similarity among four breeds of sheep in Egypt detected by random amplified polymorphic DNA markers. African Journal of Biotechnology Vol. 2 (7):194–197.

Aljumaah RS, Musthafa MM, Al-Shaikh MA, Badri OM and Hussein MF, 2012. Genetic diversity of Ardi goat based on microsatellite analysis. African Journal of Biotechnology Vol. 11(100):16539-16545.

Alubel Alemu, 2015. On-farm phenotypic characterizations and performance evaluation of Abergelle and Central highland goat breeds as an input for designing community-based breeding program. MSc theis. Haramaya University Haramaya, Ethiopia.

Álvarez I, Gutiérrez JP, Royo LJ, Fernández I, Gómez E, Arranz JJ and Goyache F, 2004. Testing the usefulness of the molecular coancestry information to assess genetic relationships in livestock using a set of Spanish sheep breeds. J. Anim. Sci. 2005. 83:737–744.

Amills M, Capote J, Toma`s A, Kelly L, Obexer-Ruff G, Angiolillo A and Sanchez A, 2004. Strong phylogeographic relationships among three goat breeds from the Canary Islands. Journal of Dairy Research 71:257–262.

Amills M, Ramı´rez O, Toma`s A, Badaoui B, Marmi J, Acosta J, Sa`nchez A and Capote J, 2008. Mitochondrial DNA diversity and origins of South and Central American goats. Animal Genetics, 40:315–322.

An XP, Han D, Hou JX, Li G, Wang JG, Yang MM, Song YX, Zhou GQ, Wang YN and Ling L, 2009. GnRHR gene polymorphisms and their effects on reproductive performance in Chinese goats. Small Rumin Res 85:130-134.

An XP, Ma T, Hou J, Fang F, Han P, Yan Y, Zhao H, Song Y, Wang J and Cao B, 2013. Association analysis between variants in KISS1gene and litter size in goats. BMC Genetics 14:63.

Araújo AM, Guimarães SEF, Machado TMM, Lopes PS, Pereira CS, Francisco Luiz Ribeiro da Silva FLR, Rodrigues MT, Columbiano VS and Fonseca CG, 2006. Genetic diversity between herds of Alpine and Saanen dairy goats and the naturalized Brazilian Moxotó breed. J. Genetics and Molecular Biology, 29(1): 67-74.

Armstrong DG, Gong JG and Webb R, 2003. Interactions between nutrition and ovarian activity in cattle: physiological, cellular and molecular mechanisms. Reproduction 61:403-414.

Azor PJ, Monteagudo LV, Luque M, Tejedor MT, Rodero E, Sierra I, Herrera M, Rodero A and Arruga MV, 2005. Phylogenetic relationships among Spanish goats breed. Animal Genetics, 36:423–425.

Bai WL, Wang J and Yin RY, 2005. Study on genetic polymorphism of HaeIII site of GH gene in Chengdu-Ma goat and Boer goat. Heilongjiang Anim Sci and Vet Med 8:13-14.

Balding DJ, 2006. A tutorial on statistical methods for population association studies. Nature review, genetics Vol.7.

Barker JSF, 1994. A global protocol for determining genetic distances among domestic livestock breeds. In: V World Congress on Genetics Applied to Livestock Production, Gelph: sd, v. 21:501-508.

Barker JSF, Tan SG, Moore SS, Mukherjee TK, Matheson JL and Selvaraj OS, 2001. Genetic variation within and relationships among populations of Asian goats (Capra hircus). Journal of Animal Breeding and Genetics 118:213–23.

Barluenga M, Stolting KN, Salzburger W, Muschick M, Meyer A: Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature 439:719-723.

Beaty TH, Fallin MD, Hetmanski JB, McIntosh I, Chong SS, Ingersoll R, Sheng X, Chakraborty R and Scott AF, 2005. Haplotype diversity in 11 candidate genes across four populations genetics 171: 259–267.

Behnke R, 2010. The Contribution of Livestock to the Economies of IGAD Member States: study findings, application of the methodology in Ethiopia and recommendations for further work. IGAD LPI Working Paper No. 02 -10.

Behzadi S, Miraei-Ashtiani SR, Sadeghi M, Zamani P and Abdoli R, 2015. Association of IGF-I Gene Polymorphisms with Carcass Traits in Iranian Mehraban Sheep Using SSCP Analysis. Iranian Journal of Applied Animal Science. 5(1): 121-126.

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

Berhanu Gebremedhin, Hoekstra D and Samson Jemaneh, 2007. Heading towards commercialization? The case of live animal marketing in Ethiopia. Improving Productivity and Market Success (IPMS) of Ethiopian Farmers Project Working Paper 5. ILRI (International LivestockResearch Institute), Nairobi, Kenya. 73 pp.

Bernardi G, Olofsson B, Filipski J, Zerial M, Salinas J, Cuny G, Meunier-Rotival M and Rodier F, 1985. Science 228, 953-958.

Boron WF and Boulpaep EL, 2011. Medical Physiology. 2nd edition. https: //www.inkling.com/read/medical-physiology-boron-boulpaep-2nd/chapter-55/ch55-section-2

Botstein D, White RL, Skolnick M, Davis RW, 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics, 32:314–331.

Brehem A, Khadem M, Jesus J, Andrade P and Vicente L, 2001. Lack of congruence between morphometric evolution and genetic differentiation suggests a recent dispersal and local habitat adaptation of the Madeiran lizard Lacerta dugesii. Genetics Selection Evolution 33:671–85.

Brito LF, Jafarikia M, Grossi DA, Kijas JW, Porto-Neto LR, Ventura RV, Salgorzaei M and Schenkel FS, 2015. Characterization of linkage disequilibrium, consistency of gametic phase and admixture in Australian and Canadian goats. BMC Genetics 16:67. DOI 10.1186/s12863-015-0220-1.

Brookes AJ, 1999. The essence of SNPs. Gene 234:177–86.

Brown MS, Dana SE, and Goldstein JL, 1973. Regulation of 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase Activity in Human Fibroblasts by Lipoproteins (cholesterol biosynthesis/enzyme suppression and stimulation/low-density and high-density lipoproteins). Vol. 70(7):2162-2166

Brown RP and Thorpe RS, 1991. Description of within-island microgeographic variation in body dimensions and scalation of the skink Chalcides sexlineatus, which testing of causal hypotheses. Bio.J. of the Linnean Society 44:47-64.

Browning SR and Weir BS, 2010. Population Structure with Localized Haplotype Clusters. Genetics 185:1337–1344.

Bruford MW, Bradley DG and Luikart G, 2003. DNA markers reveal the complexity of livestock domestication. Nature Reviews Genetics, 4: 900–910.

Callen DF, Thompson AD, Shen Y, Phillips HA, Richards RI, Mully JC, Sutherland GR, 1993. Incidence and origin of "null" alleles in the (AC)n Microsatellite markers. Am. J. Hum. Genet., 52:922-927.

Cañón J, Garcı´a D, Garcı´a-Atance MA, Obexer-Ruff G, Lenstra JA, Ajmone-Marsan P, Dunner S and The ECONOGENE Consortium, 2006. Geographical partitioning of goat diversity in Europe and the Middle East. J.Animal Genetics, 37:327–334.

Cao GL, Chu MX, Fang L, Di R, Feng T, Li N, 2010. Analysis on DNA sequence of KISS1 gene and its association with litter size in goats. Mol Biol Rep 37:3921–3929.

Cao GL, Chu MX, Fang L, Feng T, Di R and Li N, 2011. Analysis on DNA sequence of GPR54 gene and its association with litter size in goats. Mol Biol Rep 38:3839–3848.

Capote J, Fresno MR and Álvarez S, 1999. The Canarian goat breeds: characterization and present situation. Ovis 62:11–22.

Capote J, Tejera A, Amills M, Argüello A, Fresno M, López JL, 2004. Influencia históricay actual de los genotipos canarios en la población caprina americana AGRI- 35: 49–60.

Cappuccio I, Pariset L, Ajmone-Marsan P, Dunner S, Cortes O, Erhardt G, G. Lühken G, K Gutscher K, Joost S, Nijman IJ, Lenstra JA, England PR, Zundel S, Obexer-Ruff G, Beja-Pereira A, Valentini A and The econogene consortium, 2006. Allele frequencies and diversity parameters of 27 single nucleotide polymorphisms within and across goat breeds. Molecular Ecology Notes 6, 992–997.

Cardellino RA and Boyazoglu J, 2009. Research opportunities in the field of animal genetic resources, Livestock Science, 120:166–173.

Carillier C, Larroque H, Palhiere I, Clement V, Rupp R, Robert-Granie C, 2013. A first step toward genomic selection in the multi-breed French dairy goat population. J Dairy Sci. 96(11):7294–305. doi:10.3168/jds.

Chen SY, Su YH, Wu SF, Sha T, Zhang YP, 2005. Mitochondrial diversity and phylogeographic structure of Chinese domestic goats. Mol Phylogenet Evol 37(3): 804–814.

Chenyambuga SW, Hanotte O, Hirbo J, Watts PC, Kemp SJ, Kifaro GC, Gwakisa PS, Petersen PH and Rege JEO, 2004. Genetic Characterization of Indigenous Goats of Sub-saharan Africa Using Microsatellite DNA Markers. Asian Australasian Journal of Animal Sciences). 17(4).

Chiejina SN, Behnke JM and Fakae BB, 2015. Haemonchotolerance in West African Dwarf goats: contribution to sustainable, anthelmintics-free helminth control in traditionally managed Nigerian dwarf goats.

Christina MM, 2013. The Role of Kisspeptin and KNDy Cells in the Reproductive Neuroendocrine System. Electronic Thesis and Dissertation Repository. The University of Western Ontario, Canada.

Christopher AP, 2011. Variation in Y chromosome, mitochondrial DNA and labels of identity in Ethiopia. PhD dissertation. The center for Genetic Anthropology, department of Genetics Evolution and Environment, University College, London.

Chu MX, Jiao CL, He YQ, Wang JY, Liu ZH and Chen GH, 2007. Association between PCR-SSCP of bone morphogenetic protein 15 gene and prolificacy in Jining Grey goats. Anim Biotechnol 18:263-274.

Chu M, Xiao C, Feng T, Fu Y, Cao G, Fang L, Di R, Tang Q, Huang D, Ma Y, Li K and Li N, 2012. Polymorphisms of KiSS-1 and GPR54 genes and their relationships with litter size in sheep. Mol Biol Rep 39:3291-3297.

Clutton-Brock J, 1999. Domesticated animals from early times. The Natural History Museum: Cambridge University Press. 238 p.

Clutton-Brock J, 2000. Cattle, sheep, and goats south of the Sahara: an archaezoological perspective. In: Blench RM, MacDonald KC, editors. The origins and development of African livestock: archaeology, genetics, linguistics and ethnography. London: UCL Press. p. 30–37.

Colli L, Lancioni H, Cardinali I, Olivieri A, Capodiferro MR, Pellecchia M, RzepusM, Zamani W, Naderi S, Gandini F, Vahidi SMF, AghaS, Randi E, Battaglia V, Sardina MT, Portolano B, Rezaei HR, Lymberakis P, Boyer F, Coissac E, Pompanon F, Taberlet P, Marsan PA and

Achilli A, 2015. Whole mitochondrial genomes unveil the impact of domestication on goat matrilineal variability. BMC Genomics 16:1115

Cymbron T, Freeman AR, Isabel Malheiro M, Vigne JD and Bradley DG, 2005. Microsatellite diversity suggests different histories for Mediterranean and northern European cattle populations. Proceedings of the Royal Society of London Series B, Biological Sciences 272:1837–43.

Dadi H, Tibbo M, Takahashi Y, Nomura K, Hanada H and Amano T, 2008. Microsatellite analysis reveals high genetic diversity but low genetic structure in Ethiopian indigenous cattle populations. International Society for Animal Genetics, Animal Genetics, 39: 425–431

Dadi H, Kim JJ, Yoon D, Kim KS, 2012. Evaluation of single nucleotide polymorphisms (SNPs) genotyped by the Illumina Bovine SNP50K in cattle focusing on Hanwoo breed. Asian-Aust J Anim Sci 25:28-32.

Davila SG, Gil MG, Resino-Talavan and Campo JL, 2009. Evaluation of diversity between different Spanish chicken breeds, a tester line and White Leghorn population based on microsatellite markers. Poult.Sci., 88: 2518-2525.

Deldar-Tajangookeh H, Shahneh AZ, Zamiri MJ, Daliri M, Kohram H and Nejati-Javaremi A, 2009. Study of BMP-15 gene polymorphism in Iranian goats. African Journal of Biotechnology Vol. 8 (13):2929-2932.

DeRoux NE, Genin J, Carel F, Matsuda J, Chaussain and Milgrom E, 2003. Hypogonadotropic hypogonadism due to loss of function of the KiSS1-derived peptide receptor GPR54. Proceedings of the National Academy of Sciences of the United States of America. Proc. Natl. Acad. Sci USA. 100:10972-10976.

Devendra C, and Nozawa K, 1976. Goat in South East Asia – their status and production. Zeitschrift fu¨r Tierzu¨chtung und Zu¨chtungsbiologie: Organ der Reichsarbeitsgemeinschaft Tierzucht im Forschungsdienst 93: 101–20.

Devendra C and McLeroy GB, 1982. Goat and Sheep Production in the Tropics. Intermediate Tropical Agriculture Series, 271 pp. Longman, London.

Dhillo WS, Chaudhri OB, Thompson EL, Murphy KG, Patterson M, Ramachandran R, Nijher GK, Amber V, Kokkinos A, Donaldson M, et al., 2007. Kisspeptin-54 stimulates gonadotropin release most potently during the preovulatory phase of the menstrual cycle in women. J Clin Endocr Metab 92:3958–3966.

Di R, Farhad Vahidi SM, Ma YH, He XH, Zhao QJ, Han JL, Guan WJ, Chu MX, Sun W and Pu YP, 2010.Microsatellite analysis revealed genetic diversity and population structure among Chinese cashmere goats. J. Animal Genetics. doi: 10.1111/j.1365-2052.2010.02072.x.

Diegisser T, Seitz A, Johannesen J, 2006.. Phylogeographic patterns of host-race evolution in Tephritis conura (Diptera: Tephritidae). Mol Ecol 15:681-694.

Dixit SP, Verma NK, Aggarwal RAK, Kumar S, Chander R, Vyas MK and Singh KP, 2009. Genetic Structure and Differentiation of Three Indian Goat Breeds. Asian-Aust. J. Anim. Sci. Vol. 22(9): 1234 – 1240.

Dong Y, Min Xie M, Jiang Y, Xiao N, Du X, Zhang W, Tosser-Klopp G, Wang J, Yang S, Liang J, Chen W, Chen J, Zeng P, Hou Y, Bian C, Pan S, Li Y, Liu X, Wang W, Servin B, Sayre B, Zhu B, Sweeney D, Moore R, Nie W, Shen Y, Zhao R, Zhang G, Li J, Faraut T, Womack J, Zhang Y, Kijas J, Cockett N, Xu X, Zhao S, Wang J and Wang W, 2013. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (Capra hircus). Nature biotechnology vol.31:2

Dorji T, Kameda M, Lin BZ, Takahashi A, Oyama K, Sasazaki S, Yamamoto Y, Kawamoto Y and Mannen H, 2010. Mitochondrial DNA variation and genetic construction of indigenous cattle population in Bhutan. The Journal of Animal Genetics 38:77–81.

Dorji N, Duangjinda M and Phasuk Y, 2012. Genetic characterization of Bhutanese native chickens based on an analysis of Red Junglefowl (Gallus gallus gallus and Gallus gallus spadecieus), domestic Southeast Asian and commercial chicken lines (Gallus gallus domesticus). Genetics and Molecular Biology 35(3): 603-609.

Doro MG, Piras D, Leoni GG, Casu G, Vaccargiu S, Parracciani D, Naitana S, Pirastu M, Novelletto A, 2014. Phylogeny and Patterns of Diversity of Goat mtDNA Haplogroup A Revealed by Resequencing Complete Mitogenomes. PLoS ONE 9(4): e95969. doi:10.1371/journal.pone.0095969.

Dungan HM, Clifton DK, Steiner RA, 2006. "Minireview: kisspeptin neurons as central processors in the regulation of gonadotropin-releasing hormone secretion". Endocrinology 147 (3): 1154–8.

Edea Z, Dadi H, Kim SW, Dessie T, Kim KS, 2012. Comparison of SNP Variation and Distribution in Indigenous Ethiopian and Korean Cattle (Hanwoo) Populations. Genomics Inform 10(3):200-205.

Edea Z, Dadi H, Kim SW, Dessie T, Lee T, Kim K, Kim JJ and Kim KS, 2013. Genetic diversity, population structure and relationships in indigenous cattle populations of Ethiopia and Korean Hanwoo breeds using SNP markers. Frontier Genetics, Volume 4:35.

Els JF, Kotze A and Swart H, 2004. Genetic diversity of indigenous goats in Namibia using microsatellite markers: preliminary results. South African Journal of Animal Science 34.

Epstein H, 1971. The Origin of Domestic Animals of Africa. Vol. II. Africana Publishing Corporation, New York. 957 pp.

Evans DM and Cardon LR, 2005. A comparison of linkage disequilibrium patterns and estimated population recombination rates acrossmultiple populations. Am. J. Hum. Genet. 76: 681–687.

Excoffier L, Schneider S, 1999. Why hunter-gatherer populations do not show signs of pleistocene demographic expansions? Proc Natl Acad Sci USA 96:10597-10602.

Excoffier L, Laval G and Schneider S, 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. Online 1:47–50.

Excoffier L and Lischer HEL, 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources. 10: 564-567.

FAO (Food and Agriculture Organization of the United Nations), 2004. Livestock sector brief: Ethiopia. Livestock information, sector analysis and policy branch (AGAL), FAO, Rome, Italy.

FAO (Food and Agricultural Organization), 2007a. The State of the World's Animal Genetic Resources for Food and Agriculture. FAO, Rome.

FAO (Food and Agricultural Organization), 2007b. Organization for economic cooperation and development food and agriculture organization of the United nations.

FAO (Food and Agricultural Organization), 2011. Guidelines on Molecular Genetic Characterization of Animal Genetic Resources. FAO, Rome.

FAO (Food and Agricultural Organization), 2012.Contribution of agricultural growth to reduction of poverty, hunger and malnutrition: The role of agricultural growth in economic

growth, and poverty and hunger reduction. http://www.fao.org/docrep/016/i3027e/i3027e04.pdf.

FARM-Africa, 1996. Goat Types of Ethiopia and Eritrea. Physical Description and Management Systems. FARM-Africa, London, UK and ILRI (International Livestock Research Institute), Nairobi, Kenya.

Fatima S, Bhong CD, Rank DN and Joshi CG, 2008. Genetic variability and bottleneck studies in Zalawadi Gohilwadi and Surti goat breeds of Gujarat (India) using microsatellites. Small Ruminant Research 77: 58-64.

Fay JC and Wu CI, 2000. Hitchhiking under positive Darwinian selection. Genetics 155:1405–1413.

Fedlu Hassen, Endashaw Bekele, Workneh Ayalew and Tadelle Dessie, 2007. Genetic variability of five indigenous Ethiopian cattle breeds using RAPD markers. African Journal of Biotechnology Vol. 6 (19), pp. 2274-2279. http://www.academicjournals.org/AJB

Feng T, Zhao YZ, Chu MX, Zhang YJ, Fang L, Di R, Cao GL, Li N, 2009. Association between sexual precocity and alleles of KISS1 and GPR54 genes in goats. Anim. Biotechnoly 20:172–176.

Fernández H, Hugues S, Vigne JD, Helmer D, Hodgins G, et al, 2006. Divergent mtDNA lineages of goats in an Early Neolithic site, far from the initial domestication areas. Proc Natl Acad Sci USA 103:15375–15379.

Fields PD, McCauley DE, McAssey EV and Taylor DR, 2014. Patterns of cyto-nuclear linkage disequilibrium in Silene latifolia: genomic heterogeneity and temporal stability. Heredity 112:99–104.

Fox JM and Erill I, 2010.Relative Codon Adaptation: A Generic Codon Bias Index for Prediction of Gene Expression. DNA Research 17:185–196.

Frankham R, Briscoe DA and Ballou JD, 2002. Introduction to conservation genetics. Cambridge University Press, New York, New York, USA.

Freeman AR, Meghan CM, Machugh DE, Loftus RT, Achukwi MD, et al, 2004. Admixture and diversity in West African cattle populations. Mol Ecol 13: 3477–3487.

Fresno M, Capote JF, Camacho ME, Darmanin N and Delgado JV, 1992. The Canary Islands breeds: Past, present and future. Arch. Zootec. 41 (extra): 513-518. 1992.

Fryxell KJ, 2008. The evolution of codon bias in the cyanobacteria. Nova Hedwigia Bieheft 133: 285–304.

Fu YX and Li WH, 1993. Statistical Tests of Neutrality of Mutations. Genetics 133: 693-709

Fu Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925.

Funes S, Hedrick J, Vassileva G, Markowitz L, Abbondanzo S, Golovko A, Yang S, Monsma F and Gustafson E, 2003. The KiSS-1 receptor GPR54 is essential for the development of the murine reproductive system. Biochem. Biophys. Res. Commun. 312:1357-1363.

Garrine CMLP, Kotze A, Els H and Grobler JP, 2010. Genetic characterization of the indigenous Landim and Pafuri goat breeds from Mozambique. African Journal of Agricultural Research 5(22): 3130-3137.

Getinet Ameha, Hegde BP, Bekele Taffese, Enyew Negusse and Workneh Ayalew, 2005. Phenotypic Characterization of Goat Types in Northwestern Ethiopia. Eth. J. Anim. Prod. 5(1): 13-31.

Getinet M and Adebabay K, 2015. A review on indigenous cattle genetic resources in Ethiopia: adaptation, status and survival. Online J. Anim. Feed Res., 5(5): 125-137.

Getinet M, Gizaw S, Dessie T, Mwai O, DjikengA and Tesfaye K, 2016. A Review on Current Knowledge of Genetic Diversity of Domestic Goats (Capra hircus) Identified by Microsatellite Loci: How those Efforts are Strong to Support the Breeding Programs? J. Life Sci. Biomed. 6(1): 22.

Getinet Mekuriaw, Kassahun Tesfaye, Wenguang Zhang, Tadelle Dessie, Appolinairre Djikeng and Okeyo Mwai, nd. Documenting Radiocarbon Evidences, Y-chromosome, Mitochondria DNA and Autosomal Markers on Origin of Domestication and Routes of Goat Global Divergence: a Review. Impress.

Gifford-Gonzalez D. and Hanotte O, 2011. Domesticating animals in Africa: Implications of genetic and archaeological findings. J World Prehist 24:1–23

Girma Amare, 1988. Drought and Famine in Ethiopia and the Effort to Overcome It, 3 Notre Dame J.L. Ethics and Pub. Pol'y 71.

Gizaw S, Van Arendonk JAM, Komen H, Windig JJ, and Hanotte O, 2007. Population structure, genetic variation and morphological diversity in indigenous sheep of Ethiopia. Animal Genetics, 38:621–628.

Gizaw S, 2009. Goat breeds of Ethiopia: A guide for identification and utilization. Ethiopia sheep and goat productivity improvement program. Techinical bulletin, Number 27.

Glowatzki-Mullis ML, Muntwyler J, Bäumle E, Gaillard C, 2008. Genetic diversity measures of Swiss goat breeds as decision-making support for conservation policy. Small Rum. Res. 74(1-3):202-211.

Goddard KAB, Hopkins PJ, Hall JM and Witte JS, 2000. Linkage Disequilibrium and Allele-Frequency Distributions for 114 Single-Nucleotide Polymorphisms in Five populations. Am. J. Hum. Genet. 66:216–234

Gong Y, Beitelshees AL, Wessel J, Langaee TY, Schork NJ and Johnsona JA, 2007. Single nucleotide polymorphism discovery and haplotype analysis of Ca2+-dependent K+ channel beta-1 subunit Pharmacogenet Genomics. 17(4): 267–275.

Gottsch ML, Cunningham MJ, Smith JT, Popa SM, Acohido BV, Crowley WF, Seminara S, Clifton DK, Steiner RA, 2004. A role for kisspeptins in the regulation of gonadotropin secretion in the mouse. Endocrinology 145:4073–4077.

Gottsch ML, Clifton DK, Steiner RA, 2009a. "From KISS1 to kisspeptins: An historical perspective and suggested nomenclature". Peptides 30 (1): 4–9.

Gottsch ML, Navarro VM, Zhao Z, Glidewell-Kenney C, Weiss J, Jameson JL, Clifton DK, Levine JE, Steiner RA, 2009b. Regulation of KISS1 and dynorphin gene expression in the murine brain by classical and non-classical estrogen receptor pathways. Journal of Neuroscience. 29(29): 9390-9395.

Gour DS, Malik G, Ahlawat SPS, Pandey AK, Sharma R, Gupta N, Gupta SC, Bisen P and Kumar D, 2006. Analysis of genetic structure of Jamunapari goats by microsatellite markers. Small ruminant research. 66:140-149.

Grossen C, Keller L, Biebach I, The International Goat Genome Consortium, Croll D, 2014. Introgression from Domestic Goat Generated Variation at the Major Histocompatibility Complex of Alpine Ibex. PLoS Genet 10(6): e1004438. doi:10.1371/journal.pgen.1004438.

Grum Gebreyesus, Aynalem Haile and Tadelle Dessie, 2013. Breeding scheme based on community-based participatory analysis of local breeding practices, objectives and constraints

for goats around Dire Dawa, Ethiopia. Livestock Research for Rural Development 23 (3): http://www.lrrd.org/lrrd25/3/cont2503.htm.

Gübitz T, Thorpe RS and Malhotra A, 2000. Phylogeographic and natural selection in the Tenerife gecko Tarentola delalandii: testing historical and adaptive hypotheses. Molecular Ecology 9, 1213–21.

Guilaine J, 2003. De la Vague á la Tombe. La Conquéte Néolithique de la Méditerranée Le Seuil, Paris. ISBN: 2-02-055388-0; 375P.

Gupta N, Ahlawat SPS, Kumar D, Gupta SC, Pandey A and Malik G, 2007. Single nucleotide polymorphism in growth hormone gene exon-4 and exon-5 using PCR-SSCP in Black Bengal goats - A prolific meat breed of India. Meat Sci 76:658-665.

Habitamu Ayalew, 2014 (2006 E.C.). Country and politics (Amharic version:). First edition.

Hulunm Gatew, 2014. On-farm phenotypic characterization and performance evaluation of Bati, Borena and Short eared Somali goat populations of Ethiopia. MSc theis. Haramaya University Haramaya, Ethiopia.Addis Ababa Ethiopia

Halima Hassen Mogesse, 2007. Phenotypic and genetic characterization of indigenous chicken populations in Northwest Ethiopia. PhD thesis submitted to the Faculty of Natural and Agricultural Sciences Department of Animal, Wildlife and Grassland Sciences University of the Free State, Bloemfontein, South Africa.

Hall SJG and Bradley DG, 1995. Conserving livestock breed biodiversity. Trends. Ecol. Evol. 10: 267-270.

Han L, Yu HX, Cai DW, Shi HL, Zhu H, Zhou H, 2010. Mitochondrial DNA analysis provides new insights into the origin of the Chinese domestic goat. J. Small Ruminant Research 90: 41-46.

Hanotte O, Bradley DG, Ochieng JW, Verjee Y, Hill EW and Rege JEO, 2002. African pastoralism: genetic imprints of origins and migrations. Science 296, 336–9.

Harpending H, Sherry S, Rogers A, Stoneking M, 1993. The genetic structure of ancient human populations. Current Anthropology 34: 483–496.

Harpending HC, 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum Biol 66:591-600.

Harris DR, 1962. The distribution and ancestry of the domestic goat. Proc. Linn. Soc. London 173:79.

Hasan AU, Suguri S, Fujimoto C, Itaki RL, Harada M, Kawabata M, Bugoro H and Albino B, 2008. Genetic diversity in two sibling species of the Anopheles punctulatus group of mosquitoes on Guadalcanal in the Solomon Islands. BMC Evolutionary Biology 8:318. doi:10.1186/1471-2148-8-318.

Hassan FA, 2000. Climate and cattle in North Africa: a first approximation. In: Blench RM, MacDonald KC, editors. The origins and development of African livestock: archaeology, genetics, linguistics and ethnography. London: UCL Press. p. 61–86.

Hashizume T, Saito H, Sawada T, Yaegashi T, Ezzat AA, Sawai K and Yamashita T, 2010. Characteristics of stimulation of gonadotropin secretion by kisspeptin-10 in female goats. Animal Reproduction Science Volume 118, Issue 1 37–41.

Hassen H, Lababidi S, Rischkowsky R, Baum M and Tibbo M, 2012a. Molecular characterization of Ethiopian indigenous goat populations. Trop Anim Health Prod 44:1239–1246.

Hassen H, Samer Lababidi, Barbara Rischkowsky, Michael Baum and Markos Tibbo, 2012b. Phenotypic characterization of Ethiopian indigenous goat populations. Afr. J. of Biotech. 11(73), 13838-13846.

Hedrick PW and Kim KJ, 2000. Genetics of complex polymorphisms: parasites and maintenance of the major histocompatibility complex variation. In: Evolutionary Genetics: from Molecules to Morphology (Ed. by R.S. Singh and C.B. Krimbas), pp. 204–234.Cambridge University Press, Cambridge.

Hedrick P and Kumar S, 2001. Mutation and linkage disequilibrium in human mtDNA. Eur J Hum Genet 9:969–972

Heuertz M, De Paoli E, Källman T, Larsson H, Jurman I, Morgante M, Lascoux M, Gyllenstrand N, 2006. Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [Picea abies (L.) Karst]. Genetics 174:2095–2105.

Hill WG and Robertson A, 1966. The effect of linkage on limits to artificial selection. Genet. Res. 8, 269–294.

Hoda A, Hyka G, Dunner S, Obexer-Ruff G, et al, 2011. Genetic diversity of Albanian goat breeds based on microsatellite markers. Archivos de Zootecnia 60(231): 607-615.

Hoda A, Sena L and Hykaj G, 2012. Genetic diversity revealed by AFLP markers in Albanian goat breeds. Archives of Biological Sciences 64(2): 799-807.

Hoffman I, Marsan PA, Barker JSF, Cothran EG, Hanotte O, Lenstra JA, Milan D, Weigend S and Simianer H, 2004. New MoDAD marker sets to be used in diversity studies for the major farm animal species: recommendations of a joint ISAG/FAO working group. In: Proceedings of the 29th International Conference on Animal Genetics. Meiji University, Tokyo, Japan.

Horwitz LK,Tchernov E, Ducos P, Becker C, Von den Driesch A, Martin L, Garrard A, 2000. Animal domestication in the Southern Levant. Pale´ orient 25(2):63–80.

Hou JX, An XP, Wang JG, Song YX, Cui YH, Wang YF, Chen QJ, Cao BY, 2011. New genetic polymorphisms of KiSS-1 gene and their association with litter size in goats. Small Ruminant Res 96:106–110.

Hua GH, Chen SL, Yu JN, Cai KL, Wu CJ, Li QL, Zhang CY, Liang AX, Han L and Geng LY, 2008. Polymorphism of the growth hormone gene and its association with growth traits in Boer goat bucks. Meat Sci 81:391-395.

Hudson RR, 1987. Estimating the recombination parameter of a finite population model without selection. Genet. Res. **50:** 245-250.

Hughes S, Ferna´ndez H, Cucchi T, Duffraisse M, Casabianca F, Istria D, Pompanon F, Vigne JD, Hänni C, Taberlet P, 2012. A Dig into the past mitochondrial diversity of Corsican goats reveals the influence of secular herding practices. PLoS ONE 7(1): e30272. doi:10.1371/journal.pone.0030272.

Huijbregts L, Roze C, Bonafe G, Houang M, Le Bouc Y, Carel JC, Leger J, Alberti P, de Roux N, 2012. DNA polymorphisms of the KiSS1 3 ' Untranslated region interfere with the folding of a G-rich sequence into G-quadruplex. Mol Cell Endocrinol 351:239–248.

Hull KL and Harvey S, 2000. Growth hormone: a reproductive endocrine-paracrine regulator? Rev.Reprod. 5: 175-182.

Hull KL and Harvey S, 2002. GH as a co-gonadotropin: The relevance of correlative changes in GH secretion and reproductive state. J. Endocrinol 172:1-19.

Huson DH and Bryant D, 2006. Application of Phylogenetic Networks in Evolutionary Studies. Mol. Biol. Evol. 23(2):254–267. Huson H, Sonstegard T, Silverstein J, Woodward-Greene M,

Masiga C, Muchadeyi F, et al, 2014.. Genetic and phenotypic characterization of African goat populations to prioritize conservation and production efforts for small-holder farmers in Sub-Saharan africa. Vancouver, Canada: 10th World Congress on Genetics Applied to Livestock Production; Asas.

Hussein Hassen Toyib, 2015. Phenotypic characterization and breeding practices of Arsi-Bale goat population in selected districs of Arsi and Bale zones, Oromia regional state, Ethiopia. MSc thesis. Bahir Dar University, Bahir Dar, Ethiopia.

Hykaj G, Hoda A, Dobi P, Papa L, Bakiu R and Novaku M, 2013. Population structure and genetic distances between six Albanian local goat breeds using 26 SNP markers. Livestock research for Rural development 25:4.

IBC (Institute of Biodiversity Conservation), 2004. The State of Ethiopia's Farm Animal Genetic Resources: Country Report. A Contribution to the First Report on the State of the World's Animal Genetic Resources. IBC. Addis Ababa, Ethiopia.Ikemura T, 1985. Codon usage and tRNA content in unicellular and multicellular organisms. Mol.Biol.Evol 2(1):13-34.

Jelena R, Mio B, Curkovic M, Vesna P, Ivankovic A and Medugorac I, 2011. Genetic diversity measures of the Croatian spotted goat. Acta Veterinaria (Beograd), Vol. 61, No. 4, 373-382, 2011.

Joshi MB, Rout PK, Mandal AK, Tyler-Smith C, Singh L and Thangaraj K,.2004. Phylogeography and origin of Indian domestic goats. Molecular Biology and Evolution, 21: 454–462.

Kalinowski ST, 2002. How many alleles per locus should be used to estimated genetic distances? Heredity, 88: 62-65.

Kalinowski ST, 2004. Counting alleles with rarefaction: Private alleles and hierarchical sampling designs. Conservation Genetics 5: 539–543.

Kalinowski ST, 2005. Do polymorphic loci require large sample sizes to estimate genetic distances? Heredity 94:33–36

Karp A, Kresovich S, Bhat KV, Ayad WG, Hodgkin T, 1997. Molecular tools in plant genetic resources conservation: A guide to the technologies; In: IPGRI Technical Bulletin No. 2. International Plant genetic Resources Institute, Rome Italy.

Kelly JK, 1997. A Test of Neutrality Based on Interlocus Associations. Genetics 146: 1197-1206.

Kevorkian SEM, Georgescu SE, Manea MA, Zaulet M, Hermenean AO, Costache M, 2010. Genetic diversity using microsatellite markers in four Romanian autochthonus sheep breeds. Rom. Biotechnol. Lett. 15(1):5059-5065.

Khatkar MS, Nicholas FW, Collins AR, Zenger KR, Cavanagh JAL, Barris W, Schnabel RD, Taylor JF and Raadsma HW, 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics 2008, 9:187.

Kibegwa FM, Githui KE, Jung'a JO, Badamana MS and Nyamu MN, 2015. Mitochondrial DNA variation of indigenous goats in Narok and Isiolo counties of Kenya. J. Anim. Breed. Genet: 1–10.

Kijas JW, Townley D, Dalrymple BP, Heaton MP, Maddox JF, et al, 2009. A Genome Wide Survey of SNP Variation Reveals the Genetic Structure of Sheep Breeds. PLoS ONE 4(3): e4668. doi:10.1371/journal.pone.0004668.

Kijas JW, OrtizJS, McCulloch R, James A, Brice B, Swain B, Tosser-Klopp G and International Goat Consortium, 2013. Genetic diversity and investigation of polledness in divergent goat populations using 52088 SNPs. Anim. Genet. 44, 325–335.

Kim KI, Yang YH, Lee SS, Park C, Ma R, Bouzat JL and Lewin HA, 1999. Phylogenetic relationships of Cheju horses to other horse breeds as determined by mtDNA D-loop sequence polymorphism. Animal Genetics 30:102–108

Kim KS, Yeo JS, Lee JW, Kim JW and Choi CB, 2002. Genetic diversity of goats from Korea and China using microsatellite analysis. Asian-Australasian Journal of Animal Science 4: 461-465

Kim S and Misra A, 2007. SNP Genotyping: Technologies and Biomedical Applications. Annu. Rev. Biomed. Eng. 9:289–320

Kimura M, 1980. A Simple Method for Estimating Evolutionary Rates of Base Substitutions Through Comparative Studies of Nucleotide Sequences. J. Mol. Evol. 16, 111--120.

Kirilov M, ClarksonJ, Liu X, Roa J, Campos P, Porteous R, Schütz G and Herbison AE,' 2013. Dependence of fertility on kisspeptin–Gpr54 signaling at the GnRH neuron. Nature Communications 4, Article number:2492.

Kotze A, Swart H, Grobler JP and Nemaangani A, 2004. A genetic profile of the Kalahari Red goat breed from Southern Africa. South African Journal of Animal Science 34(1):10-12.

Kumar D, Dixit SP, Sharma R, Pandey AK, Sirohi G, Patel AK, Aggarwal RAK, Verma NK, Gour DS and Ahlawat SPS, 2005. Population structure genetic variation and management of Marwari goats. Small Ruminant Research. 59:41-48.

Kumar S, Gupta T, Kumar N, Dikshit K, Navani N, Jain P, Nagarajan M, 2006. Genetic variation and relationships among eight Indian riverine buffalo breeds. Mol. Ecol., 15: 593-600.

Kumar S, Dixit SP, Verma NK, Singh DK and Pande A, Kumar S, Chander R and Singh LB, 2009. Genetic diversity analysis of the Gohilwari breed of Indian goat (Capra hircus) using microsatellite markers. Am. J. Anim. Vet. Sci., 4: 49-57. DOI: 10.3844/ajavsp.

Landguth EL, Fedy BC, Oyler-Mccance SJ, Garey AL, Emel SL, Mumma, Wagner HH, Fortin MJE and Cushman SA, 2012. Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. Molecular Ecology Resources 12: 276–284.

Larmer S, Ventura R, Buzanskas ME, Sargolzaei M, Schenkel FS, 2014 Assessing admixture by quantifying breed composition to gain historical perspective on dairy cattle in Canada. Vancouver, Canada: 10th World Congress on Genetics Applied to Livestock Production; 2014. Asas.

Lashmar SF, Visser C and van Marle-Köster E, 2015. Validation of the 50k Illumina goat SNP CHIP in the South African Angora goat. South African Journal of Animal Science 2015, 45 (No. 1).

Lee JH, Miele ME, Hicks DJ, Phillips KK, Trent JM, Weissman BE, Welch DR, 1996. "KISS1, a novel human malignant melanoma metastasis-suppressor gene". J. Natl. Cancer Inst. 88 (23): 1731–7.

Lewontin RC, 1974. The genetic basis of evolutionary change. Columbia U.P., N.Y.

Li WH and Sadler LA, 1991. Low nucleotide diversity in man. Genetics 129:513–23

Li, MH, Zhao SH, Bian C, Wanng HS, Wei H, Liu B, Mei Yu M, Fan B, Chen SL, Zhu MJ, Li SJ, Xiong TA, Li K, 2002. Genetic relationships among twelve Chinese indigenous goat populations based on microsatellite analysis. Genet. Sel. Evol. 34: 729–744.

Li MY, Pan QH, Pan YC, Shen W, Min LJ, and Ren ZY, 2004. Polymorphism analysis of goat growth hormone (GH) gene by PCR-RFLP. Journal of Laiyang Agricultural Colleague, 21(1): 6–9 (in Chinese, with English abstract).

Li XL, Wu ZL, Liu ZZ, Gong YF, Zhou RY and Zheng GR, 2006. SNP Identification and Analysis in Part of Intron 2 of Goat MSTN Gene and Variation within and among Species. Journal of Heredity97:3.

Librado P and Rozas J, 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics applications note. Vol. 25 no. 11: 1451–1452

Liu RY, Lei CZ Liu SH and Yang GS, 2007. Genetic diversity and origin of Chinese domestic goats revealed by complete mtDNA D-loop sequence variation. Asian-Aust. J. Anim. Sci. Vol. 20, No. 2: 178 – 183.

LMP (Ethiopian Livestock Master Plan), 2015. Institutions and policies to implement the Ethiopia livestock master plan Ethiopia LMP Brief. https://cgspace.cgiar.org/bitstream/handle/10568/68038/LMP_5.pdf?sequence=1

Lobry JR and Gautier C, 1994 Hydrophobicity, expressivity and aromaticity are the major trends of amino-acid usage in 999 Escherichia coli chromosome-encoded genes. Nucleic Acids Res 22:3174–3180.

Lohse K and Kelleher J, 2009. Measuring the degree of starshape in genealogies—summary statistics and demographic inference. Genet Res 91:281–292.

Luan XH, Zhou YX, Wang W, Yu H, Li P, Gan XH, Wei DZ, Xiao JH, 2007. Association study of the polymorphisms in the KISS1 gene with central precocious puberty in Chinese girls. Eur J Endocrinol, 157:113–118.

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

Luikart G, Fernández H, Mashkour M, England PR and Taberlet P, 2006. Origins and diffusion of domestic goats inferred from DNA markers example analyses of mtDNA, Y Chromosome, and microsatellites. Documenting documentations: New genetic and archaeological paradigms. University of California press. Berkeley Los Angeles London.

Luna-González A, Hernández-Arteaga S, Sánchez-Garza M, López-Revilla R, Medina-Esparza L, Cruz-Vázquez C, 2012. Microsatellite loci and paternity analysis in Nubia and Boer goats. Arch Med Vet 44, 123-127.

Lv FH, Agha S, Kantanen J, Colli L, Stucki S, Kijas JW, Joost S, Li MH, and Marsan PM, 2014. Adaptations to climate-mediated selective pressures in sheep. Mol. Biol. Evol. 31(12):3324–3343

Machado TMM, Chakir M and Lauvergne JJ, 2000. Genetic distances and taxonomic trees between goats of Ceará State (Brazil) and goats of the Mediterranean region (Europe and Africa). Genetics and Molecular Biology, 23, 1, 121-125.

MacHugh DE and Bradley DG, 2001. Livestock genetic origins: Goats buck the trend. PNAS Vol. 98, no. 10, Pp 5382–5384.

Mahmoudi B and Babayev MS, 2009. The investigation of genetic variation in Taleshi goat using microsatellite marker. Res. J. Biol. Sci. 4(6):644-646.

Mahmoudi B, Mansour B, Reza S, Majnun and Hamed, 2010. Genetic Diversity among Three Goat Populations Assessed by Microsatellite DNA Markers in Iran. J. Global Veterinaria 4 (2): 118-124.

Mahmoudi B, Babayev MSh, Hayeri Khiavi F, Pourhosein A and Daliri M, 2011. Breed characteristics in Iranian native goat populations. Journal of Cell and Animal Biology Vol. 5(7), pp. 129-134.

Majo FC, 2008. Selection and linkage desequilibrium tests under complex demographies and ascertainment bias. Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra (UPF). PhD dissertation. Barcelona.

Maletsanake D, Nsoso SJ and Kgwatalala PM, 2013. Genetic variation from 12 microsatellite makers in an indigenous Tswana goat flock in South-eastern Botswana. Livestock Research for Rural Development 25 (2).

Manceau V, Després L, Bouvet J, Taberlet P, 1999. Systematics of the genus Capra inferred from mitochondrial DNA sequence data. Mol. Phylogenet. Evol. 13, 504–510.

Manel S, Berthier P and Luikart G, 2002. Detecting wildlife poaching: identifying the origin of individuals with Bayesian assignment test and multilocus genotypes. Conservation Biology 16, 650–9.

Mannen H, Nagata Y and Soichi Tsuji S, 2001. Mitochondrial DNA reveal that domestic goat (*Capra hircus*) are genetically affected by two subspecies of bezoar (*Capra aegagurus*). Biochemical Genetics Vol. 39, Nos. 5/6.

Mannen H, Kohno M, Nagata Y, Tsuji S, Bradley DG, et al, 2004. Independent mitochondrial origin and historical genetic differentiation in North Eastern Asian cattle. Mol Phylogenet Evol 32: 539–544.

Marroni F, Pinosio S, Zaina G, Fogolari F, Felice N, Cattonaro F and Morgante M, 2011. Nucleotide diversity and linkage disequilibrium in Populus nigra cinnamyl alcohol dehydrogenase (CAD4) gene. J. Tree Genetics and Genomes 7:1011–1023.

Marshall TC, Slate J, Kruuuk LEB and Pemberton JM, 1998. Statistical confidence for likelihood-based paternity inference in natural population. Molecular Ecology 7: 639-655

Martínez AM, Carrera MP, Acosta JM, Rodríguez-Gallarda PP, Cabello A, Camacho E, Delgado JV, 2004. Genetic characterization of the Blanca Andaluza goat based on microsatellite markers. S. Afr. J. Anim. Sci. 34(1):17-19.

Matukumalli LK, Cynthia T, Robert L, Schnabel D, Taylor JF, Allan MF, Heaton MP, O'Connell J, Moore SS, SmithTPL, Sonstegard TS and Van Tassell CP, 2009. Development and characterization of a high density SNP genotyping assay for cattle. PLoSONE 4, e5350. doi:10.1371/journal.pone.0005350.

Maynard Smith J and Haigh J, 1974. The hitch-hiking effect of a favourable gene. Genet. Res. 23, 23–35.

Mazdarani FH, Akbari MT, Fard RMN, Hessari M, Pour KC, 2014. Molecular identification of *Capra hircus* in East Chia Sabz, an Iranian pre-pottery Neolithic site, Central Zagros, based on mtDNA. The Journal of Animal & Plant Sciences, 24(3): 945-950.

McRae AF, McEwan JC, Dodds KG, Wilson T, Crawford AM, Slate J, 2002. Linkage disequilibrium in domestic sheep. Genetics 160 (3):1113-1122.

Meadows JRS, Li K, Kantanen J, Tapio M, Sipos W, et al, 2005. Mitochondrial Sequence Reveals High Levels of Gene Flow Between Breeds of Domestic Sheep from Asia and Europe. J Hered 96(5): 494–501.

Meadows JRS, Cemal I, Karaca O, Gootwine E, Kijas JW, et al, 2007. Five ovine mitochondrial lineages identified from sheep breeds of the Near East. Genetics 175: 1371–1379.

Meghen C, MacHugh DE and Bradley DG, 1994. Genetic characterization and West African cattle. World Anim. Rev.78: 59-66.

Mekuriaw G, Gizaw S and Tegegne A, 2016. Genetic-phenotypic and production system diversity in goat populations of Ethiopia: Options for sustainable goat production systems. LIVES Working Paper 22. Nairobi, Kenya: International Livestock Research Institute (ILRI).

Melinda ZA, Emshwiller E, Smith BD and Bradley DG, 2006. Documenting domestication: the intersection of genetics and archaeology. Trends Genet 22:139–155.

Messager S, Chatzidaki EE, Ma D, Hendrick AG, Zahn D, Dixon J, Thresher RR, Malinge I, Lomet D, Carlton MB, Colledge WH, Caraty A, Aparicio SA, 2005. "Kisspeptin directly stimulates gonadotropin-releasing hormone release via G protein-coupled receptor 54". Proc. Natl. Acad. Sci. U.S.A. 102 (5): 1761–6.

Mestawet TA, Girma, Ådnøy T, Dagnachew B, Sundsaasen KK, Lien S, Kent MP, Devold TG, Narvhus JA,. Vegarud GE, 2013. New genetic polymorphism at the αs1-casein gene region in Ethiopian indigenous goat breeds. Small Ruminant Res., http://dx.doi.org/10.1016/j.smallrumres.2013.03.012

Mpofu N, 2002. The multiplication of Africa's indigenous cattle breeds internationally: The story of the Tuli and Boran breeds. AGTR Case Study. Nairobi, Kenya: ILRI. https://cgspace.cgiar.org/handle/10568/3600.

Msoffe PLM, Minga UM, Olsen JE., Yongolo MGS, Juul-Madsen HR, Gwakisa PS, Mtambo MMA, 2001. Phenotypes including immunocompetence in scavenging local chicken ecotypes of Tanzania. Tropical Animal Health and Production 33: 341-354.

Msoffe PLM, Mtambo MMA, Minga UM, Olsen JE, Juul-Madsen HR, Gwakisa PS, Mutayoba SK, Katule AM, 2004. Productivity and reproductive performance of the free-range local domestic fowl ecotypes in Tanzania. Livestock Research for Rural Development. Vol.16, Art.#67.http://www.cipav.org.co/lrrd/lrrd16/9/msof16067.htm.

Mucha S, Mrode R, Coffey M, Conington J, 2014. Estimation of genomic breeding values for milk yield in UK dairy goats. 10th world congress on genetics applied to livestock production. Canada: Vancouver. Asas.

Mueller JC, 2004. Linkage disequilibrium for different scales and applications. Brief Bioinform 5:355–364.

Muema EK, Wakhungu JW, Hanotte O and Jianlin Han, 2009. Genetic diversity and relationship of indigenous goats of Sub-saharan Africa using microsatellite DNA markers. Livestock Research for Rural Development. Volume 21, Article #28. http://www.lrrd.org/lrrd21/2/muem21028.htm

Nachman MW, 2002. Variation in recombination rate across the genome: evidence and implications. Curr. Opin. Genet. Dev. 12:657-663.

Naderi S, Rezaei HR, Taberlet P, Zundel S, Rafat SA, Naghash HR, Mohamed A, El-Barody A, Ertugrul O, Pompanon F and, the Econogene Consortium, 2007. Large-scale mitochondrial DNA analysis of the domestic goat reveals six maternal lineages with high haplotype diversity. PLoS ONE 10:e1012.

Naderi S, Rezaei HR, Pompanon FO, Blum MGB, Negrini R, Naghash HR, Balkız O, Mashkour M, Gaggiotti OE, Ajmone-Marsan P, Kence A, Vigne JD and Taberet P, 2008. The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. PNAS vol. 105, no. 46, 17659–17664.

Nassiry MR, Javanmard A and Tohidi R, 2009. Application of statistical procedures for analysis of genetic diversity in domestic animal populations. Am J Anim Vet Sci 4:136-141.

Navarro VM, Castellano JM, Fernandez-Fernandez R, Tovar S, Roa J, Mayen A, Nogueiras R, Vazquez MJ, Barreiro ML, Magni P, et al, 2005. Characterization of the potent luteinizing

hormone-releasing activity of KISS1 peptide, the natural ligand of GPR54. Endocrinology 146:156–163.

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

Nei M, Chakravarti A and Tateno Y, 1977. Mean and variance of Fst in a finite number of incompletely isolated populations. Theor. Pop. Biol. 11: 291-306.

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

Netsanet Zergaw, 2014. On-farm phenotypic characterizations and performance evaluation of Central highland goat and Woyto-Guji goat types for designing community-based breeding strategies. MSc theis. Haramaya University.

Nicholls JA and Austin JJ, 2005. Phylogeography of an east Australian wet-forest bird, the satin bowerbird (Ptilonorhynchus violaceus), derived from mtDNA, and its relationship to morphology. J.Mol Ecol, 2005 Apr;14(5):1485-96.

Nickerson DA, Taylor SL, Weiss KM, Clark AG, Hutchinson RG, Stengård J, Salomaa V, Vartiainen E, Boerwinkle E and Sing CF, 1998. DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. Nat. Genet. 19:233–40.

Nicoloso L, Bomba L, Colli L, Negrini R, Milanes M, Mazza R, Sechi T, Frattini S, Talenti A, Coizet B, Chessa S, Marletta D, D'Andrea M, Bordonaro S, Ptak G, Carta A, Pagnacco G, Valentini A, Pilla F, Ajmone-Marsan P, Crepaldi P and the Italian Goat Consortium, 2015. Genetic diversity of Italian goat breeds assessed with a medium-density SNP CHIP. Genetics Selection Evolution 47:62.

Nomura K, Ishii K, Dadi H, Takahashi Y, Minezawa M, Cho CY, Sutopo, Faruque MO, Nyamsamba D and Amano T, 2012. Microsatellite DNA markers indicate three genetic lineages in East Asian indigenous goat populations. Animal Genetics, 43, 760–767.

Nomura K, Yonezawa T, Mano S, Kawakami S, Shedlock AM, et al, 2013. Domestication Process of the Goat Revealed by an Analysis of the Nearly Complete Mitochondrial Protein-Encoding Genes. PLoS ONE 8(8): e67775.

Novoa EM, Ribas de Pouplana L, 2012. Speeding with control: codon usage, tRNAs, and ribosomes. Trends Genet 28(11): 574–81.

Nsoso SJ, Machete JB, Molatole M, Ndebele RT, Lebani NN, Chabo RG, Kalake A M, Jacyna L, Segadimo BW and Mine OM, 2001. The impact of traditional management on seasonal internal parasite burdens and productivity of indigenous Tswana goats in southern Botswana. Onderstepoort Journal of Veterinary Research 68: 101-104.

Ojango JM, Mpofu N, Marshall K and Andersson-Eklund L, 2011. Quantitative methods to improve the understanding and utilisation of animal genetic resources In:Animal Genetics Training Resource, version 3. Ojango, J.M., Malmfors, B. and Okeyo, A.M. (Eds). International Livestock Research Institute, Nairobi, Kenya, and Swedish University of Agricultural Sciences, Uppsala, Sweden.

Okamura H, Tsukamura H, Ohkura S, Uenoyama Y, Wakabayashi Y, Maeda K, 2013. Kisspeptin and GnRH Pulse Generation. Advanced in experimental medicine and biology. Volume 784: 297-323

Ola SI, Ai JS, Liu JH, Wang Q, Wang ZB, Chen DY and Sun QY, 2008. Effects of gonadotrophins, growth gormone, and activin A on enzymatically isolated follicle growth, oocyte chromatin organization, and steroid secretion. Mol Reprod Dev 75:89-96.

Pagani L, Kivisild T, Tarekegn A, Ekong R, Plaster C, Romero IG, Ayub Q, Mehdi SQ, Thomas MG, Luiselli D, Bekele E, Bradman N, Balding DJ and Tyler-Smith C, 2012. Ethiopian genetic diversity reveals linguistic stratification and complex influences on the Ethiopian gene pool. The American journal of Human genetics 91:83–96.

Pariacote F, 2006. The goat in South America and development perspectives. In: international symposium goat farming Central, Eastern European countries, present and future. IGA-CAPRIROM. Constanta, Romania.

Pariset L, Cuteri A, Ligda C, Ajmone-Marsan P, Valentini A and ECONOGENE Consortium, 2009a. Geographical patterning of sixteen goat breeds from Italy, Albania and Greece assessed by Single Nucleotide Polymorphisms. BMC Ecology9:20.

Pariset L, Joost S, Marsan PA, Valentini A, 2009b. Landscape genomics and biased FST approaches reveal single nucleotide polymorphisms under selection in goat breeds of North-East Mediterranean. BMC genetics 10, (1): 7.

Payne WJA and Wilson RT, 1999. An introduction to Animal Husbandry in the tropics. Fifth edition. Blackwell Science Ltd: London. Pereria F., Pereira L., Van Asch B., Bradley D. G. and Amorim A, 2005. The mtDNA catalogue of all Portuguese autochthonous goat (Capra hircus) breeds: high diversity of female lineages at the western fringe of European distribution. Molecular Ecology 14:2313–2318.

Payseur BA, Nachman MW, 2000. Microsatellite variation and recombination rate in the human genome. Genetics 156:1285–1298.

Pei YF, Zhang L, Liu J and Deng HW, 2009. Multivariate association test using haplotype trend regression. Ann Hum Genet. 73(4): 456–464.

Pemberton JM, Slate J, Bancroft DR, Barrett JA, 1995. Non amplifying alleles at microsatellite loci: A caution for parentage and population studies. Mol. Ecol., 4: 249-252.

Pereira F, Pereira L, Van-Asch B, Bradley DG and Amorim A, 2005. The mtDNA catalogue of all Portuguese autochthonous goat (Capra hircus) breeds: high diversity of female lineages at the western fringe of European distribution. Molecular Ecology 14: 2313–2318.

Pereira F, Queirós S, Gusmão L, Nijman IJ, Cuppen E, Lenstra JA, Econogene Consortium, Davis SJM, Nejmeddine F and Amorim A, 2009. Tracing the History of Goat Pastoralism: New clues from mitochondrial and Y Chromosome DNA in North Africa. Mol. Biol. Evol. 26(12): 2765–2773.

Pidancier N, Jordan S, Luikart G, Taberlet P, 2006. Evolutionary history of the genus Capra (Mammalia, Artiodactyla): Discordance between mitochondrial DNA and Y-chromosome phylogenies. Molecular Phylogenetics and Evolution 40: 739–749.

Polley S, De S, Batabyal S, Kaushik R, Yadav P, Arora JS, Chattopadhyay S, Pan S, Brahma B and Datta TK, 2009. Polymorphism of fecundity genes (BMPR1B, BMP15 and GDF9) in the Indian prolific Black Bengal goat. Small Rumin Res 85:122-129.

Porter V, 1996. Goats of the World. Farming Press, Ipswich, U.K.

Pramod KR, Manjunath BJ, Ajoy M, Laloe D, Singh L and Thangaraj K, 2008. Microsatellite-based phylogeny of Indian domestic goats. BMC Genetics 9:11.

Pritchard JK, Stephens M and Donnelly P, 2000. Inference of population structure from multilocus genotype data. Genetics 155, 945–59.

Pritchard JK and Przeworski M, 2001. Linkage disequilibrium in humans: models and data. Am. J. Hum. Genet. 69:1–14, 2001.

Prunier A and Quesnel H, 2000. Influence of the nutritional status on ovarian development in female pigs. Anim Reprod Sci 61:185-197.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ and Sham PC, 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. American journal of human genetics, 81. http://pngu.mgh.harvard.edu/purcell/plink/.

Qi Y, J Luo XF Han Zhu YZ, Chen C, Liu and JX and Sheng HJ, 2009. Genetic diversity and relationships of 10 Chinese goat breeds in the Middle and Western China. Small Rumin. Res. 82:88-93.

Qian W, Yang JR, Pearson NM, Maclean C, Zhang J, 2012. Balanced codon usage optimizes eukaryotic translational efficiency. PLoS Genet 8(3).

Ramamoorthi J, Thilagam K, Sivaselvam SN and Karthickeyan SMK, 2009. Genetic characterization of Barbari goats using microsatellite markers. Journal of Veterinary Science 10(1): 73-76.

Rege JEO, 1999. The state of African cattle resources I. Classification framework and identification of threatened and extinct breeds. Animal Genetic Resources Bulletin 25:1-25. population structure, assignment, and demographic analyses of bowhead whale populations. SNPs for bowhead whale population genetics. SC/64/AWMP3.

Rege JEO, 2002. Defining livestock breeds in the context of community-based management of farm animal genetic resources. In: Animal Genetic Resources Virtual Library CD-ROM. GTZ, SACCAR and ILRI. ILRI-Ethiopia@cgiar.org, Nairobi.

Rendo F, Iriondo M, Jugo BM, Mazo´n LI, Aguirre A, Vicario A and Estonba A, 2004. Tracking diversity and differentiation in six sheep breeds from the North Iberian Peninsula through DNA variation. Small Ruminant Research 52, 195–202.

Reynolds J, Weir BS, and Cockerham CC, 1983. Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics 105:767-779.

Robles-Sikisaka R, Rivera R, Nollens HH, Leger JSt, Durden WN, Stolen M, Burchell J, Wellehan Jr, James FX, 2012. Evidence of recombination and positive selection in cetacean papilloma viruses. Virology 427: 189–197.

Rodero A, Delgado JV and Rodero E, 1992. Primitive Andalusian livestock: their implications in the discovery of America. Archivos de Zootecnia 41: 383–400.

Rogers AR, Harpending H, 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552-569.

Rozas J, Gullaud M, Blandin G and Aguade M, 2001. DNA variation at the rp49 gene region of Drosophila simulans: Evolutionary inferences from an unusual haplotype structure. Genetics 158: 1147-1155.

Rozas J, Sanchez-DelBarrio JC and Messeguer X, 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496-2497.

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

Sardina MT, Ballester M, Marmi J, Finocchiaro R, van Kaam JBCHM, Portolano B and Folch JM, 2006. Phylogenetic analysis of Sicilian goats reveals a new mtDNA lineage. Anim Genet 37(4): 376–378.

Schlumbaum A, Campos PF, Volken S, Volken M, Hafner A, Schibler J, 2010. Ancient DNA, a Neolithic legging from the Swiss Alps and the early history of goat, J. Archaeol. Sci. doi:10.1016/j.jas.2009.12.025.

Schwaiger FW, Buitkamp J, Weyers E and Epplen JT, 1993. Typing of MHC-DRB genes with the help of intronic simple repeated DNA sequences. Molecular Ecology 2: 55–59.

Schwerin M, Brockmann G, Vanselow J and Seyfert HM,. 1995. Perspectives of molecular genome analysis in live-stock improvement. Arch. Fur. Tierz. 38, 21-31.

Serrano M, CalvoJH, Martínez M, Marcos-Carcavilla A, Cuevas J, González C, Jurado JJ, de Tejada PD, 2009. Microsatellite based genetic diversity and population structure of the endangered Spanish Guadarrama goat breed. BMC Genet.10: 61.

Service SK, Temple Lang DW, Freimer NB and Sandkuijl LA, 1999. Linkage-disequilibrium mapping of disease genes by reconstruction of ancestral haplotypes in founder populations. Am. J. Hum. Genet. 64:1728–1738.

Shahab M, Mastronardi C, Seminara SB, Crowley WF, Ojeda SR, Plant TM, 2005. Increased hypothalamic GPR54 signaling: a potential mechanism for initiation of puberty in primates. Proc Natl Acad Sci USA . 102:2129–2134.

Sharp PM and Li WH, 1987. The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acid Research, 15,1281-1295.

Shifman S, Bronstein M, Sternfeld M, Pisante-Shalom A, Lev-Lehman E, Weizman A, Reznik I, Spivak B, Grisaru N, Karp L, Schiffer R, Kotler M, Strous RD, Swartz-Vanetik M, Knobler HY, Shinar E, Beckmann JS, Yakir B, Risch N, Zak NB, Darvasi A, 2002. A highly significant association between a COMT haplotype and schizophrenia. Am J Hum Genet;71: 1296-1302.

Shinde DB, Gujar BV, Patil HB, Satpute SA and Kashid UB, 2008. Protocol for isolation of genomic DNA from livestock blood for microsatellite analysis. Indian J. Anim. Res., 42 (4): 279-281.

Silva JRV, Figueiredo JR and van den Hurk R, 2009. "Involvement of growth hormone (GH) and insulin-like growth factor (IGF) system in ovarian folliculogenesis," Theriogenology, vol. 71, no. 8, pp. 1193–1208.

Silveira LG, Noel SD, Silveira-Neto AP, Abreu AP, Brito VN, Santos MG, Bianco SDC, Kuohung W, Xu S, Gryngarten M, Escobar ME, Arnhold IJP, Mendonca BB, Kaiser UB, and Latronico AC, 2010. Mutations of the KISS1 gene in disorders of puberty. J Clin Endocr Metab, 95:2276–2280.

Simpson GG, 1945. The principles of classification of mammals. Bull. Am. Mus. Nat. Hist. 85, 1-350.

Sirotkin AV, Mertin D, Süvegová K, Makarevich AV and Mikulová E, 2003. Effect of GH and IGF-I treatment on reproduction, growth, and plasma hormone concentrations in domestic nutria (Myocastor coypus). Gen Comp Endocrinol 131:296-301.

Sisay Gezahegn, 1996. Characterization of some indigenous cattle breeds of Ethiopia using blood protein polymorphisms. M.Sc. Thesis. Alemaya University of Agriculture, Ehtiopia.

Slatkin M, 1981. Estimating levels of gene flow in natural populations. Genetics 99 323-335.

Slatkin M, 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics, 139: 457.

Slatkin M, 2008. Linkage disequilibrium: understanding the evolutionary past and mapping the medical future. Nature Reviews Genetics 9: 477-485 doi:10.1038/nrg2361.

Smith AB, 1992. Origins and spread of pastoralism in Africa. Ann Rev Anthropol. 21:125–141.

Smith BD, 1998. The emergence of agriculture. New York: Scientific American Library.

Smith JT, Clifton DK, Steiner RA, 2006. "Regulation of the neuroendocrine reproductive axis by kisspeptin-GPR54 signaling". Reproduction 131 (4): 623–30.

Smith JT, Saleh SNH, Clarke IJ, 2009. Seasonal and cyclical change in the luteinizing hormone response to kisspeptin in the ewe. Neuroendocrinology 90:283–291

Smith JT, Li Q, Yap KS, Shahab M, Roseweir AK, Millar RP, Clarke IJ, 2011. Kisspeptin is essential for the full preovulatory LH surge and stimulates GnRH release from the isolated ovine median eminence. Endocrinology 152:1001–1012.

Sodhi M, Mukesh M, Mishra BP, Mitkari KR, Prakash B and Ahlawat SP, 2005. Evaluation of genetic differentiation in Bos indicus cattle breeds from Marathwada region of India using microsatellite polymorphism. Animal Biotechnology, 16: 127–137.

Sokal RR and Rohlf FJ, 1981. Biometry. Second Edition. W. H. Freeman and Company. New York.

Solomon AG, 2014. Design of community based breeding programs for two indigenous goat breeds of Ethiopia. Doctoral Thesis: University of natural resources and applied life sciences, department of sustainable agricultural systems division of livestock sciences. Vienna, Austria.

Soule M, 1987. Viable populations for conservation. Cambridge Univ. press, Cambridge.

Städler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P, 2009. The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. Genetics 182:205-216.

Sultana S, Mannen H, Tsuji S, 2003. Mitochondrial DNA diversity of Pakistani goats. Anim Genet 34(6): 417–421.

Sunnucks P, 2000. Efficient genetic markers for population biology. Trends in Ecology and Evolution, 15: 199-206.

Tadelle Dessie, 2003. Phenotypic and genetic characterization of local chicken eco-types in Ethiopia. Dissertation, Landwirtschaftlich-Gaertnerischen Fakultaet der Humboldt-Universitaet zu Berlin, Germany, ISBN 3-89574-497-2, 208pp.

Tajima F, 1989a. The effect of change in population size on DNA polymorphism. Genetics 123:597–601

Tajima F, 1989b. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123 (3): 585–95.

Takada T, Kikkawa Y, Yonekawa H, Kawakami S, Amano T, 1997. Bezoar (Capra aegagrus) is a matriarchal candidate for ancestor of domestic goat (Capra hircus): evidence from the mitochondrialDNAdiversity. Biochem Genet. 35: 315-326.

Takahashi H, Nyamsamba D, Mandakh B, Zagdsuren Yo, Amano T, Nomura K, Yokohama M, Ito S and Minezawa M, 2008. Genetic structure of Mongolian goat populations using microsatellite loci analysis. Asian-Aust. J. Anim. Sci. 21:947-953.

Tamura K, Stecher G, Peterson D, Filipski A and Kumar S, 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30(12):2725–2729.

Tapio M, Tapio I, Grislis Z, Holm LE, Jeppsson S, Kantanen J, Miceikiene I, Olsaker I, Viinalass H and Eythorsdottir E, 2005. Native breeds demonstrate high contributions to the molecular variation in northern European sheep. Molecular Ecology, 14: 3951–3963.

Tesfaye Alemu Tucho, 2004. Genetic characterization of indigenous goat populations of Ethiopia using microsatellite DNA markers. A PhD thesis submitted to the nationa dairy research institute, Deemed University. Karnal-Haryana, India.

Tesfaye Getachew, 2015. Genetic diversity and admixture analysis of Ethiopian Fat-tailed and Awassi sheep using SNP markers for designing crossbreeding schemes. University of Natural Resources and Life Sciences, Vienna, Austria.

Thompson JD Higgins DG and Gibson TJ, 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res., 22, 4673–4680.

Thorpe R., Black H and Malhotra A, 1996. Matrix correspondence tests on the DNA phylogeny of the Tenerife Lacertid elucidate both historical causes and morphological adaptation. Systematic Biology 45: 335–43.

Tibbo M, 2006. Productivity and Health of Indigenous Sheep Breeds and Crossbreds in the Central Ethiopian Highlands. PhD thesis, Swedish University of Agricultural Sciences.

Tokarska M, Marshall T, Kowalczyk R, Wojcik JM, Pertoldi C, Kristensen TN, Loeschcke V, Gregersen VR and Bendixen C, 2009. Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: the case of European bison. Heredity 103, 326–332.

Tomikawa J, Uenoyama Y, Ozawa M, Fukanuma T, Takase K, Goto T, Abe H, Ieda N, Minabe S, Deura C, Inoue N, Sanbo M, Tomita K, Hirabayashi M, Tanaka S, Imamura T, Okamura H, Maedaf K and Tsukamura H, 2012. Epigenetic regulation of Kiss1 gene expression mediating estrogen-positive feedback action in the mouse brain. www.pnas.org/cgi/doi/10.1073/pnas.1114245109

Toro M A, Fernández J, and Caballero A, 2006. Scientific basis for policies in conservation of farm animal genetic resources. Proc. 8th World Congr. Genet. Appl. Livest. Prod., Belo Horizonte, Brazil [CD-ROM ed.]. 8th WCGALP Secretariat, Belo Horizonte,MG, Brazil.

Tosser-Klopp G, Bardou P, Bouchez O, Cabau C, Crooijmans R, Dong Y, et al, 2014. Design and characterization of a 52 K SNP CHIP for goats. PLoS One:9, e86227Uzzaman Md.R.,

Trizio I, Crestanello B, Galbusera P, Wauters A, Tosi G, Matthysen E and Hauffe C, 2005. Geographical distance and physical barriers shape the genetic structure of Eurasian squirrels (Sciurus vulgaris) in the Italian Alps. Molecular Ecology 14: 469–81.

Uzzaman MdR, Edea Z, MdSA, Walker J, Bhuiyan AKFH and Kim KS, 2014. Genome-wide single nucleotide polymorphism analyses reveal genetic diversity and structure of wild and domestic cattle in Bangladesh. Asian Australas. J. Anim. Sci. Vol. 27, No. 10: 1381-1386.

Väli Ülo, Annika Einarsson, Lisette Waits and Hans Ellegren, 2008. To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? Molecular Ecology 17, 3808–3817.

Verma NK, Dixit SP, Aggarwal RAK, Chander R, Kumar S, Ahlawat SPS, 2007. Genetic analysis of the Sirohi breed of Indian goat (*Capra hircus*). Korean J. Genetics. 29:129-136.

Voight BF, Kudaravalli S, Wen X, Pritchard JK, 2006. A map of recent positive selection in the human genome. PLoS Biol;4:e72.

Voruz JL, 1999. Chronologies Néolithiques. De 6000 á 2000 avant nótre ére dans le Bassin Rhodanien. Revue archéologique du Center de la France, Année 1999, Volume 38, Numéro 1. P. 236-237.

Weir BS, 1996. Genetic Data Analysis II. Sinauer Associates, Inc, Sunderland, MA.

Weir BS and Basten CJ, 1990. Sampling strategies for distances between DNA sequences. Biometrics 46:551-582.

Weir BS, Cardon LR, Anderson AD, Nielsen DM, Hill WG, 2005. Measures of human population structure show heterogeneity among genomic regions. Genome Res; 15:1468-1476.

Weir BS and Cockerham CC, 1984. Estimating F-statistics for the analysis of population structure. Evolution, 38, 1358–1370.

Whiteley AR, Spruell P and Allendorf FW, 2004. Ecological and life history characteristics predict population genetic divergence in two salmonids in the same landscape. Molecular Ecology 13, 3675–88.

Widmer A and Lexer C 2001. Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. TRENDS in Ecology and Evolution Vol.16 No.6

Workneh Ayalew, 1992. Preliminary survey of indigenous goat types and goat husbandry practices in Southern Ethiopia. M.Sc. Thesis, Alemaya University of Agriculture.

Wright S, 1943. Isolation by distance. Genetics 38 : 114 March 1943

Wu WS, Hua GH, Yang LG, Wen QY, Zhang CY, Khairy MZ and Chen SL, 2009. Association analysis of the INHA gene with litter size in Boer goats. Small Rumin Res 82:139-143.

Yadav A and Yadav BR, 2008. DNA Fingerprint: Genetic relationship in six Indian goat breeds. Indian J. Biotechnol. 7:487-490.

Yilma Deressa, 1967 (1959 E.C.). የኢትዮጵያ ታሪክ በ16ኛው ክ.ዘ. በብርሀንና ሰላም ቀዳማዊ ኃይለ ሥላሴ ማተሚያ ቤት ታተመ፡፡ አዲስ አበባ ህዳር 20 1959 ዓ.ም.

Yu A, Zhao C, Fan Y, Jang W, Mungall AJ, Deloukas P, Olsen A, Doggett NA, Ghebranious N, Broman KW, Weber JL, 2001. Comparisons of human genetic and sequence-based physical maps. Nature 409:951–953.

Yuqin W, Li YX, Zhang NN, Wang ZB, Bai JY, 2011. Polymorphism of exon 2 of BMP15 gene and its relationship with litter size of two Chinese goats. Asian Austral J Anim, 24: 905–911.

Zeder MA, 2008. "Animal domestication in the Zagros: An update and directions for future research" in Archaeozoology of the Near East, VIII: Proceedings of the Eighth International Symposium on the Archaeozoology of Southwestern Asia and Adjacent Area. Edited by Emanuelle Vila, Lionel Goucherin, Alice Choyke, and Hilke Buitenhuis, pp. 243–78.

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

Zerabruk M, Vangen O and Haile M, 2007. The status of cattle genetic resources in North Ethiopia: On-farm characterization of six major cattle breeds. Animal Genetic Resources Information, No. 40.

Zeuner FE, 1963. A history of domestic animals. Hutchinson and Co.Ltd:London

Zhang C, Zhang W, Luo H, Yue W, Gao M and Jia Z, 2008a. A New Single Nucleotide Polymorphism in the IGF-I Gene and its association with growth traits in the Nanjiang Huang goat. Asian-Aust. J. Anim. Sci. Vol. 21(8): 1073-1079.

Zhang W, Sun X, Yuan H et al. 2008b. The pattern of insertion/deletion polymorphism in Arabidopsis thaliana. Mol. Genet. Genomics 280: 351–361.

Zhang C Liu Y, Huang K, Zeng W, Xu D, Wen Q and Yang L, 2011. The association of two single nucleotide polymorphisms (SNPs) in growth hormone (GH) gene with litter size and superovulation response in goat-breeds. Genetics and Molecular Biology, 34(1): 49-55.

Zhou HM, Allain D, Li JQ, et al, 2003. Effects of non-genetic factors on production traits of Inner Mongolia Cashmere goats in China. J. Small Ruminant Research, 47(1): 85-89.

Zhu M, Zhu B, Wang YH, Wu Y, Xu L, Guo LP, Yuan ZR, Zhang LP, Gao X, Gao HJ, Xu SZ and Li JY, 2013. Linkage disequilibrium estimation of Chinese beef Simmental cattle using high-density SNP panels. Asian Australas. J. Anim. Sci. Vol. 26(6): 772-779.

Zidi A, Abo-Shady H, Molina A, Menéndez-Buxadera A, Sánchez-Rodríguez M, Díaz C, Carabaño MJ and Serradilla JM, 2014. Genome wide association for heat stress tolerance/susceptibility in Florida dairy goats. Vancouver, Canada: 10th World congress on genetics applied to livestock production.

Zilháo J, 2001. Radiocarbon evidence for maritime pioneer colonization at the origin of farming in west Mediterranean Europe. Proc. Natl. Acad. Sci. USA 98:14180-14185.

10 Appendix

10.1 Appendix table

Appendix Table A. Reference sequences employed for haplogroup analysis

Country of origin	Sample size	Haplogroup	Accession number (Reference)
Ethiopia	309	A, G	This study
Iran	25	A, G	EF617945, EF617863-EF618084 (Naderi et al., 2007)
Iraq	7	A	AJ317762-68 (Luikart et al., 2001)
Pakistan	40	A,B,C,D	AB110552–AB110591 (Sultana et al., 2003)
Kenya	58	A, G	KP120622-KP120681 (Kibegwa et al., 2015)
Saudi Arabia	43		AJ317752-59 (Luikart et al., 2001); EF618309-45
Egypt	26	A, G	AJ317780-83; AJ317795-801 (Luikart et al., 2001); EF617711-28
Nigeria	12		AJ317810-811; AJ317823-25 (Luikart et al., 2001); EP618246-52
India	3	A, D	AY155721, AY155708, AY155952 (Joshi et al., 2004)
Turkey	1	G	EF618535 (Naderi et al 2007)
Sicily	2	F	DQ241349;DQ241351 (Sardina et al 2006)
China	3	B, C, D	DQ121578 (Liu et al 2006); DQ188892, DQ188893 (Liu et al 2005)
Austria	1	D	EF617701 (Naderi et al 2007)
Mongolia	1	B	AJ317833 (Luikart et al 2001)
Azerbaijan	1	B	EF617706 (Naderi et al 2007)
Laos	1	B	AB044303 (Mannen et al 2001)
Jordan	1	A	EF618200 (Naderi et al. 2007)
France	1	A	EF617779 (Naderi et al 2007)
Italy	1	A	EF618134 (Naderi et al., 2007)
Switzerland	1	C	AJ317838 (Luikart et al 2001)
Spain	1	C	EF618413 (Naderi et al 2007)

Appendix Table B. Primers employed for D-loop amplification and sequencing

Category	Name	Sequence	Purpose
External primers	tRNA-Phenaylalanine-F	5’-CACCATCAACCCCAAAGCTG-3’	Amplification and sequencing
	tRNA-Proline-R	5’-CAGTGCCTTGCTTTGGTTAAGC-3’	
	BDG-F	5’- CATCTGCTTCTTCTTCAG GGCCATC-3’,	
	HC3-R	5’-TGGACTCAGCTATGGCCGTC-3’	
Internal primers	GDLS-1F	5’-GCGGACATACAGCCTTCATA-3’	
	GDLS-1R	5’-ATATCTAGGAGGG AGCGTGT-3’	
	GDLS-2F	5’-ACCT AAAATCGCCCACTC-3’	
	GDLS-2R	5’-TGATCTAG TGGACGGGATAC-3	

Appendix Table C. Haplotypes shared by populations

Shared haplotype	No.of sequences	population	Number
ET8	7	Hararghe Highland	1
		Ambo	2
		Abergelle	1
		Arsi-Bale	1
ET14	3	Abergelle	1
		Gumuz	1
		Ambo	1
ET30	3	Long eared Somali	2
		Hararghe highland	1
ET31	7	Abergelle	2
		Nubian	2
		Gumuz	1
		Afar	1
		Agew	1
ET38	7	Kaffa	6
		Agew	1
ET39	4	Ambo	3
		Woyto-Guji	1
ET41	4	Agew	2
		Gondar	1
		Abergelle	1
ET44	3	Arsi-Bale	2
		Kaffa	1
ET48	3	Hararghe Highland	1
		Afar	2
ET52	2	Kaffa	1
		Gondar	1
ET55	2	Hararghe Highland	1
		Ambo	1
ET63	3	Nubian	1
		Kaffa	1
		Ambo	1
ET84	4	Nubian	3
		Afar	1
ET86	3	Arsi-Bale	2
		Hararghe Highland	1
ET102	5	Gumuz	1
		Nubian	4
ET107	2	Abergelle	1
		Arsi Bale	1
ET108	3	Gumuz	1
		Agew	2
ET115	2	Abergelle	1
		Gumuz	1
ET141	2	Small eared Somali	1
		Afar	1
ET147	7	Gumuz	3
		Abergelle	3
		Hararghe Highland	1
ET166	2	Gondar	1
		Abergelle	1
ET213	2	Gondar	1
		Agew	1

Appendix Table D. Distribution of monomorphic and polymorphic loci across chromosomes with respect to goat populations

	AGW		**ARB**		**ABE**		**AFR**		**WGJ**		**NBN**		**BAR**		**AMB**	
Chr	M	P	M	P	M	P	M	P	M	P	M	P	M	P	M	P
1	49	3065	46	3068	38	3076	19	3095	42	3072	18	3096	141	2973	38	3076
2	56	2652	44	2664	61	2647	22	2686	45	2663	13	2695	150	2558	40	2668
3	35	2222	35	2222	36	2221	16	2241	33	2224	17	2240	105	2152	25	2223
4	45	2269	36	2280	44	2272	16	2298	40	2274	13	2301	121	2193	32	2282
5	41	2108	43	2108	47	2104	24	2125	45	2104	15	2134	86	2063	45	2104
6	49	2268	47	2270	51	2266	36	2281	48	2269	16	2301	111	2208	41	2276
7	39	2063	31	2071	43	2059	17	2085	38	2064	13	2089	107	1995	36	2066
8	32	2221	46	2207	40	2213	14	2239	43	2210	19	2234	119	2134	37	2216
9	32	1795	29	1798	25	1802	18	1809	23	1804	9	1818	71	1756	29	1798
10	39	1963	24	1978	42	1960	18	1984	43	1959	16	1986	97	1905	32	1970
11	42	2010	33	2019	28	2024	16	2036	46	2006	16	2036	108	1944	28	2024
12	38	1630	49	1619	39	1629	12	1656	38	1630	12	1656	92	1576	35	1633
13	32	1548	27	1553	27	1553	16	1564	30	1550	14	1566	90	1490	22	1558
14	32	1548	27	1553	27	1553	16	1564	30	1550	14	1566	90	1490	22	1558
15	28	1520	28	1520	20	1528	10	1538	25	1523	11	1537	75	1473	25	1523
16	19	1509	12	1516	19	1509	6	1522	17	1511	3	1525	56	1472	17	1511
17	26	1385	22	1389	25	1386	8	1403	24	1387	5	1406	65	1346	18	1393
18	31	1167	31	1167	21	1177	16	1182	23	1175	7	1191	60	1138	28	1170
19	19	1134	17	1136	20	1133	7	1146	23	1130	6	1147	64	1089	17	1136
20	32	1400	29	1403	30	1402	17	1415	28	1404	12	1420	63	1369	26	1406
21	23	1351	16	1358	19	1355	14	1360	16	1358	7	1367	71	1303	17	1357
22	23	1103	25	1101	24	1102	10	1116	27	1099	7	1119	49	1077	21	1105
23	15	984	15	984	22	977	10	989	12	987	8	991	45	954	14	985
24	21	1241	16	1246	15	1247	11	1251	17	1245	7	1255	56	1206	16	1246
25	22	790	19	793	22	790	10	802	20	792	10	802	50	762	21	791
26	24	975	16	983	15	984	7	992	19	980	4	995	39	960	18	981
27	12	882	11	883	21	873	5	889	15	879	2	892	20	874	11	883
28	14	873	10	877	10	877	4	883	14	873	4	883	31	856	10	877
29	29	897	17	909	24	902	7	919	15	911	6	920	47	879	21	905
µ	**31.0**	**1606.0**	**27.6**	**1609.5**	**29.5**	**1607.6**	**13.9**	**1623.1**	**28.9**	**1608.0**	**10.5**	**1626.5**	**78.6**	**1558.4**	**25.6**	**1611.1**
X	85	1588	73	1600	98	1575	37	1636	109	1564	32	1641	175	1498	94	1579

Appendix Table D. Continued ….

Chr	GON		LES		HGH		KFF		GMZ		SES		HBD		CSH		Average	
	M	P	M	P	M	P	M	P	M	P	M	P	M	P	M	P	M	P
1	63	3051	52	3062	29	3085	56	3058	68	3046	39	3075	24	3090	100	3014	**51.4**	**3062.6**
2	56	2652	51	2657	19	2689	75	2633	54	2654	38	2670	53	2655	129	2579	**56.6**	**2651.4**
3	36	2221	32	2225	17	2240	40	2217	44	2213	30	2227	29	2228	104	2155	**39.6**	**2216.9**
4	41	2273	42	2272	16	2300	66	2248	49	2265	25	2289	24	2290	79	2235	**43.1**	**2271.3**
5	56	2093	52	2097	21	2128	60	2089	43	2106	33	2116	31	2118	113	2036	**47.2**	**2102.1**
6	54	2263	38	2279	36	2281	69	2248	59	2258	43	2274	44	2273	129	2188	**54.4**	**2262.7**
7	44	2058	39	2063	23	2079	60	2042	55	2047	33	2069	42	2060	102	2000	**45.1**	**2056.9**
8	37	2216	35	2218	23	2230	70	2183	48	2205	31	2222	29	2224	82	2171	**44.1**	**2208.9**
9	26	1801	33	1794	14	1813	45	1782	48	1779	27	1800	41	1786	93	1734	**35.2**	**1791.8**
10	31	1971	36	1966	21	1981	60	1942	38	1964	38	1964	44	1958	98	1904	**42.3**	**1959.7**
11	30	2022	41	2011	21	2031	54	1998	37	2025	26	2026	42	2010	95	1957	**41.4**	**2011.2**
12	47	1621	36	1632	14	1654	54	1614	36	1632	29	1639	25	1643	74	1594	**39.4**	**1628.6**
13	31	1549	29	1551	19	1561	55	1525	40	1540	25	1555	26	1554	77	1503	**35.0**	**1545.0**
14	31	1549	29	1551	19	1561	55	1525	40	1540	25	1555	26	1554	77	1503	**35.0**	**1545.0**
15	22	1526	30	1518	13	1535	39	1509	37	1511	25	1523	19	1529	71	1477	**29.9**	**1518.1**
16	17	1511	25	1503	8	1520	33	1495	27	1501	19	1509	19	1509	62	1466	**22.4**	**1505.6**
17	27	1384	25	1386	15	1396	43	1368	31	1380	24	1387	17	1394	51	1360	**26.6**	**1384.4**
18	28	1170	27	1171	16	1182	42	1156	31	1167	24	1174	21	1177	88	1110	**30.9**	**1167.1**
19	13	1140	19	1134	6	1147	35	1118	25	1128	20	1133	21	1132	47	1106	**22.4**	**1130.6**
20	27	1405	30	1402	21	1411	50	1382	35	1397	16	1416	21	1411	54	1378	**30.7**	**1401.3**
21	16	1358	16	1358	11	1363	32	1353	23	1351	18	1356	25	1349	64	1309	**24.3**	**1350.4**
22	18	1108	21	1105	10	1116	31	1095	27	1099	16	1110	16	1110	56	1070	**23.8**	**1102.2**
23	13	986	17	982	8	991	17	982	17	982	17	982	14	985	37	962	**17.6**	**981.4**
24	23	1239	20	1242	6	1256	32	1230	22	1240	11	1251	24	1238	62	1200	**22.4**	**1239.6**
25	26	786	22	790	10	802	34	778	29	783	18	794	19	793	55	757	**24.2**	**787.8**
26	22	977	10	989	8	991	24	975	22	977	11	988	20	979	61	938	**20.0**	**979.0**
27	14	880	16	878	7	887	18	876	18	876	11	883	12	882	33	861	**14.1**	**879.9**
28	13	874	12	875	7	880	11	876	14	873	7	880	17	870	52	835	**14.4**	**872.6**
29	19	907	15	911	7	919	25	901	25	901	7	919	18	908	39	887	**20.1**	**905.9**
μ	**30.4**	**1606.6**	**29.3**	**1607.7**	**15.3**	**1621.7**	**44.3**	**1593.0**	**35.9**	**1601.4**	**23.7**	**1613.3**	**26.3**	**1610.7**	**75.3**	**1561.7**		
X	79	1594	92	1581	45	1628	135	1538	90	1583	68	1605	65	1608	108	1565	86.6	1586.4

Key: ABO=Ambo, GON=Gondar, LES=Long eared Somali, HGH=Hararghe highland, KFF= Kaffa, GMZ= Gumuz, SES= Small eared Somali, HBD= Ibex-Cashmere hybrid, CSH= Cashmere; Ibex= European Ibex

Appendix Table E. Monomorphic (below diagonal) and polymorphic (above diagonal) share by population matrix

Pop$_i$	N1	AGW	ARB	ABE	AFA	WOG	NBN	BAR	ABO	GON	LES	HGE	KAF	GMZ	SES	Hybrid	CASH	Ibex	N2
AGW	1885		47343	47530	47619	46526	47701	45795	47445	47340	47196	47616	47265	47210	47372	47356	46017	665	47911
	%		99.02	99.41	99.60	97.31	99.77	95.78	99.23	99.01	98.71	99.59	98.86	98.74	99.08	99.05	96.25	1.39	%
ARB	1717	1150		47637	47786	46648	47847	45900	47488	47406	47351	47787	47400	47180	47524	47521	46165	665	48079
	%	66.98		99.1	99.4	97.0	99.52	95.47	98.77	98.60	98.49	99.39	98.59	98.13	98.85	98.84	96.02	1.38	%
ABE	1834	1087	1026		48004	46768	48060	46068	47703	47681	47539	47985	47522	47380	47692	47766	46385	670	47961
	%	59.27	55.94		100.1	97.51	100.21	96.05	99.46	99.42	99.12	100.05	99.08	98.79	99.44	99.59	96.71	1.40	%
AFA	1059	768	767	735		46938	48382	46241	47824	47690	47776	48269	47677	47487	47984	48145	46772	674	48737
	%	72.52	72.43	69.41		96.3	99.27	94.88	98.13	97.85	98.03	99.04	97.83	97.44	98.45	98.79	95.97	1.38	%
WOG	1787	1246	1200	1070	832		47002	45231	46653	46574	46665	46922	46635	46386	46749	46642	45335	654	48008
	%	69.73	67.15	59.88	46.56		97.90	94.00	97.18	97.01	97.20	97.74	97.14	96.62	97.38	97.15	94.43	1.37	%
NBN	878	669	647	610	524	715		46293	47887	47752	47792	48310	47977	47579	48011	48309	46901	672	48917
	%	76.20	73.69	69.48	59.68	81.44		94.64	97.89	97.62	97.70	98.76	98.08	97.26	98.15	98.76	95.88	1.37	%
BAR	3339	1224	1160	1078	844	1404	715		45926	45859	45821	46185	45804	45662	45978	45972	44717	644	46459
	%	36.66	34.74	32.29	25.28	42.05	21.41		98.85	98.71	98.63	99.41	98.59	98.28	98.96	98.95	96.25	1.39	%
ABO	1688	1223	1098	1064	776	1176	659	1158		47476	47354	47815	47420	47283	47527	47555	46192	668	48107
	%	72.45	65.05	63.03	45.97	69.67	39.04	68.60		98.69	98.43	99.39	98.57	98.29	98.79	98.85	96.02	1.39	%
GON	1812	1242	673	647	766	1221	647	1215	1181		47270	47700	47296	47169	47420	47445	46084	667	47984
	%	68.54	37.14	35.71	42.27	67.38	35.71	67.05	65.18		98.51	99.41	98.57	98.30	98.82	98.88	96.04	1.39	%
LES	1774	1061	1047	985	814	1274	649	1139	1021	1061		47727	47279	47056	47509	47477	46129	663	48022
	%	59.81	59.02	55.52	45.89	71.82	36.58	64.21	57.55	59.81		99.39	98.45	97.99	98.93	98.87	96.06	1.38	%
HGE	1183	889	892	840	716	940	576	912	891	900	879		47667	47472	47935	48039	46650	675	48613
	%	75.15	75.40	71.01	60.52	79.46	48.69	77.09	75.32	76.08	74.30		98.05	97.65	98.61	98.82	95.96	1.39	%
KAF	2357	1173	1140	1012	759	1288	667	1166	1131	1131	1077	873		47168	47443	47427	47445	663	47438
	%	49.77	48.37	42.94	32.20	54.65	28.30	49.47	47.98	47.98	45.69	37.04		99.43	100.01	99.98	100.01	1.40	%
GMZ	2022	1322	1124	1073	772	1243	684	1228	1198	1208	1056	882	1213		47239	47229	45887	661	47774
	%	65.38	55.59	53.07	38.18	61.47	33.83	60.73	59.25	59.74	52.23	43.62	59.99		98.88	98.86	96.05	1.38	%
SES	1514	978	960	878	1514	1098	608	1036	934	951	1002	836	980	980		47722	46356	668	48282
	%	64.60	63.41	57.99	100.00	72.52	40.16	68.43	61.69	62.81	66.18	55.22	64.73	64.73		98.84	96.01	1.38	%
Hybrid	788	234	230	225	196	265	181	304	236	252	244	215	238	244	229		47180	670	49008
	%	29.70	29.19	28.55	24.87	33.63	22.97	38.58	29.95	31.98	30.96	27.28	30.20	30.96	29.06		96.27	1.37	%
CASH	2350	457	437	407	386	520	334	611	435	451	457	387	449	464	425	523		655	47446
	%	19.45	18.60	17.32	16.43	22.13	14.21	26.00	18.51	19.19	19.45	16.47	19.11	19.74	18.09	22.26		1.38	%
IBEX	49111	1866	1698	1453	1049	2600	866	3300	1672	1795	1753	1174	1797	2022	1498	774	2321		684
	%	3.80	3.46	2.96	2.14	5.29	1.76	6.72	3.40	3.65	3.57	2.39	3.66	4.12	3.05	1.58	4.73		

Key: N1= Number of monomorphic loci in each goat population; N2= Number of monomorphic loci in each goat population; AGW=Agew, ARB=Arsi-Bale, ABE=Abergelle, AFA=Afar, WOG=Woyto-Guji, NBN=Nubian, BAR=Barka, ABO=Ambo, GON=Gondar, LES=Long eared Somali, HGE=Hararghe highland, KAF= Kaffa, GMZ= Gumuz, SES= Small eared Somali, Hybrid= Ibex-Cashmere hybrid, CASH= Cashmere; Ibex= European Ibex

Appendix Table F. Proportion of genetic backgrounds/clusters of study goat populations

Populations	Clusters							
	1	2	3	4	5	6	7	8
Nubian	0.4477	0.0260	0.1192	0.0378	0.0071	0.0949	0.0294	0.2379
Barka	0.1383	0.0533	0.4065	0.0321	0.0121	0.0555	0.0281	0.2742
Abergelle	0.0069	0.0220	0.9014	0.0003	0.0133	0.0186	0.0007	0.0367
Gondar	0.0054	0.0769	0.7704	0.0013	0.0335	0.0775	0.0009	0.0342
Ambo	0.0019	0.1667	0.5458	0.0021	0.1273	0.0800	0.0004	0.0758
Agew	0.0017	0.1470	0.5645	0.0003	0.0420	0.2252	0.0003	0.0190
Gumuz	0.0073	0.1046	0.2663	0.0005	0.0229	0.5858	0.0006	0.0121
Arsi-Bale	0.0089	0.2169	0.4096	0.0012	0.0343	0.0328	0.0013	0.2950
Hararghe highland	0.0358	0.1076	0.2812	0.0185	0.0193	0.0274	0.0121	0.4981
Afar	0.0713	0.0025	0.1669	0.0405	0.0104	0.0091	0.0298	0.6695
Small eared Somali	0.0399	0.0901	0.1193	0.0168	0.0087	0.0093	0.0087	0.7072
Long eared Somali	0.0168	0.1230	0.0294	0.0028	0.0049	0.0038	0.0009	0.8185
Woyto-Guji	0.0025	0.3912	0.0186	0.0002	0.0099	0.0151	0.0001	0.5623
Kaffa	0.0025	0.9539	0.0181	0.0002	0.0044	0.0101	0.0001	0.0108
Ibex-Cashmere hybrid	0.0090	0.0007	0.0034	0.1597	0.0012	0.0014	0.8201	0.0045
Cashmere	0.0025	0.0000	0.0005	0.9741	0.0000	0.0003	0.0215	0.0010

Appendix Table G. Primers designed for analysis of KISS1 gene

Region	Name given	Sequence	Lgth	GC%	T_m
Exon1_R1	CH_KISS1_Exon1_F2	5′-TTATGTCACTGCAGCTGG-3′	18	50.0	52.3
	CH_KISS1_Exon1_R1	5'- CTTGCTACTCACTGGCTG -3'	18	55.6	52.9
Exon1_R2	CH_KISS1_Exon1_F1	5'- AGCGCTGAGCTTCCTAG -3′	17	58.8	54.3
	CH_KISS1_Exon1_R3	5′-GGCAATGGTCAGCATCATC-3′	19	52.6	54.1
Exon2	Chi_KISS1_Exon2_F	5'-CACTGTCCCACTGCATCTC-3'	19	57.9	55.5
	Chi_KISS1_Exon2_R	5′-GTAACGGCAGAAGAGCCTC-3′	19	57.9	55.5

Key: lgth=length (bp)

Appendix Table H. Analysis of variance of exon1 and exon2 regions of KISS1 gene

Exon1					Exon2				
Source	**Df**	**Type III SS**	**M. Square**	**F Val.**	**Source**	**Df**	**Type III SS**	**M.Square**	**F Val.**
Parity	4	2.46961285	0.61740321	2.93**	Parity	4	6.83085662	1.70771415	8.85***
Population	1	10.60811916	10.60811916	50.34***	Population	1	14.80968402	14.80968402	76.79***
Genotype at G3416C	1	0.41633801	0.41633801	1.98**	Genotype at T950C	1	0.16954025	0.16954025	0.88**
Genotype at C3811T	2	0.38815068	0.19407534	0.92 *					
Genotype at T3963C	2	0.78901259	0.39450630	1.87**					
Error	166	34.98043349	0.21072550			235	45.32326482	0.19286496	
Total	176	52.28248588				241	69.55371901		

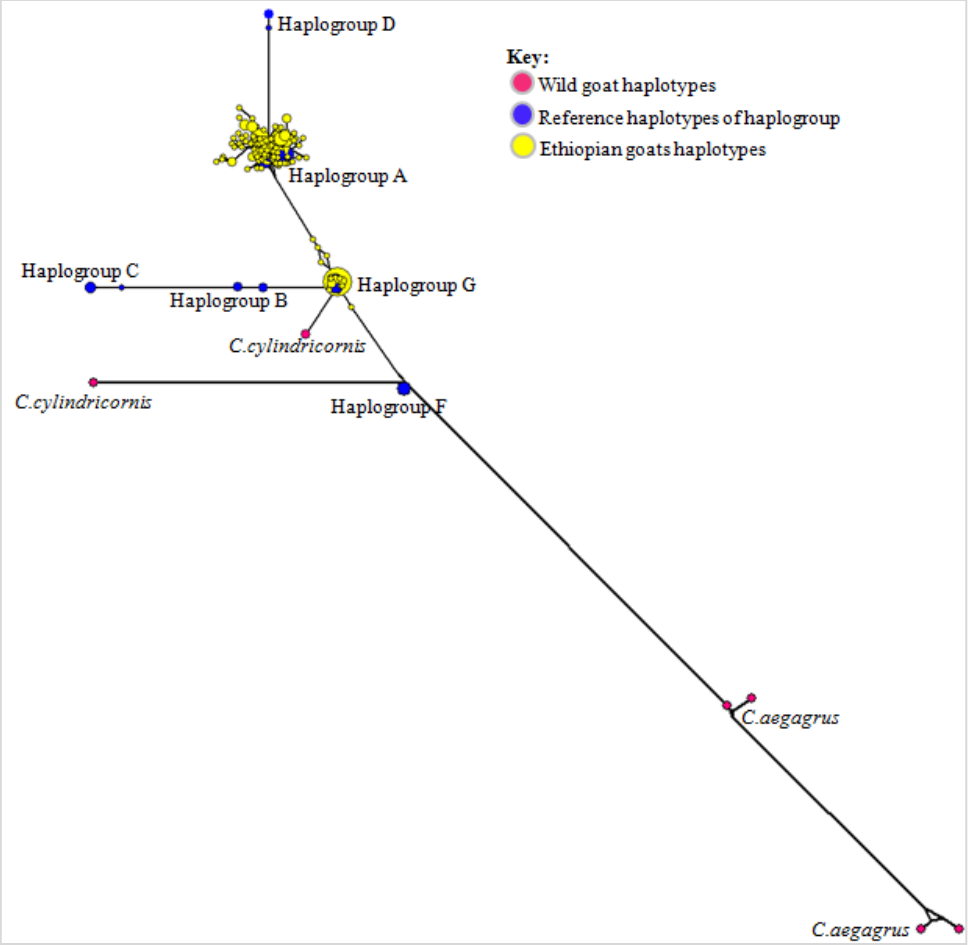
R2= 0.331; CV= 34.87%; **=P<0.01; ***=P<0.001; R2=35% CV=33.01%

Appendix Table I. Summary of linkage disequilibrium measures of exon1 for both goat populations

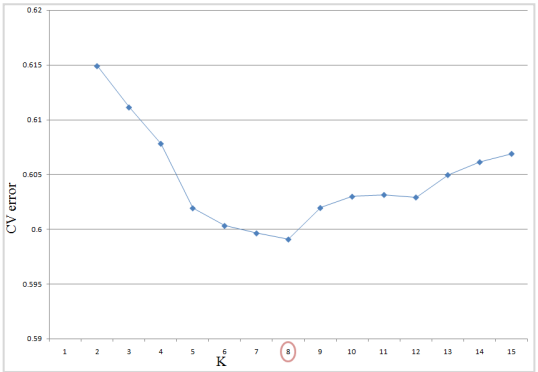
Locus1	Locus2	Dist	D	D'	R	Fisher	Chi-sq	Locus1	Locus2	Dist	D	D'	R	Fisher	Chi-sq
74	136	62	0.000	-1.000	-1.000	1.000	0.049	369	528	159	-0.118	-0.694	-0.492	0.000***B	41.907***B
74	253	179	0.000	-1.000	-0.006	1.000	0.006	369	531	162	0.024	0.316	0.137	0.085	3.230
74	369	295	0.003	1.000	0.080	0.474	1.116	369	647	278	0.000	0.049	0.006	1.000	0.005
74	416	342	0.000	-0.000	-0.006	1.000	0.006	369	683	314	-0.011	-0.096	-0.052	0.595	0.459
74	490	416	0.000	-1.000	-0.006	1.000	0.006	369	709	340	0.174	0.715	0.698	0.000***B	84.390***B
74	503	429	0.000	-1.000	-0.008	1.000	0.012	416	490	74	0.000	-1.000	-0.006	1.000	0.006
74	528	454	0.004	1.000	0.102	0.358	1.801	416	503	87	0.000	-1.000	-0.008	1.000	0.012
74	531	457	-0.001	-1.000	-0.031	1.000	0.170	416	528	112	-0.002	-1.000	-0.057	1.000	0.562
74	647	573	0.000	-1.000	-0.008	1.000	0.012	416	531	115	-0.001	-1.000	-0.031	1.000	0.170
74	683	609	-0.001	-1.000	-0.043	1.000	0.322	416	647	231	0.000	-1.000	-0.008	1.000	0.012
74	709	635	0.003	1.000	0.082	0.462	1.169	416	683	267	-0.001	-1.000	-0.043	1.000	0.322
136	253	117	0.000	-1.000	-0.017	1.000	0.049	416	709	293	-0.003	-1.000	-0.071	1.000	0.865
136	369	233	0.013	0.525	0.122	0.152	2.563	490	503	13	0.000	-1.000	-0.008	1.000	0.012
136	416	280	0.000	-1.000	-0.017	1.000	0.049	490	528	38	-0.002	-1.000	-0.057	1.000	0.562
136	490	354	0.000	-1.000	-1.017	1.000	0.049	490	531	41	-0.001	-1.000	-0.031	1.000	0.170
136	503	367	-0.001	-1.000	-0.024	1.000	0.098	490	647	157	0.000	-1.000	-0.008	1.000	0.012
136	528	392	-0.017	-1.000	-0.165	0.052	4.685*	490	683	193	-0.001	-1.000	-0.043	1.000	0.322
136	531	395	-0.007	-1.000	-0.090	0.370	1.417	490	709	219	0.003	1.000	0.082	0.462	1.169
136	647	511	-0.001	-1.000	-0.024	1.000	0.098	503	528	25	0.002	0.221	0.032	1.000	0.176
136	683	547	-0.011	-1.000	-0.125	0.201	2.689	503	531	28	-0.002	-1.000	-0.044	1.000	0.342
136	709	573	0.013	0.535	0.127	0.146	2.790	503	647	144	0.000	-1.000	-0.012	1.000	0.024
253	369	116	-0.003	-1.000	-0.072	1.000	0.906	503	683	180	-0.003	-1.000	-0.061	1.000	0.649
253	416	163	0.000	-1.000	-0.006	1.000	0.006	503	709	206	-0.005	-1.000	-0.100	0.500	1.741
253	490	237	0.000	-1.000	-0.006	1.000	0.006	528	531	3	0.018	0.190	0.104	0.183	1.880
253	503	250	0.000	-1.000	-0.008	1.000	0.012	528	647	119	0.002	0.221	0.032	1.000	0.176
253	528	272	0.004	1.000	0.102	0.358	1.801	528	683	155	0.034	0.221	0.167	0.041*	4.838*
253	531	278	-0.001	-1.000	-0.031	1.000	0.170	528	709	181	-0.108	-0.651	-0.451	0.000***B	35.250***B
253	647	394	0.000	-1.000	-0.008	1.000	0.012	531	647	116	-0.002	-1.000	-0.044	1.000	0.342
253	683	430	-0.001	-1.000	-0.043	1.000	0.322	531	683	152	0.104	0.947	0.688	0.000***B	81.773***B
253	709	456	-0.003	-1.000	-0.071	1.000	0.865	531	709	178	0.026	0.330	0.146	0.081	3.707
369	416	47	-0.003	-1.000	-0.072	1.000	0906	647	683	36	-0.003	-1.000	-0.061	1.000	0.649
369	490	121	0.003	1.000	0.080	0.474	1.116	647	709	62	-0.005	-1.000	-0.100	0.500	1.741
369	503	134	0.006	1.000	0.114	0.223	2.245	683	709	26	0.044	0.336	0.205	0.008**	7.264

10.2 Appendix figure

Appendix Figure A. Phylogenetic network analysis of domestic and wild goats



Appendix Figure B. Graph of cross validation error



Appendix Figure C. PCR conditions for exon1 and 2 regions of KISS1 gene amplification

a. Exon2

PCR program:

• 95°C	5 minutes	
• 94°C	15 seconds	5 cycle
• 64°C - 60°C	30 seconds	
• 72 °C	45 seconds	
• 94°C	15 seconds	35 cycle
• 58°C	45 seconds	
• 72°C	1 minute	
• 72°C	10 minutes	
• 4°C	∞	

PCR reaction:

Primer (10pM/µl) F	0.4µl
Primer (10pM/µl) R	0.4µl
BSA	1.0µl
Hi-Di	0.5µl
H_2O (Nuclease free)	17.3µl
Template	0.4µl

b. Exon1_R1

PCR program:

• 95°C	5 minutes	
• 94°C	15 seconds	5
• 62-58°C	30 seconds	
• 72 °C	45 seconds	
• 94°C	15 seconds	35
• 56°C	45 seconds	
• 72°C	1 minute	
• 72°C	10 minutes	
• 4°C	∞	

PCR reaction:

Primer (10pM/µl) F	0.4µl
Primer (10pM/µl) R	0.5µl
BSA (100X: 10mg/ml)	2.0µl
Hi-Di	0.5µl
H_2O (Nuclease free)	16.3µl
Template	0.4µl

c. Exon1_R2

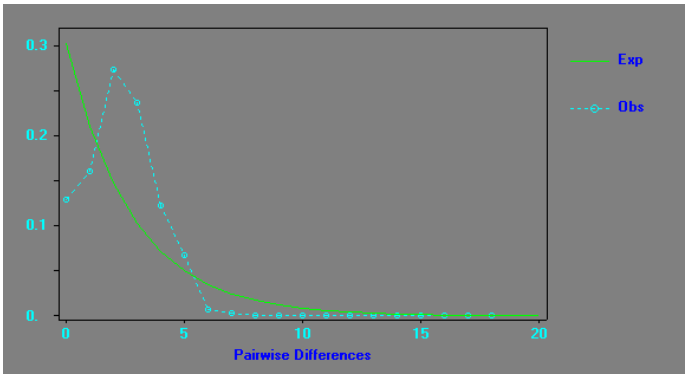
PCR program:

• 95°C	5 minutes	
• 94°C	15 seconds	5
• 62-58°C	30 seconds	
• 72 °C	1.0 minute	
• 94°C	15 seconds	35
• 56°C	1.0 minute	
• 72°C	1.0 minute	
• 72°C	10 minutes	
• 4°C	∞	

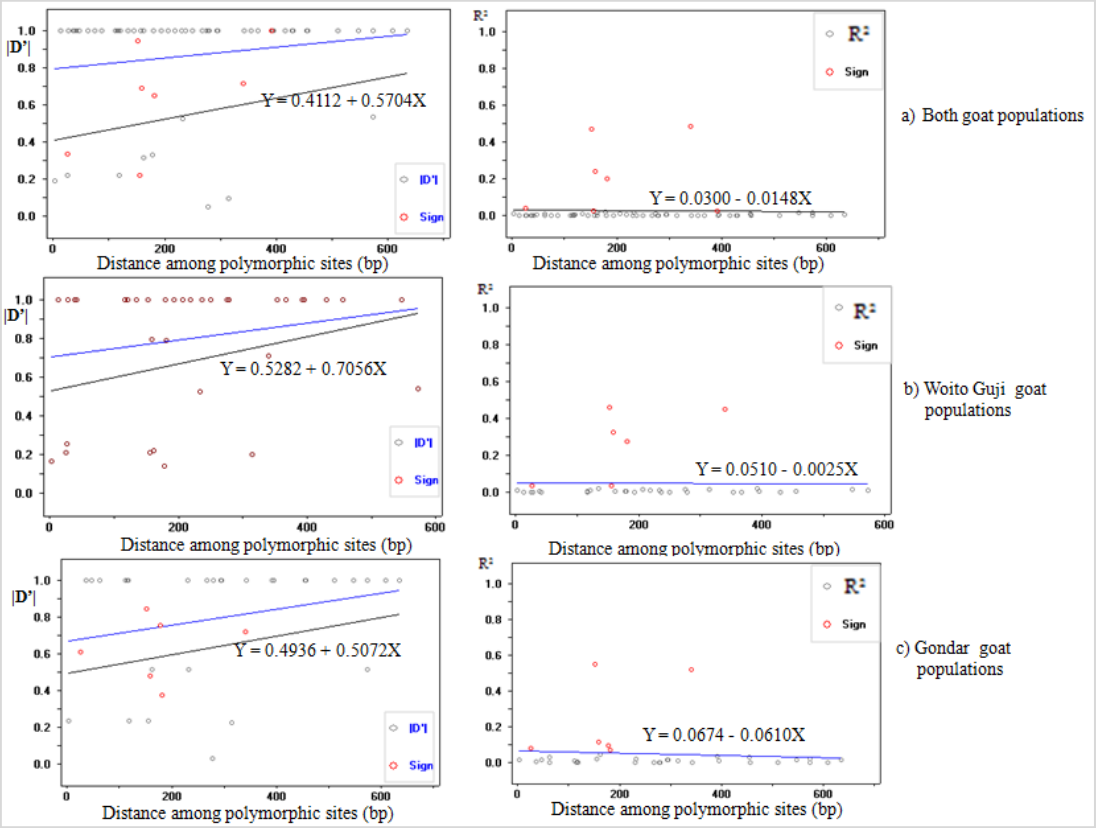
PCR reaction:

Primer (10pM/µl) F	0.4µl
Primer (10pM/µl) R	0.4µl
BSA (100X: 10mg/ml)	2.0µl
Hi-Di	0.5µl
H_2O (Nuclease free)	16.3µl
Template	0.6µl

Appendix Figure D. Trend of observed and expected haplotype heterozygosities



Appendix Figure E. Measures of LD decays ($|D'|$ and R^2) with respect to distances among segregating/polymorphic sites



11 DECLARATION

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any other university and all sources of material used for the thesis have been duly acknowledged.

Name: Getinet Mekuriaw Tarekegn **Signature**: 

Place: College of Natural Sciences, Addis Ababa University

Date: May 24, 2016

This thesis has been submitted for examination with my approval as a university advisor.

Kassahun Yesfaye (PhD):  **Date**: May 24, 2016

Name of your thesis advisor