ADDIS ABABA UNIVERSITY

INSTITUTE OF BIOTECHNOLOGY

Genome-Wide Signature of Positive Selection, Breed-specific SNPs and Linkage Disequilibrium in Ethiopian Indigenous and European Beef Cattle Breeds

MSc Thesis

By;

Dejenie Mengistie Yirsaw

Addis Ababa, Ethiopia

September 2021

Genome-wide Signature of Positive Selection, Breed-specific SNPs and Linkage Disequilibrium in Ethiopian Indigenous and European Beef Cattle Breeds

A Thesis Submitted to Institute of Biotechnology, Addis Ababa University in Partial
Fulfillment of the Requirements for the degree of Master of Science in Biotechnology

By;

Dejenie Mengistie Yirsaw

Addis Ababa, Ethiopia

September 2021

Thesis Approval Sheet

We certify that Dejenie Mengistie Yirsaw MSc thesis research entitled “Genome-Wide Signature of Positive Selection, Breed-specific SNPs and Linkage Disequilibrium in Ethiopian Indigenous and European Beef Cattle Breeds” has been conducted under our direct supervision. Therefore, we kindly request the Institute of Biotechnology, Addis Ababa University to officially approve the thesis.

Supervisors:	Signature:	Date:
Tesfaye Sisay (PhD, Prof.)	________________	______________
Hailu Dadi (PhD)	________________	______________
Zewdu Edea (PhD)	________________	______________

We, undersigned, members of the boards of Examiners of the final open defense by Dejenie Mengistie Yirsaw have read and evaluated his thesis entitled “Genome-wide Signature of Positive Selection, Breed-specific SNPs and Linkage Disequilibrium in Ethiopian Indigenous and European Beef Cattle Breeds. This is therefore to certify that the thesis has been accepted in partial fulfillment of the requirements for the Degree Master of Science in Biotechnology.

Tileye Feyissa (PhD)	________________	______________
Name of Examiner	Signature	Date
Getint Mekuriaw (PhD)	________________	______________
Name of Examiner	Signature	Date

Tesfaye Sisay (PhD, Prof.)	________________	______________
Director, Institute of Biotechnology	Signature	Date

DEDICATION

I am honored to dedicate this thesis research to my late Mother, ENATNESH ATALAY, for upbringing me with love and affection. Her selflessness will be remembered!

STATEMENT OF AUTHOR

First and foremost, I declare that this thesis is my original work and that all sources of information used in this thesis have been properly acknowledged. This thesis was submitted as part of the requirements of a Master of Science degree in Biotechnology at Addis Ababa University, and it has been placed in the university library to be made available for the reader. I hereby certify that I am not submitting this thesis to any other school to receive an academic degree, diploma, or certificate. Short quotes from this thesis are permitted without specific permission as long as the source is properly acknowledged.

Dejenie Mengistie Yirsaw

September 2021

Acknowledgments

“No work is done by a man alone

The path is visible when a path is shown”

Writing the acknowledgment signals my approaching the next milestone of my academic journey. Over feelings with gratitude and a sense of veneration, I would like to thank soulfully the “Almighty God” who bestowed me with his blessing and benevolence at every step without which nothing could have been accomplished. I would like to thank St. Virgin Mary (ቅድስት ድንግል ማርያም) for helping me to live in God’s grace, guidance, protection, and wisdom.

Gratitude is expressed to members of my advisory committee, Professor TESFAYE SISAY, DR. HAILU DADI, and DR. ZEWDU EDEA for their excellent guidance and mentorship which were indispensable for the fruitful execution of this work. They were always available tirelessly for open discussions and pertinent scientific questions without which this thesis would not have been possible. It gives me a great honor to thank NATIONAL AGRICULTURAL BIOTECHNOLOGY RESEARCH CENTER for sponsoring my graduate study. I convey special thanks to CHUNGBUK NATIONAL UNIVERSITY and INTERNATIONAL LIVESTOCK RESEARCH INSTITUTE for providing me the SNP data used for this study. I would like to thank ETHIOPIAN BIOTECHNOLOGY INSTITUTE for hosting me as a graduate fellow. I thank Addis ABABA UNIVERSITY for admitting me to the Biotechnology Graduate Program. I am also indebted to the assistance conveyed by many research personnel at the Ethiopian Institute of Agricultural Research.

Finally, I would like to extend my heartfelt appreciation to so many friends and colleagues who in one way or the other provided their unreserved support.

Table of Contents

List of Tables

List of Figures

List of abbreviations

BP	Biological Process
CLR	Composite Likelihood Ratio
CSA	Central Statistical Agency
DAVID	Database for Annotation, Visualization and Integrated Discovery
FST	Fixation Index
GEBV	Genomic Estimated Breeding Values
GS	Genomic Selection
GWAS	Genome-Wide Association Study
GO	Gene Ontology
GGP	Gene Seeks Genomic Profiling
HWE	Hardy Weinberg Equilibrium
HD	High Density
HS	Subpopulations Heterozygosity
HT	Heterozygosity Total
IHS	Integrated Haplotype Score
KEGG	Kyoto Encyclopedia of Genes and Genomes
LD	Linkage Disequilibrium
MAF	Minor Allele Frequency

MAS	Marker-Assisted Selection
IMF	Meat Intramuscular Fat
MSTN	Myostatin
mtDNA	Mitochondrial DNA
NGS	Next-Generation Sequencing
NJ	Neighbor-Joining
PIC	Polymorphic Information Content
PLAG1	Pleomorphic Adenoma Gene 1
QTL	Quantitative Trait Locus
RFLP	Restriction Fragment Length Polymorphism
SSR	Simple Sequence Repeats
ROH	Runs of Homozygosity
SNP	Single Nucleotide Polymorphism

Abstract

Title: Genome-Wide Signature of Positive Selection, Breed-specific SNPs and Linkage Disequilibrium in Ethiopian Indigenous and European Beef Cattle Breeds

By: Dejenie Mengistie

September 2021

Addis Ababa University

The detection of selection signatures in Ethiopian indigenous breeds can contribute to the identification of regions of the genome that are, or have been, functionally important and, as a consequence, have been targeted by selection. Despite the availability of genomic methods for determining the origin and divergence of domestic cattle in East Africa, particularly Ethiopia, knowledge regarding their genetic adaptability and divergence remain limited. To investigate signatures of selection, breed-specific SNPs, and linkage disequilibrium (LD), Ethiopian indigenous cattle populations, three Ethiopian cattle populations (Begait n = 40; Boran n = 40; and Fogera n = 43; were genotyped with a high-density 80K SNP array. Three European beef cattle breeds (Angus, n = 42), Hereford (n = 35), and Charolais (n = 37) were also used for comparison purposes. The average minor allele frequency was 0.19 ± 0.17, 0.20 ± 0.17, 0.21± 0.17, 0.31 ± 0.13, 0.32 ± 0.12, 0.32 ± 0.13 for Angus, Herford, Charolais, Boran, Fogera, and Begait cattle, respectively. Across the Ethiopian and European cattle populations, a common variant minor allele frequency ($\geq$0.10 and $\leq$ 0.5) accounted for 94% and 62%, of the total SNPs, respectively. A total of 7759 and 48 SNPs were identified as breed-specific in the Ethiopian cattle populations and European beef cattle, respectively. These SNPs were associated with 3364 and 17 genes, respectively. Across the sampled Ethiopian cattle populations, the mean observed and expected heterozygosity were 0.403 and 0.400, respectively. Similarly, for European cattle, the observed and expected heterozygosity were 0.25 and 0.26 respectively. Principle component analysis (PCA), and neighbor-joining (NJ)-tree revealed a separation of

Ethiopian cattle breeds from European beef breeds. Similarly, the NJ-tree grouped the study cattle according to their breed group with close clustering of the three Ethiopian cattle populations. The average r2 values were 0.22 ± 0.25, 0.23 ± 0.25, and 0.22 ± 0.25 in Boran, Begait, and Fogera, respectively. For Angus, Herford and Charolais it was 0.17 ± 0.28, 0.17 ± 0.28 and 0.18 ± 0.29, respectively. It was observed that average r2 values decreased as physical distance increased. The top 1% FST values were considered to delimit genomic regions under positive selection. Some of the candidate genes are involved in biological processes and pathways linked to meat quality attributes. The candidate genes are involved in meat tenderness (*CAPZB, DCTN2, KIF5A, DCTN2*, *RGS20*, *STIM1*, *KLHL3*, and *ITGB1BP1*), IMF (*DGAT1*, *LPCAT1*, *PLCL2*, *CTNNA1*, and *B4GALNT1*) meat color (*ASAP3*, *ADAP1*, *APBB1IP*, *RTKN2*, and *STAP2*). Furthermore, some of the candidate genes are associated with tropical adaptation — heat tolerance, and resistance to disease (*SOD1*, *DCTN2*, *GNAS, IGF1*, *TLN1*, *RRAS, GNAS,* and *ASAP3*). The significant degree of SNP variation found in this study suggests that these markers could be useful for genetic research in Ethiopian cattle breeds. The high genetic variability observed in Ethiopian cattle breeds can be used as a potential source for future breeding programs. The identified candidate genes have to be further confirmed through GWAS, sequencing, and omics studies.

Key words; ***Adaptation, Section pressure, Selection signature***

CHAPTER ONE

1. INTRODUCTION

In several parts of the developing countries, livestock production is one of the major means of achieving improved living standards. Livestock plays a critical role in both the national economies and the livelihoods of rural communities in sub-Saharan African countries (Getachew, 2010). In Africa, Ethiopia has the largest cattle population (70 million) annual livestock survey (2020/21). According to the Ethiopian Biodiversity Institute, 2016 report, there are more than 28 native cattle breeds/ecotypes that have been recognized to exist in the country. These populations can be further classified into four groups: zebu (*B. indicus*), sanga (*zebu* × *B. taurus*), zenga (*sanga* × *zebu*), and humpless *B. taurus* (Rege, 1999). Many of them are named after the community maintaining the population or geographical location where they predominate, the true genetic relationship between the main populations has not yet been well defined or documented.

The productivity of indigenous cattle populations is influenced by several factors such as poor genetics, shortage of feed, water, abundant health problems, and a poor housing system that result in low production and low reproductive performance (Ayalew *et al*., 2018).

The determination of the allelic distribution of markers associated with economically significant traits can be a powerful tool for acquiring immediate knowledge for the selection of superior animals and to make successful early decisions (Shor-shimoni *et al*., 2017). Identification of functional variations such as missense variants and variants in indigenous cattle within downstream and upstream genic regions would allow these variants to be identified for their effects on complex traits (Zwane *et al*., 2019). In the absence of phenotypic data, comparisons of breeds that have been subjected to various selective pressures may aid in the identification of genomic areas and genes that regulate qualitative and complicated traits (Stafuzza *et al*., 2017; Van Den Berg *et al.*, 2020).

Therefore, the investigation of differences between and within cattle breed is an important initial guide for promoting the best use of genetic resources for farm animals and allows successful genetic enhancement to meet the needs of production strategies and to plan and integrate enhancement programs in the context of the unique efficiency of a population (Hu *et al.*, 2018).

In indigenous cattle, mapping and identification of candidate genes associated with economically significant traits help to attain rapid genetic gain in the beef industry (Moreira *et al.*, 2019). Single nucleotide polymorphism (SNP) markers analysis has become the standard approach in recent years for genome-wide association studies and selection signatures analysis. The commercial availability of a large range of genome-wide SNP panels offers the opportunity to fine map genes impacting complex quantitative traits (Stranger *et al*., 2011).

Therefore, this study was designed to investigate selection signatures likely associated with beef and ecological traits, breed-specific SNPs, and linkage disequilibrium (LD) in Ethiopian indigenous and European beef cattle breeds.

1.2. Objectives

1.2.1. General objective

To identify selection signatures and breed-specific SNPs, and estimate linkage disequilibrium in Ethiopian and European beef cattle breeds using genome-wide SNPs

1.2.2. Specific objectives

- To identify signatures of positive selection in Ethiopian and European beef cattle breeds exposed to diverse selective pressures
- To identify breed-specific SNPs in Ethiopian indigenous cattle and European beef cattle breeds
- To estimate linkage disequilibrium in Ethiopian indigenous and European beef cattle breeds

CHAPTER TWO

2. LITERATURE REVIEW

2.1. Origins and domestication of cattle

More than 200,000 years ago, cattle, also known as *B. taurus* (humpless) or *B. indicus* (humped), were split into two distinct lineages (Daetwyler *et al*., 2014). Since domestication, cattle have been essential in the transformation of human society from nomadic hunter-gatherers to sedentary farming societies across much of Europe, Asia, and Africa. Although our knowledge of ancestral population relationships is poor, cattle domestication is thought to have occurred on two or three occasions, giving rise to the taurine (*B. taurus*) and indicine (*B. indicus*) species, which share a common ancestor with the aurochs (*B. primigenius*) 250,000 years ago (Pitt *et al*., 2019).

Around 10,000 years ago, according to archaeological and genomic data the ancestors of taurine cattle (*B. taurus*) were domesticated from *B. primigenius* in the Fertile Crescent during the Neolithic period (MacHugh *et al*., 2017). However, 1,500 years later, in the Indus Valley, a second domestication event occurred from *B. primigenius* nomadicus, which was isolated from the taurine branch between 250 years ago, eventually giving rise to the extant indicine cattle (*B. indicus*), also known as zebu cattle (Pitt *et al*., 2019).

The first short horns cattle appeared in Mesopotamia around 3000 years BC. This phenotype was more appropriate for those habitats and substituted for most long-horn forms in Asia and neighboring continents by the second wave of migrations (Pitt *et al*., 2019). The distribution of cattle in various regions of the world has resulted in the creation of many ecotypes adapted to their local environments (Theunissen and Lenstra, 2015). Several different "agrotypes" were developed by human selection, which preceded the development of breeds that differ in coat color, horn, and docility (Mwai *et al.*, 2015).

In the last 200 years, through a systematic selection of isolated populations that became the new breeds, cattle diversity has been increased. Many cattle acquired, for example, large udders when dairy and beef production began and this type of domestication process resulted in a decrease in the size of the breed (Theunissen and Lenstra, 2015, Pitt *et al.*, 2019).

Cattle have been migrated to Africa from around 4000-5000 years BC (Mwai *et al*., 2015). *B. taurus* is humpless and includes two classes of shorthorns and longhorns that are found primarily in Central and West Africa. *B. indicus* is humped and in Africa, they are essential types of cattle. *B. indicus* cattle predominate in western and eastern parts of Africa (Gifford-Gonzalez and Hanotte, 2011). Zebu breeds have been made more attractive to local farmers by their substantial body mass and higher productivity in tsetse-free areas, which somewhat explains the proliferation of these breeds and wide distribution across Africa (Tarekegn *et al.*, 2018).

On the African continent, there is no pure *B. indicus*, as all cattle bear taurine mitochondrial DNA (Mwai *et al*., 2015). According to mtDNA sequences from Europe, Africa, and India, both African and European cattle are in one large taurine cluster, while all Indian breeds are in the other. Only taurine mitochondrial genomes were discovered in African breeds, with no evidence of female zebu imports. A taurine Y-chromosome was maintained by most sangas, suggesting that male zebu introgression in these cattle was only partial (Li *et al*., 2013, Koolmees and Lenstra, 2014).

In Sub-Saharan Africa, 180 cattle breeds have been recognized, including 150 indigenous cattle breeds and newly introduced exotic and commercial composites (Mwai *et al*., 2015). However, since the genetic differentiation between these cattle breeds is often unclear, it may be more fitting to refer to African cattle populations or ecotypes.

The indigenous cattle of Africa include different crosses between longhorn cattle (*B. taurus*), and zebu cattle (*B. indicus*). Crossbreeding in East Africa led to the creation of cervico-thoracically humped sanga cattle which spread southward and reached most areas of the south-east part of Africa during 250-500 AD (Koolmees and Lenstra, 2014; Gifford-Gonzalez and Hanotte, 2011). New cattle breeds (zebu and sanga-types) were established as the result of the migration of humans across Africa (Angress and Reed, 2014).

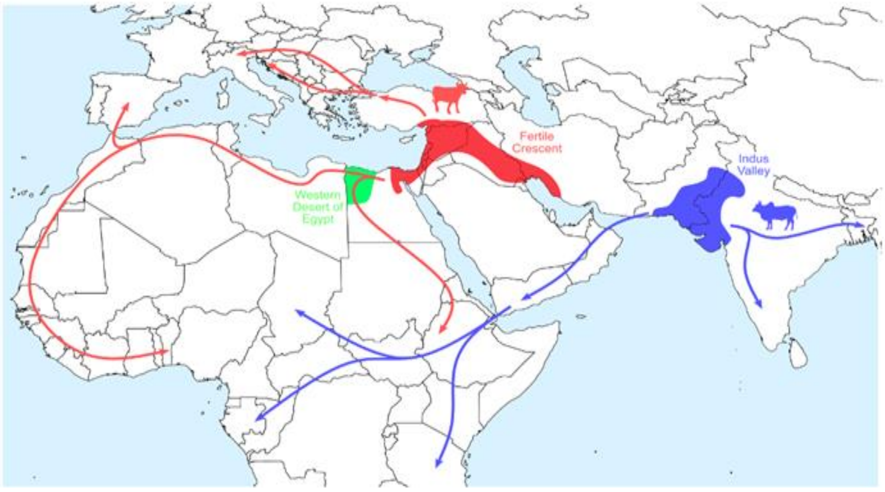


Figure 1: Main domestication sites and migration routes of taurine (*B.taurus*) and indicine (*B. indicus*) cattle (Pitt *et al*., 2019)

East Africa including Ethiopia is considered the cradle of the Near-East *B. taurus* as well as the Arabian and Indian *B. indicus* cattle migration corridors and is often referred to as the secondary hybridization zone. The country is home to over 28 cattle breeds or populations, which can be grouped into four categories: zebu (*B. indicus*), sanga (*zebu* × *B. taurus*), zenga (*sanga* × *zebu*), and the humpless (*B. taurus)* (Rege, 1999).

Three key lines of reasoning were used to justify this study: To begin with, Ethiopia has one of the most diverse populations of indigenous cattle in the world. Second, it is strategically located near the Horn and the East Coast, which serve as cattle entry points into Africa. Finally, microsatellite data analysis showed Ethiopian cattle to be hybrid populations (Dadi *et al.*, 2008). All Ethiopian cattle mtDNA sequences converged on one maternal lineage (T1), which corresponds to African *B. taurus* cattle; no zebu mtDNA haplotypes have been discovered in the Ethiopian population (Tarekegn *et al.*, 2018).

2.2. Molecular markers

2.2.1. Genetic markers

Genetic markers are used to differentiate individuals and populations from each other. Initially karyotyping, is a procedure that assesses the number and shape of chromosomes to identify abnormalities. Some African zebu cattle were characterized by karyotyping (Mwai *et al.*, 2015). Other markers, such as protein electrophoresis, have been used to study the population structures of different cattle populations. Both of these genetic marker groups display low variance to perform population genetic analysis effectively (Mwai *et al.*, 2015).

Restriction fragment length polymorphisms (RFLP) were among the first genetic markers used in animal genetics and nowadays RFLP is applied for candidate genes. The advent of the polymerase chain reaction (PCR) in 1983 allowed DNA fragment amplification to be used as genetic markers for sequence variations in various types of DNA. In the 1990s, laborious methods for detecting RFLPs were replaced by microsatellites.

In cattle population studies, microsatellites, tandem repeats of very brief (one to six base pair) nucleotide motifs, are commonly used genomic markers. They have been used to define the evolutionary relationships between cattle subspecies, population levels of genomic admixture, migration history, as well as to map genomic quantitative trait loci (QTL) within species, due to the high level of polymorphisms usually observed at microsatellite loci (Hanotte *et al.*, 2003). Although microsatellite markers have shown great success in enhancing our understanding of cattle population structure and history, their relatively restricted bovine genome coverage is a downside. It has opened new avenues for scientists to further study the genetic history of cattle populations by incorporating recent developments in genomic tools into bovine full genome characterization. For both population genetic and linkage mapping studies in animal genetics, microsatellites have been commonly used (Groeneveld *et al.*, 2010) and are still used, for example, in parentage testing and population genetics. Progress in genetic science has increased dramatically with the publication of the first draft of the human genome sequence and the advancement of next-generation sequencing (NGS) methodologies in the early 21st century (the first NGS machine was commercially available in 2004). The first version of the cattle genome sequence was published in 2009 (Bovine Genome Sequencing and Analysis Consortium 2009) and microsatellites were replaced by single nucleotide polymorphisms (SNPs). An SNP is a variation in the genome between the maternal haplotype and the paternal haplotype at a single nucleotide position, which occurs in at least 1% of individuals within a species (Liu *et al.*, 2020). In theory, one SNP can show up to four alleles, but at each locus, an SNP is biallelic in most cases (Albers and McVean, 2018).

In 2008, the first commercial version of the whole-genome SNP array (more than 50,000 SNPs) came on the market (Matukumalli *et al.*, 2009) and in 2010, a higher density version of the array (more than 700,000 SNPs) followed. Whole-genome SNP arrays enable the analysis of genetic population histories and detection of chromosomal regions under selection more accurately than was previously possible (Pérez-Enciso *et al.*, 2015). Compared to microsatellites, the drawback of SNPs is the small number of alleles; polymorphic information content (PIC) is high for microsatellites compared to SNPs (Teneva *et al.*, 2013). Microsatellites are typically mostly neutral (i.e., not causing a difference in phenotype but SNPs are potentially causative). Relatively low mutation rates for SNPs are (10^{-8} to 10^{-9}) (Albers and McVean, 2018). The most common type of genetic variation is SNPs, and their presence across the whole genome makes them perfect for studying genomic region inheritance (Liu *et al*., 2020). Genotyping of the bovine genome with SNPs is one of these developments. The low SNP density array (BovineSNP50 Genotyping Bead Chip, versions 1 and 2), which genotypes more than 54,000 SNPs (Bejarano *et al*., 2018), and the higher density array (Bovine HD Genotyping Bead Chip), which genotypes more than 777,000 SNPs, are two examples of these arrays for cattle built by Illumina (Rincón *et al*., 2011). Centered on the dual-color, single-base extension infinium HD assay, both of these arrays are genotyped SNPs (Bejarano *et al*., 2018). For genome-wide association research purposes, SNPs genotyped by these two arrays are validated principally in commercial taurine cattle breeds. This bias of European taurine breeds, which is more important in the lower density array, will make this method less successful in analyzing the genome of zebu cattle and indigenous taurine cattle populations outside of Europe (Bejarano *et al.*, 2018).

The most common type of genetic variation is SNPs, and their presence across the whole genome makes them perfect for studying genomic region inheritance (Liu *et al*., 2020). An SNP is a variation in the genome between the maternal haplotype and the paternal haplotype at a single nucleotide position, which occurs in at least 1% of individuals within a species (Liu *et al*., 2020). Relatively low mutation rates for SNPs are (10^{-8} to 10^{-9}) (Albers and McVean, 2018). There are two forms of SNP: transformation, which is a change between two bases of purine or two pyrimidines, and transversion, which is a change between the base of purine and pyrimidine. In theory, one SNP can show up to four alleles, but at each locus, an SNP is biallelic in most cases (Albers and McVean, 2018).

SNPs are helpful genetic markers for testing their relationship with diseases or quantitative qualities of economic importance, and SNP chips are excellent for large-scale genotyping to help improve the genetics of farm animals and plants. They have been frequently used in cattle for quantitative trait loci (QTL) mapping, genome-wide association studies, and genomic selection (Magnabosco *et al.*, 2016). There are *B. indicus* SNP chips on the market, such as the Gene Seek Genomic Profiler High-Density *B. indicus* (GGP HD) SNP chip (80K) and the GGP low-density *B. indicus* (50K) SNP chip, but they were designed with both *B. taurus* and *B. indicus* content and were not optimally designed for *B. indicus* cattle, due to limited genotype data (Lashmar *et al.*, 2019).

The imputation accuracies were close to those published by Junior *et al*. (2017), who reported accuracies of 0.888 to 0.993 when SNP genotypes were imputed from 3K, 7K, 20K, 50K, and 80K SNPs to 777K SNPs in Gyr (*B. indicus*) dairy cattle, respectively. Their findings revealed that imputation accuracy rose as the LD chip's SNP density increased, but that the gain in accuracy became relatively lower as the LD chip size increased beyond 15K SNPs because imputation accuracy was already high at this density (Neves *et al.*, 2014). The success of using highly dense commercial DNA markers (SNP chips) is dependent on several parameters, including, (i) Because there are many composite synthetic commercial breeds accessible, as well as certain breeds that are a blend of multi breeds, the chips must be produced, (ii) Many commercial species are genetically distinct, and their surroundings differ; significant changes in the effects of DNA markers in separated populations are expected (this would be an expression of genotype-environment interactions), (iii) In genomic technology, the cost of genotyping is a crucial issue.

The qualities are chosen in a variety of ways. For example, milk selection differs significantly from beef selection, and the value of selected replacement of beef bulls, which are generally utilized in natural mating in over 90% of herds, differs significantly from that of dairy bulls employed in artificial insemination (Gomes *et al*., 2012).

2.3. Measurements of genetic diversity and differentiation

Genetic diversity is an important feature of population dynamics since it is closely correlated with the population's evolutionary potential and the deleterious effects of inbreeding (Greenbaum *et al*., 2014). However, many different forms of genetic diversity measurements exist most notably heterozygosity-based measures and allelic richness-based measures (Greenbaum *et al*., 2014). Allelic richness (number of alleles) is a genetic diversity indicator suggesting the long-term adaptability and persistence potential of a population. It is used as a genetic diversity metric less frequently than heterozygosity, partly because the stochastic mechanism of genetic drift for allelic richness is more difficult to take into account mathematically (Greenbaum *et al*., 2014). According to the studies by Greenbaum *et al.,* (2014) on allelic diversity (Allelic richness) using stochastic modeling by integrating gene flow and genetic drift into a source population and newly formed population, and in their studies, they stated that genetic drift and gene flow are experiencing the allelic richness of a newly founded population. For example, it is known that the number of segregating alleles in a population offers basic information on past population size fluctuations (Armando and Aurora, 2013).

In gene-frequency-diversity or allelic-diversity components, the division of diversity contributes to very different conservation strategies (Caballero *et al*., 2010), indicating a complementarity between the two forms of measures of diversity. In general, the components of quantitative genetic variance for characteristics and gene-frequency-diversity measures are more strongly correlated with short-term selection response, whereas allelic-diversity measures are more correlated with long-term and total selection response (Armando and Aurora, 2013).

Traditionally, the information provided by genetic markers has been used to measure the parameters associated with the distribution of genetic variation within subdivided populations, e.g., within and among domestic breeds. For instance, the heterozygous average number of alleles or FIS statistics are measures of diversity within breeds (Carroll *et al.*, 2018). In addition, the genetic relationships between breeds, between species, can be determined by evaluating genetic distances. Heterozygosis, which estimates the proportion of heterozygous individuals for a single marker or, in extended terms, for an average set of markers, is one of the parameters traditionally used as a measure of genetic diversity. There are two methods in which heterozygous can be measured; one is through measuring the proportion of heterozygous individuals through counting the number of heterozygous genotypes, or by recording their genetic frequencies, defined as observed heterozygosity (Ho). The second is also known as "gene diversity," the "expected heterozygosity (He)," which is characterized as the heterozygous value that would be expected under Hardy-Weinberg equilibrium conditions. In general, a subdivided population shows lower levels of heterozygosity than expected; the reduction in heterozygosity observed can be used to measure the degree or extent of differentiation between subpopulations (Carroll *et al.*, 2018).

In any population analysis, to estimate genetic differentiation within the species as a whole, the ideal first step is to collect samples of the species over its entire range. However, most studies concentrate on selective sampling for economic or sampling constraints in particular areas with an economic or conservation interest. Genotypes or haplotypes are scored for the individuals sampled, depending on the markers used, and the data is analyzed in a variety of ways to measure genetic variation between populations (Rodriguez-Pena *et al.*, 2018).

For the minimum number of individuals to sample per venue, there is no universal law. In estimating population structure, the individuals seem to be a reasonable trade-off between sampling cost and bias. Several genetic differentiation markers exist. F-statistics, which was created by Wright (1949/1950), is one of the most commonly used measures.

Wright's F-statistics: We need to identify a key parameter in population genetics before implementing population structure measures: the fixation index also called the coefficient of inbreeding. To compare how much heterozygosity is present in the actual population compared to predicted levels of heterozygosity under random mating (and other HW equilibrium conditions), a number symbolized F, is widely used:

$$F= \frac{(He-Ho)}{He} \tag{1}$$

Where H_e is the expected frequency of heterozygotes based on the frequency of the population allele, Ho is the frequency of heterozygotes observed. F is placed on a convenient scale of -1 and +1 by dividing by the predicted heterozygosity. Negative values indicate heterozygote excess and positive values indicate homozygote excess. The analysis of genetic differentiation between populations includes many new versions of the fixation index, the so-called F-statistics, to account for the divergence of sub-populations. During the 1920s, Sewall Wright introduced the idea of F-statistics but later suggested the three parameters as we know them now. These indicators have been developed to identify the genetic composition of diploid species in the population. The basic assumptions are that all populations are of the same size and that any population has equal opportunities to exchange people with any other population. First, for each biallelic loci, heterozygosity is measured and then averaged according to the scale considered (total population, subpopulation). So, a series of heterozygosity hierarchical measures were defined: HI: means heterozygosity observed across subpopulations, HS: mean predicted heterozygosity among randomly mating subpopulations within each subpopulation, HT: predicted heterozygosity within the total population, with random mating. HT and HS have maximum values of 0.5, and HI will differ between 0 (no observed heterozygotes) and 1 provided the biallelic loci (all observed individuals are heterozygous). Now, three hierarchical F-statistics are described based on HI, HS, and HT: FIS, FST, and FIT.

FST is fixation index and estimates as:

$$\mathrm{FST} = \frac{(HT-HS)}{HT},\ 0 \leq \mathrm{FST} \leq 1 \qquad (2)$$

The FST coefficient reflects the discrepancy between the average predicted subpopulation heterozygosity and the expected total population heterozygosity, so it calculates the decrease in heterozygosity due to the allele frequency divergence of the subpopulation. It is possible, at a lower stage, that two alleles randomly sampled from a single subpopulation are similar by descent. An FST value close to 0 means that there is no distinction between subpopulations, and a value close to 1 means that subpopulations are fully differentiated. Although FST has a theoretical range between 0 and 1, the limit observed is typically much less than 1. The following are qualitative FST interpretation guidelines; little genetic differentiation, 0 to 0.05, moderate genetic differentiation, 0.05 to 0.15, great genetic differentiation, 0.15 to 0.25 and very large genetic differentiation, > 0.25 (Wright,1965).

2.4. Linkage disequilibrium

Linkage disequilibrium (LD) is the non-random association of alleles at different loci in a given population. When the frequency of association of their various alleles is higher or lower than what would be predicted if the loci were separate and randomly linked, loci are said to be in linkage disequilibrium (Hui and Burt, 2020). The Hardy-Weinberg principle, also known as the Hardy-Weinberg equilibrium, model, theorem, or law, states in population genetics that allele and genotype frequencies in a population will remain constant in the absence of other evolutionary factors from generation to generation. Genetic drift, mate preference, assortative mating, natural selection, sexual selection, mutation, gene flow, meiotic drive, genetic hitchhiking, the bottleneck of the population, founder effect, and inbreeding are among these factors (Hui and Burt, 2020).

At equilibrium, the Hardy-Weinberg equation tests the genetic variance of a population and it looks like: It is said that genes located next to each other on the same chromosome are likely linked. Due to random interaction, two alleles from different genes on the same chromosome appear to be associated with different individuals at a higher frequency than predicted, there is linkage disequilibrium between these genes. Two genes are unlinked on separate chromosomes or the same chromosome at great distances (Linkage-Equilibrium).

Linkage-Disequilibrium can be generated by several variables such as genetic drift. Genetic drift is greater when the size of the population is small. In a haplotype, mutation may occur and establish an LD between the mutated locus and this haplotype. This disequilibrium usually decreases over generations, but if genetic drift or selection occurs, it may increase. LD can also be produced by population mixing or by migration. Initially, LD is proportional to the variations in alleles between populations and is independent of the distance between markers. LD decreases over generations, but the rate of decrease varies with the relationship between loci. LD tends to vanish for independent loci, and it lasts much longer for linked loci. Selection leads to a decrease in the number of breeders as genetic drift and hence a decrease in the number of haplotypes in the population. Similarly, this reduction leads to a rise in consanguinity. Around the genome, the recombination rate is not constant. The strength of the linkage can be characterized by the level of LD. LD is large in medium recombination regions and high in low recombination regions (Espigolan *et al.*, 2013).

Linkage disequilibrium was calculated between pairs of SNPs within a chromosome using PLINK (Purcell *et al.,* 2007). The r^2 measures, which was described as the square correlation coefficient of two-loci alleles, that was selected because it was independent of the frequency of the allele (Lu *et al.*, 2012). Briefly, two loci, A and B, each locus having two alleles (denoted A1, A2; B1, B2, respectively) were included in its calculation (Qanbari *et al.*, 2010). The haplotype frequencies were denoted as f11, f12, f21, and f22 respectively for haplotypes A1B1, A1B2, A2B1, and A2B2, and fA1, fA2 fB1 and fB2 respectively for haplotypes A1, A2, B1, and B2. Then, r^2 will be determined from this as:

$$r^2 = \frac{(\mathrm{f11f22 - f12f21})^2}{\mathrm{fA1fA2fB1fB2}} \tag{3}$$

2.4.1. Application of linkage dis-equilibrium

With the expansion of genomic selection (GS) and genome-wide association studies (GWAS) published in recent years, there has been an increasing interest in LD analysis. The LD between markers and quantitative trait loci (QTL) is used in both GS and GWAS to estimate genomic breeding values (GEBV) or to identify regions that control traits of interest. The number of animals in the reference population, the heritability of the trait, the distribution of QTL effects, and the strength of familial connections between the reference population and the selected candidates all influence the accuracy of GEBV (Wientjes *et al.*, 2013). To assess the feasibility of employing a reference population from a different breed or cross, it's necessary to understand the LD in those breeds and crosses, as well as the LD phase's persistence across these populations and with the population of selection candidates. The accuracy of forecasts in populations with distantly related individuals is mostly influenced by LD between markers and QTL, whereas predictions in populations with closely related populations rely primarily on the direct relationship (Daetwyler *et al*., 2012). McKay *et al*. (2007) used 2670 single nucleotide polymorphism (SNP) markers genotyped in eight breeds of *B. taurus* and *B. indicus* cattle to determine the degree of LD in cattle. According to their findings, moderate LD ($r^2 = 0.20$) in cattle stretched to 40–60 kb, implying that roughly 50,000 SNPs would capture most of the LD information needed for GWAS in European *B. taurus* breeds. Thevenon *et al*. (2007) used 42 microsatellite markers on chromosomes 1, 4, and 7 to assess the degree of LD in 363 *B. indicus* and *B. taurus* cattle from western Africa. Their findings showed that LD in cattle from developed countries extended for shorter distances than previously thought. This meant that GWAS in these hybrids would require between 75,000 and 300,000 SNP markers. Edea *et al.* (2014) used the Illumina BovineSNP50 Bead Chip to assess the degree of LD in Ethiopia's indigenous breeds and found that the extent of LD was lower in Ethiopian cattle populations than in European *B. taurus*. The reduced degree of LD identified in these breeds was related to SNP ascertainment bias, which was caused by the identification of SNPs employed in the assay design in European *B. Taurus* breeds. The preservation of allele phase correlations between markers and QTL across populations is critical for the success of LD applications across populations (Mustafa *et al.,* 2018). As a result, it is critical to quantify the degree of LD and the durability of allelic phase correlations in populations where GWAS, MAS, and genomic selection will be used.

2.5. Selection signatures

Natural and artificial selection, along with adaptations to different biogeographic regions and conditions of development, have influenced allele frequencies in cattle associated with adaptation or modification of the selected characteristics. It is difficult to methodologically distinguish the impacts of natural or artificial selection (Randhawa *et al*., 2016). Selection may result in linkage imbalance (non-random association of alleles at different locations, LD) and lower genetic variability in regions close to a preferred allele, leading to observable trends that facilitate the location of the genome selective region. Continuous selective sweep signals indicate the existence of genetic variants that are likely to influence phenotypes but may also occur because of genetic drift or demographic processes, particularly in artificially selected organisms (Jacobs *et al*., 2016).

Understanding how selection affects a population will help breeders establish effective breeding plans to improve commercially valuable traits in cattle (Gurgul *et al.*, 2018). Mutations can be neutral or they may be deleterious and affect the population's fitness. Consequently, the mutation's allele frequency does not increase or become set. However, the allele frequency for those functional variants increases and can become fixed when desirable alleles are selected for a prolonged period. These trends include the hitchhiking effect (i.e., linkage imbalance) consisting of neutral position areas downstream or upstream of the functional version (Derks *et al.,* 2018).

2.5.1. Methods to detect selection signatures

Fst is locus-specific, which means it can detect which genetic variants are being selected. It is preferable to look for multiple SNPs with similar average Fst scores rather than assessing each SNP individually. Using hierarchically ordered data sets, Fst statistics can detect false positive/negative outcomes (Fariello *et al.*, 2014). Several statistical approaches are used to classify signs of selection within the genome. All statistical measures to detect positive selection are broadly focused on five signatures according to Fariello *et al*. (2014), a high proportion of mutation-altering roles, reduction in genetic diversity, high-frequency derived alleles, population variations, and long haplotypes.

Most methods for detecting selection have been established (Utsunomiya *et al*., 2015) because it creates visible footprints on the genome. In the age of genomic data, adaptive processes can now be inferred in the absence of phenotypic data. Methods of selection signatures detection are sometimes described as "genome to phenotype" approaches involving the statistical evaluation of population genomic data irrespective of phenotype to identify probable prior selection targets (Qanbari and Simianer, 2014). Depending on the test statistics, selection signatures can be found either from intergenic regions, coding regions or from both (Vitti *et al*. 2013).

2.5.2. Types of selection signals

Positive Selection: - Darwin and Wallace came up with the notion of natural selection in the 1850s. Darwin's theory of natural selection claimed that "If variations useful to any organic being ever occur, people thus characterized will surely have the best chance of being preserved in the struggle for life; and these will tend to produce offspring similar to the strong principle of inheritance" these will appear to generate similarly characterized descendants. This principle of preservation, or the survival of the fittest, is called natural selection (Flannery, 2018). If the phenotype increases the fitness of an organism, it becomes more prevalent in a population over time. This phenomenon is called Positive Selection. The allele(s) behind the beneficial phenotype would finally become more common at the population level. At the sequence stage, genomic signals of positive selection are characterized by reduced local variability, a deviated spectrum of allele frequencies, and unique patterns of linkage disequilibrium.

Positive selection is defined as a significant difference in allele frequency between populations at a given locus (Zhao *et al.,* 2015). Earlier experiments on positive selection signatures were made with microsatellite markers or small numbers of SNPs. The availability of whole genome-wide genotyping arrays and whole-genome sequences has expanded research capability. In cattle, selection signals are correlated with domestication and adaptation. Variations in coat color in cattle, for instance, are a characteristic related to domestication (Qanbari *et al.,* 2014).

Negative selection (purifying selection): - Deleterious mutations are normally eliminated from the gene pool before any measurable frequency is reached within a population so that genome regions are under intense negative selection pressure where no differences are accepted and are thus typically highly conserved across organisms (Vitti *et al*., 2013).

Balancing selection: - When multiple alleles are retained in a population at an intermediate frequency, such a phenomenon is called the selection of balancing. Owing to the very subtle effects on the genome, identification of balancing selection is more difficult, and negative selection is usually observed for conserved regions. Balancing selection will occur due to heterozygote gain (i.e., the individual heterozygote has greater fitness compared to any of the homozygotes) or frequency-dependent selection (i.e., when it is uncommon, an allele has greater fitness and several alleles will be retained in population) (Vitti *et al.*, 2013). Balancing selection is the most difficult type of selection to detect and current methods suffer from low power and high false-positive frequency (Fijarczyk and Babik, 2015).

2.5.3. Selection signatures in the cattle genome

The genome of cattle has 30 pairs of chromosomes and is 2.65 Gbps in size, with 22,000 genes (Elsik *et al*., 2009). The biological mechanisms that differentiate breeds artificially selected for distinct phenotypes and adaption to a local environment can be understood by identifying the fingerprints of natural and artificial selection pressures in the genomes of cattle. So far, a catalog of genes has been identified from a variety of cattle breeds with varying demographic histories, and these studies will presumably continue to find new genes that are linked to local adaptation qualities in various local breeds. Low Fst values among subpopulations indicate low levels of genetic divergence in the population, whereas a value of 0 indicates that the populations are not divided (no differentiation). High Fst values suggest greater genetic differentiation between subpopulations. An Fst value of 1 implies complete separation of the subpopulation from the overall population (extreme differentiation).

The Bovine HapMap Consortium (2009) analyzed the genomes of nineteen geographically and physiologically different bovine breeds, identifying genomic areas that had been impacted by domestication and artificial selection.

Various genetic areas affecting milk yield and composition, meat quality, and feed conversion efficiency in different breeds were discovered under selection in the study (The Bovine HapMap Consortium 2009). Several beef and dairy cattle breeds have a hallmark of genetic areas controlling coat color (*MCR1, KIT*) and body size and stature (*PLAG1*) (Kim *et al*., 2017). Rothammer *et al.* (2013) analyzed the genomes of ten beef and dairy cattle breeds and identified genes (*TG, ABCG2, DGAT1, GH1, GHR*, and the Casein Cluster) that are highly linked to QTL regions for beef and dairy attributes.

Lee *et al*. (2014) looked at the genome of Holstein cattle and found genomic areas related to milk yield and composition, as well as those associated with genetic diseases, that were influenced by selection (cardiovascular disease). Artificial selection for milk and beef traits has had such an impact on the genomes of dairy and beef cattle breeds that genes and gene regions associated with milk and beef traits have been identified as being under positive selection in the respective breeds. Despite current methods and technologies, the genomes of African indigenous cattle breeds are being studied less extensively. However, because of current initiatives around the world to use the whole genome and SNPCHIP data from cattle breeds, the situation is projected to swiftly alter (Kim *et al*., 2017).

Bahbahani *et al*. (2015) analyzed the genome of small East African Shorthorn Zebu cattle for positive selection signatures and detected 24 candidate genomic regions under selection with 409 annotated genes involved in biological pathways of immunity, reproduction, development, and heat tolerance. They also detected signals of selection in heat shock protein families (*HSPB9, DNAJC7, DNAJC8, DNAJC14*, and *DNAJC18*) and heat stress response (*PPP1R10*) genes (Bahbahani *et al.*, 2015). Genes associated with adaptation to tropical environments (*KRT222, KRT24, KRT25, KRT26, KRT27*, and *SPB9*), immune response (*CYM, CDC6*, and *CDK10*), and nervous system development (*WNT5B, FMOD,* and *CDK10*) in six South African cattle breeds. Strong candidate genes for immunological response (*MHC, CD79A, CXCR4, DLK1, RFX3, SEMA4A, TICAM1,* and *TRIM21*), nervous system (*NEUROD6, OLFM2, MAGI1, SEMA4A*, and *HTR4*), and skin and hair qualities (*EDNRB, TRSP1,* and *KRTAP8-1*) were identified in West African cattle breeds (Makina *et al*., 2015).

Positive selection in the African cattle genome is driven by a variety of natural selection factors. The continent's high ambient temperature could be a driving force behind the development of superior thermotolerance mechanisms in African cattle, such as improved thermoregulation, higher fertility, and faster growth rates than European cattle breeds in tropical settings (Bahbahani *et al.*, 2015). The processes of thermotolerance include molecular and cellular functions (Belhadj Slimen *et al.*, 2015), physical structures of hair and skin (Jian *et al.*, 2014), thermal sweating (Jian *et al*., 2014; Lenis Sanin *et al*., 2016), and physiological mechanisms (Jian *et al*., 2014; Lenis Sanin *et al.*, 2016). The African environment's higher disease prevalence could be a driving force behind the evolution of disease and parasite resistance in African cow breeds (Okeyo *et al.*, 2015). The tsetse fly, a vector for trypanosomiasis, is well-known in the humid and subhumid West African region (Yaro *et al*., 2016). In response to the selective pressure caused by trypanosome challenges, the cattle breeds that predominate in these areas have developed adaptive features. Shorthorn taurine cow breeds showed greater anemia control in an experiment comparing diverse West African taurine cattle breeds to indicine cattle breeds for a trypanosome challenge, demonstrating their local adaptation to trypanosomiasis (Berthier *et al*., 2015). In addition, when compared to other African zebu and commercial cattle breeds, selection signatures analysis demonstrated resistance to trypanosomiasis in West African N'Dama cattle (Kim *et al*., 2017).

2.5.4. Breed-specific SNPs

An SNP was breed-specific when it was polymorphic in one breed but fixed in other breeds. Many livestock species have evolved distinct breeds as a result of natural and artificial selection, as well as genetic drift due to limited population sizes. As a result, breed names are becoming more commercialized and used as brand names and breed validation of livestock products is becoming increasingly important in determining food safety and authenticity in global and domestic markets (Pant *et al*., 2012). This problem could be solved quickly, easily, reliably, accurately, and economically by testing an individual's genotype at unique marker loci. In recent years, there has been a lot of interest in the possible application of genetic markers to track individuals back to their source breed, and various approaches to animal traceback have been investigated. The molecular markers used to assign breed to individuals, as well as the allocation process used to assign a breed to individuals based on their genotypes at unique marker loci, assess the accuracy and efficiency of breed assignment to individuals (Zwane *et al.,* 2016). Breed-specific SNPs could, therefore, provide the livestock industry with an easy, rapid, economic, and reliable method to validate the breed of livestock individuals and products.

CHAPTER THREE

3. MATERIALS AND METHODS

3.1. Study breeds, sample collection, and DNA extraction

DNA samples were collected from three Ethiopian indigenous cattle populations (n = 123) that were kept in different government ranches: Begait (n = 40), Boran (n = 40) and Fogera (n = 43). The Begait, Boran, and Fogera cattle were sampled from Humera, Dida Tiyura, and Andassa ranches, respectively. Unrelated female and male animals were sampled based on available pedigree information. Nasal samples were collected using Performagene livestock's nasal swab DNA collection kit and DNA was extracted from nasal samples according to the manufacturer's recommendations (DNA Genotek Inc., 2012). Three European beef breeds (n =114) (Angus, n = 42), Hereford (n = 35), and Charolais (n = 37) were used as reference breeds.

3.2. Description of the breeds

Boran is a zebu (*B. indicus)* breed maintained in southern Ethiopia and neighboring areas. Usually, Boran cattle are white or grey fawns (Haile *et al*., 2011). They have a well-developed dewlap and preputial sheath; in males, they have small ears that are often pendulous and higher than in females. It has a deep chest, a somewhat messy rump, and well-developed hindquarters. They are very flexible and adapt well to different conditions. Cows are very effective converters of pasture forage into deposits of body fat that are later mobilized during feed shortage and lactation times. Consequently, during lactation or mild droughts, the cows hardly lose conditions. Depending on the improvement status, level of management, and production system, the average live weight for adults varies. Mature bulls of Boran in Ethiopia, for example, have been registered to weight ranges from 318 to 680 kg in males and 225 to 454 kg in females (Haile *et al.,* 2011).

Fogera is Zenga (*Zebu x Sanga*) breed distinguished and well known for its black-and-white or black-and-grey pied coat; short, stumpy, pointed horns; hump ranges from thoracic to cervicothoracic; dewlap is folded and moderate to wide in size; used for drought, milk, and meat (DAGRIS, 2007). One of the main characteristics of the breed is a large body with high strong bones (DAGRIS, 2007). Alberro and Haile-Mariam, (1982) classified Fogera breed as an intermediate Zebu-Sanga type formed from the interbreeding of Abyssinian Zebu from the central highlands and the neighboring Sanga (Danakil, Raya-Azebo). Fogera cattle are typically huge, being tall with long legs, one of the main features of the breed is a massive body with sturdy strong bones. Because of its lower age at first calving and ability to plough the marshy area and better dowry (exchange value) and other social purposes, farmers choose Fogera breed (Sendeku *et al.*, 2016).

Begait cattle belong to the Zebu community of North Sudan and are raised for both milk and beef and reared in Tahtay Adiabo, Kaftahumera, Welqait, Asgedet Simbla districts of northwest Ethiopia. The breed is characterized by well-developed udder, long legs, and large humps, long teats, and higher milk, 5 ± 0.5 liters/day, a live body weight of 333 ± 51kg, and 278 ± 41kg for males and females, respectively. The breed is adapted to heat stress, water shortages, long-distance walking to find feed and water.

Angus breed is originated in northeastern Scotland and mostly black cattle. They are a top-ranking beef breed. In the early 19th century, several constructive breeders developed the breed and fixed the existing type of cattle. The mature cows are 500-550 kg in weight and the live weight of the bulls can reach up to 1000 kg with a fattening weight of 750-950 kg (*https://www.Livestock of theworld.com/Cattle/).*

Charolais are originated from west-central to south-eastern France. They are a breed of large light-colored cattle which was created for draft purposes in France but is now kept for the production of beef and used for crossbreeding. They are renowned for their longevity and good fertility and they are good herd animals and have good maternal qualities, because of their calm disposition. In reality, they display the greatest fattening efficiency and high specific forage intake capacity. Large carcass yield and distinct muscularity with relatively low-fat deposits on the carcass and bulls' live body weight ranged from 1,000 to 1,650 kg, cows from 700 to 1,200 kg, are features of the Charolais (*http://afs.okstate.edu/breeds/cattle/Charolais /index.html/).*

Hereford is named from the county of Herefordshire, England's historic farming area where they originated. These breeds are medium to large and mainly meat cattle. Colored with a white mask, they are usually dark red to red-yellow. They are available with both horned and polled variants. The horned variant has horns that are short and thick. At the sides of the head, their horns are normally bent back. They are muscular, with medium to long side lengths https://www.roysfarm. com/Hereford-cattle/. The mean live body weight of mature cows is approximately 800 kg and bulls weigh about 1200 kg on average (*https://www.michigan statefairllc.com/beef-cattle*).

3.3. Genotyping and quality control

Ethiopian cattle samples were genotyped with the 80K SNP Bead Chip (Gene Seek Genomic Profiler). The SNP markers were screened for a call rate of ≥ 90%, a minor allele frequency (MAF) of > 0.01, and a sample call rate of > 90%. After the above quality management parameters had been applied, the autosomal SNP markers obtained were used for downstream analysis. Two hundred thirty-one animals were kept after removing 6 animals with a genotype completion rate of less than 90%. From an initial set of 67,491 SNPs, a subset of 67477, 67468, 67414, 66811, 66934, and 65460 SNPs for Angus, Herford, Charolais, Boran, Fogera, and Begait respectively were kept for breed-specific SNPs and MAF analysis, and 67252, 67271, 67260, 67392, 67400, 67410, 67310, 67299, and 67300 for Begait-Angus, Begait- Hereford, Begait- Charolais, Boran-Angus, Boran-Hereford, Boran-Charolais, Fogera-Angus, Fogera-Hereford and Fogera- Charolais, respectively were used for selection signatures and LD analyses.

3.4. Data analysis

To examine the basic indices of within-breed genetic variability, MAF and genetic diversity (observed and expected heterozygosity) were estimated using PLINK (Purcell *et al*., 2007). PCA was analyzed using the same software and plotted in R. Power Marker software was used to construct phylogenetic trees (Liu and Muse, 2005). Pairwise genetic differentiation (FST) values were calculated using PLINK between the following breed pairs: Boran-Angus, Boran-Hereford, Boran-Charolais, Begait-Angus, Begait- Hereford, Begait- Charolais, Fogera-Angus, Fogera- Hereford and Fogera- Charolais according to (Weir and Cockerham,1984). For all pairwise comparisons, the top 1% FST values were considered to define SNPs under positive. The Bovine UMD.3.1 genome assembly was used to annotate candidate SNPs within the top 1% FST values. Functional analysis was performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) resources (v. 6.8; https://david.ncifcrf.gov/). QTL regions that overlapped with the identified candidate regions were searched from the cattle QTL db (http://www.animalgenome.org/cgibin /QTLdb /BT/index). The physical map of the identified candidate genes associated with the Gene Ontology (GO) term was mapped using map chart software (Voorrips, 2002).

Linkage disequilibrium was calculated between pairs of SNPs within a chromosome using Using SNP & VARIATION SUITE (Golden Helix, Inc; www.goldenhelix.com). The r^2 measures, which was described as the square correlation coefficient of two-loci alleles was selected because it was independent of the frequency of the allele (Lu *et al.*, 2012). Briefly, two loci, A and B, each locus having two alleles (denoted A1, A2; B1, B2, respectively) were included in its calculation (Qanbari *et al.*, 2010). The haplotype frequencies were denoted as f11, f12, f21, and f22 respectively for haplotypes A1B1, A1B2, A2B1, and A2B2, and fA1, fA2 fB1 and fB2 respectively for haplotypes A1, A2, B1, and B2. Then, r^2 will be determined from this as:

$$r^2 = \frac{(\mathrm{f11f22} - \mathrm{f12f21})^2}{\mathrm{fA1fA2fB1fB2}}$$

Breed-specific SNP: An SNP has been declared to be breed-specific if it has an allele that is present in only one breed (Ramos *et al*., 2011). MAF was calculated in each breed using PLINK (Purcell *et al*., 2007).

CHAPTER FOUR

4. RESULTS

4.1. Minor allele frequency

Minor allele frequency (MAF) was estimated from genotypic data for autosomal markers (Table 1). The mean MAF was 0.19 ± 0.17, 0.20 ± 0.17, 0.21 ± 0.17, 0.31 ± 0.13, 0.32 ± 0.12, 0.32 ± 0.13 for Angus, Herford, Charolais, Boran, Fogera, and Begait breeds respectively (Table 1). A significant difference was detected between the indigenous Ethiopian breeds and European beef cattle breeds.

Table 1: Average MAF in Ethiopian and European cattle breeds

Breed	n	Mean ± SD
Boran	39	0.31 ± 0.13
Begait	38	0.32 ± 0.13
Fogera	40	0.32 ± 0.12
Overall	**117**	**0.31.7 ± 0.13**
Angus	42	0.19 ± 0.17
Herford	35	0.20 ± 0.17
Charolaise	37	0.21 ± 0.17
Overall	**114**	**0.20 ± 0.17**

The percentage of SNPs with a MAF = 0 varied from 1% in Ethiopian cattle breeds to 29% in Angus (Table 2). In Ethiopian cattle breeds and European beef breeds, common variant MAFs (≥0.10 and ≤0.5) accounted for 94% and 62% of the total SNPs, respectively. Angus breed had the lowest proportion of common variants SNPs (59%). Rare variant MAFs (>0 and 0.05) were noted in 2% and 9% of the total SNPs, respectively.

Table 2: MAF distribution of 80K SNP Bead Chip in Ethiopian and European cattle breeds

Breed	n	Fixed (0)		Rare (>0 and <0.05)		Intermediate (≥0.05 and<0.10)		Common (≥0.1 and ≤0.5)	
		SNP	Prop	SNP	Prop	SNP	Prop	SNP	Prop
Boran	39	328	0.01	1385	0.02	2806	0.04	62292	0.93
Begait	38	350	0.01	1077	0.02	2693	0.04	61340	0.94
Fogera	40	690	0.01	1763	0.03	2287	0.03	63176	0.94
Overall	**117**	**456**	**0.01**	**1408**	**0.02**	**2595**	**0.04**	**62269**	**0.94**
Angus	42	19737	0.29	4321	0.06	3551	0.05	39868	0.59
Herford	35	18298	0.27	3634	0.05	3127	0.05	42413	0.63
Charolais	37	9032	0.13	10739	0.16	4641	0.07	43004	0.64
Overall	**114**	**15687**	**0.23**	**6231**	**0.09**	**3773**	**0.06**	**41761**	**0.62**

Among Ethiopian breeds, SNPs with the highest MAFs were observed on BTA.4, BTA.7, BTA.8, and BTA.11 in the Begait breed. The Boran cattle showed the lowest polymorphisms across the chromosomes (Figure 2). Among European beef breeds, SNPs with the highest MAFs were observed in Charolais (Figure 3).

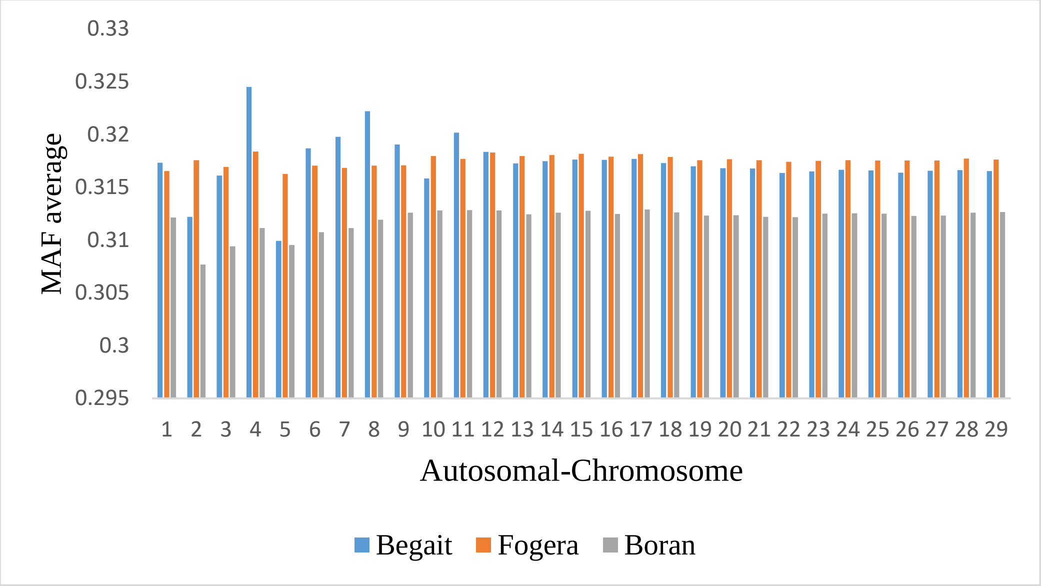


Figure 2: MAF distribution of Ethiopian cattle breeds across the autosomal chromosome, single nucleotide polymorphism

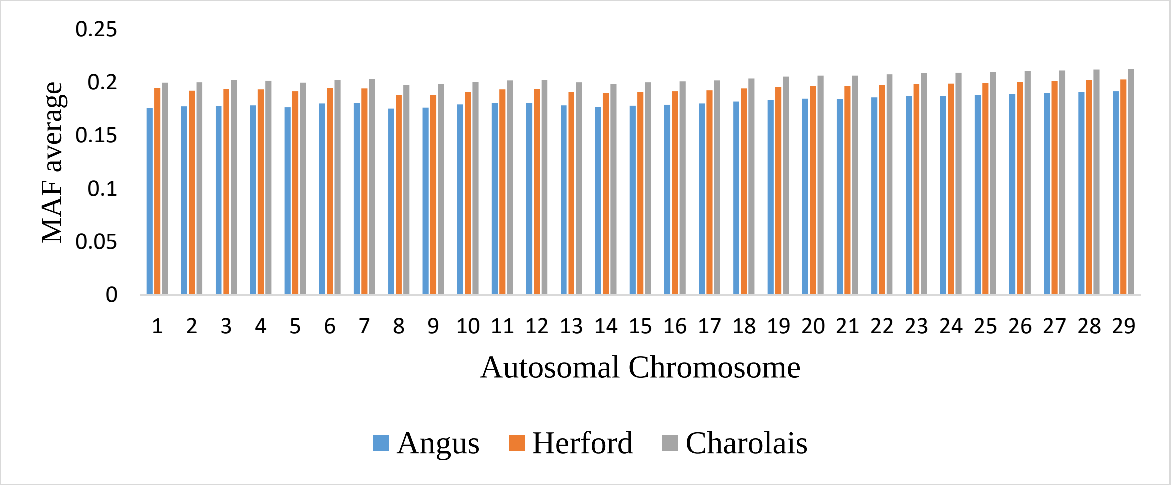


Figure 3: MAF distributions in European cattle breeds across the autosomal-chromosomes

Polymorphisms were also compared between Ethiopian indigenous cattle populations and European beef breeds. The highest polymorphism was observed in the Ethiopian cattle and the lowest polymorphism was shown by European breeds, especially the Angus breed across all chromosomes (Figure 4).

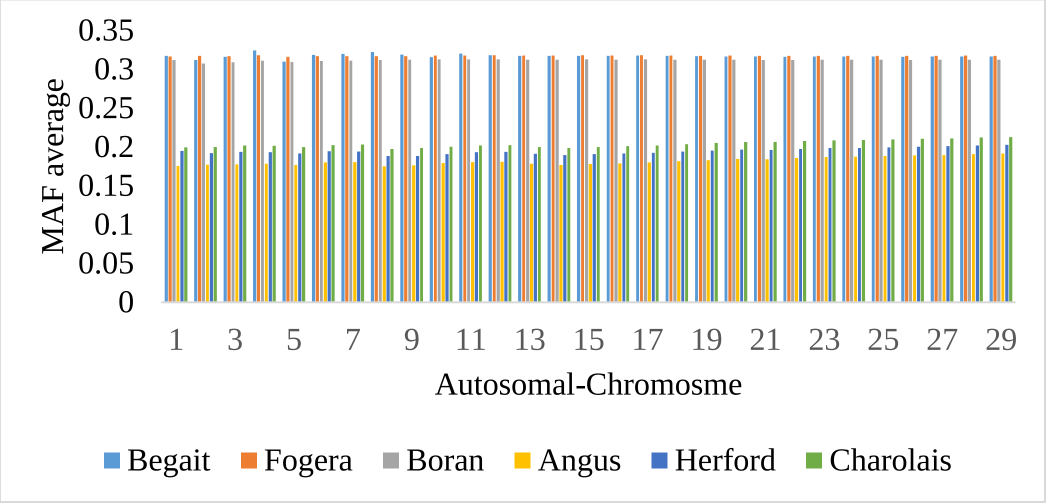


Figure 4: MAF distribution in Ethiopian and European breeds across the autosomal chromosomes

4.2. Breed-specific SNPs

Within Ethiopian cattle populations, the highest number of breed-specific SNPs was detected in the Boran breed. The lowest number of SNPs was observed in the Fogera (Table 3).

Table 3: Breed-specific SNPs detected in Ethiopian cattle breeds

Cattle breed	Number of SNPs	MAF		
		Minimum	Maximum	Average
Boran	90	0.012	0.23	0.04
Fogera	28	0.012	0.14	0.04
Begait	76	0.13	0.22	0.05

Similarly, within the European beef cattle breeds, breed-specific SNPs were identified (Table 4). The Charolais breed had the highest number of unique SNPs (8903) and the Angus breed had the lowest unique SNPs (250).

Table 4: Descriptive analysis of European beef cattle breed, breed-specific SNPs

Cattle breed	Number of SNPs	MAF		
		Minimum	Maximum	Average
Angus	251	0.01	0.44	0.06
Herford	377	0.01	0.42	0.06
Charolais	8902	0.01	0.45	0.02

Breed-specific SNPs in Ethiopian and European cattle breeds were identified. Ethiopian cattle populations had a higher number of unique SNPs, while all European beef cattle breeds showed a lower and similar number of unique SNPs (Table 5).

Table 5: Breed-specific SNPs detected in the comparison of Ethiopian cattle breeds and European beef cattle breeds

Breed	Number of SNPs	MAF		
		Minimum	Maximum	Average
Boran	8038	0.013	0.50	0.32
Begait	7869	0.013	0.50	0.33
Fogera	8031	0.013	0.50	0.33
Angus	48	0.06	0.49	0.33
Herford	48	0.10	0.49	0.33
Charolais	48	0.03	0.50	0.32

4.2.1. Annotation of European beef cattle specific SNPs

European beef cattle-specific SNPs (48) and Ethiopian cattle-specific SNPs (7759) were genomically annotated. European beef cattle-specific SNPs were associated with, 17 genes. Thirty-seven (37%) (0.70%) SNPs were found to be intergenic; 2 (0.04%) SNPs were upstream of genes, 3 (0.06%) SNPs were found downstream of the genes and the remaining 9(0.17%) SNPs were located within the coding regions (Figure 6). Molecular functions of the genes corresponding to European beef cattle breed-specific SNPs were identified as linked to a particular trait.

Interestingly among the identified traits were: the poll locus is responsible for the lack of horns and the *C1H21orf62* gene was identified as a poll locus mutant that is a candidate for longhorn growth. Olfactory transduction pathway, important for direct survival behaviors including detection of food sources, finding mates, and predator avoidance and *LOC618554* identified as a candidate gene for adaptation traits (Table 6).

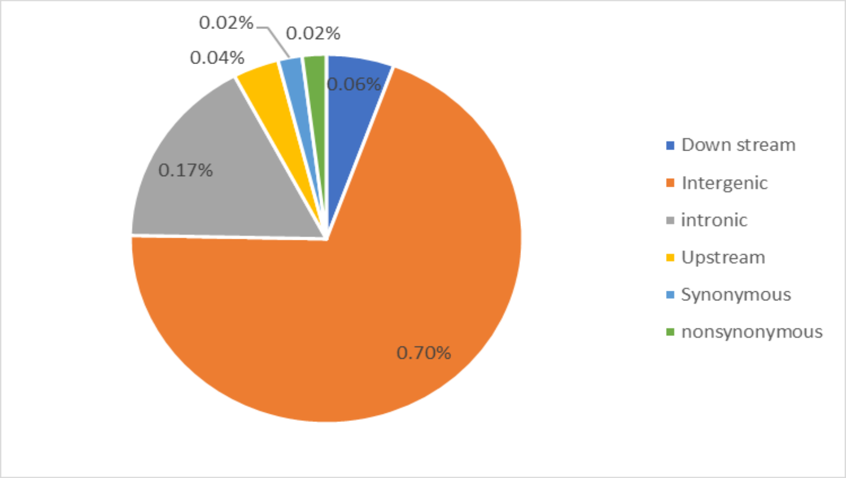


Figure 5: European beef cattle specific SNPs corresponding gene proportion across a genome

Table 6: European beef breeds, specific SNPs, and the corresponding genes and their associated traits

Chr.	Position	rsnumber	Genes	Trait	Reference
1	2049400	rs110875985	*C1H21orf62*	poll locus mutant	Liu *et al*., 2013
1	42749580	rs109175475	*DCBLD2*	Angiogenesis	He *et al*., 2020
2	31080979	rs110334343	*SCN2A*	Voltage-gated sodium channel regulation	Lena and Mantegazza., 2019
8	4270697	rs29015318	*GALNTL6*	Growth and feed production	Seabury *et al*., 2017
8	41527418	rs42495334	*RFX3*	Ciliaogenesis	Chen *et al*., 2018
14	72255136	rs109689318	*RAD54B*	DNA damage and repair	Yasuhara *et al*., 2014
25	39154079	rs110798174	*LOC618554*	Olfactory transduction pathway	Abbas and Vinberg., 2021
25	3768108	rs108945685	*MGRN1*	Homeostasis	Benvegnu *et al*., 2017

4.2.2. *Functional analysis of Ethiopian cattle specific SNPs*

Annotated SNPs identified as Ethiopian cattle-specific SNPs were residing within 3364 genes. Interestingly, some of the identified genes were associated with the relevant biological process including cell morphogenesis (GO:0000902), cellular response to heat (GO:0034605), cellular potassium ion homeostasis (GO:0030007), cellular response to forskolin (GO:1904322), metabolic process (GO:0008152), energy homeostasis (GO:0097009), actomyosin structure organization (GO:0031032), regulation of pH(GO:0006885), cell fate commitment (GO:0045165), pituitary gland development(GO:0021983) and positive regulation of Ras protein signal transduction (GO:0046579).

Table 7: Biological process of Ethiopian cattle specific SNPs corresponding candidate genes

GO Terms	Count	P-Value
(GO:0000902) cell morphogenesis	14	0.0000037
(GO:0034605) cellular response to heat	7	0.00032
(GO:0030007) cellular potassium ion homeostasis	4	0.000083
(GO:1904322) cellular response to forskolin	4	0.00034
(GO:0008152) metabolic process	23	0.00011
(GO:0097009) energy homeostasis	5	0.0060
(GO:0006885) regulation of pH	4	0.027
(GO:0046579) positive regulation of Ras protein signal transduction	7	0.034
(GO:0045165) cell fate commitment	15	0.0049
(GO:0021983), pituitary gland development	6	0.0023

4.2.2.1. *Pathway analysis of the candidate genes*

Of the total genes, 211 genes were involved in various pathways (Table 8). Among these pathways, the insulin signaling pathway enhances inhibition of protein synthesis and the promotion of lipogenesis and glycolysis. The Rap1 signaling pathway regulates inhibition of defective angiogenesis, focal adhesion, and the regulation of actin cytoskeleton, while potentially promoting adherent’s junction, cell adhesion, migration, and polarity. Toll-like receptors (TLR) recognize pathogen-associated molecular patterns originating from diverse microorganisms and play an important role in the innate immune system. Melanogenesis is the process of producing melanin pigments, which are produced by cells known as melanocytes. Platelet activation is critical in hemostasis. Platelets enhance the initial stimulus when the vessel wall is injured and aggregate to create a platelet plug, from which the fibrin component of the thrombus originates.

Table 8: KEGG pathway analysis of genes associated with SNPs detected as Ethiopian Cattle populations specific

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway	Count	P-value
Insulin secretion	26	4.1E-7
Rap1 signaling pathway	38	2.1E-4
Calcium signaling pathway	38	9.7E-5
Toll-like receptor signaling pathway	18	1.5E-3
Platelet activation	28	2.8E-3
Melanogenesis	28	1.3E-2
Endocytosis	35	2.4E-2

4.3. Genetic diversity

Genetic variability analysis was carried out to determine the level of heterozygosity within the Ethiopian and European breeds, and the results are presented in Table 9. The overall mean of observed and expected heterozygosity for Ethiopian breeds were 0.40, 0.40 respectively; while 0.25 and 0.26 values were for European beef breeds (Table 9).

Table 9: Observed (H_O) and expected heterozygosity (H_E) in Ethiopian indigenous cattle breeds and European beef cattle breeds

Breed	H_O	H_E
Begait	0.40	0.40
Boran	0.40	0.40
Fogera	0.41	0.40
Over all	**0.40**	**0.40**
Angus	0.25	0.25
Charolais	0.28	0.28
Hereford	0.23	0.26
Overall	**0.25**	**0.26**

4.3.1. Population structure and relationship

To assess the population structure among the six cattle breeds/populations, a principal component analysis was performed. The SNP data sets clearly distinguished the cattle breeds based on their geographic distributions. Accordingly, the three Ethiopian cattle breeds were clustered separately (Figure 6). As portrayed in Figure 7, PC1 explained 70.63% of the total variation and unequivocally separated the Ethiopian indigenous breeds from European beef breeds (Figure 7). Similarly, the phylogenetic analysis reiterated the PCA results. The three Ethiopian cattle breeds shared the same clade, while the European cattle breeds were found on separate clades (Figure 8). As depicted in Figure 9 the two European cattle breeds (Hereford and Charolais) were closely clustered as did for Boran and Fogera.

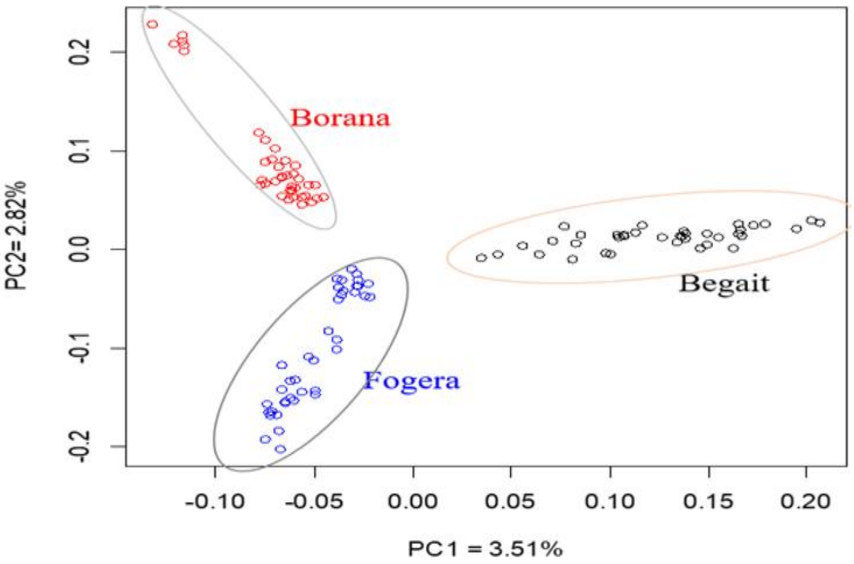


Figure 6: Clustering of Ethiopian cattle populations based on principal component analysis

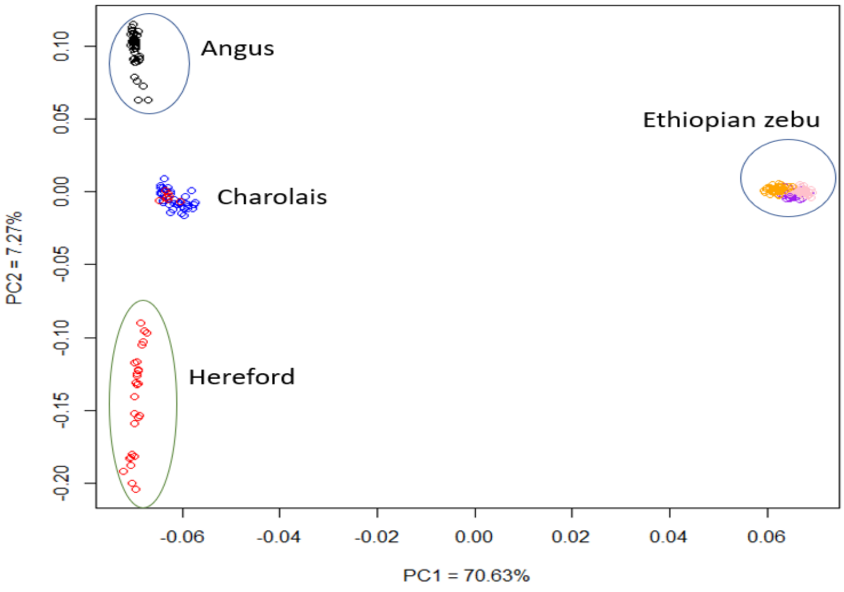


Figure 7: PCA-based Clustering of the study cattle breeds

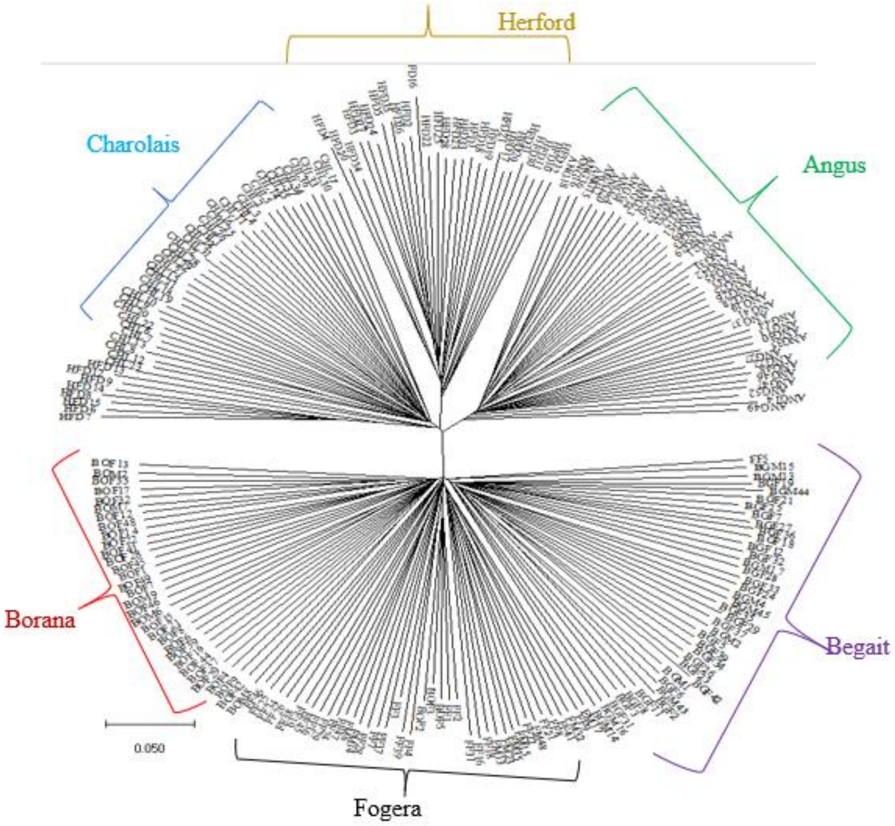


Figure 8: Genetic relationships among 6 cattle breeds constructed using a neighbor-joining tree from shared allele distance

Phylogenetic tree analysis (Figure 9) strongly supports the PCA result, and clustering analysis a clear divergence among breeds of the Ethiopian breeds and in the European beef cattle breeds.

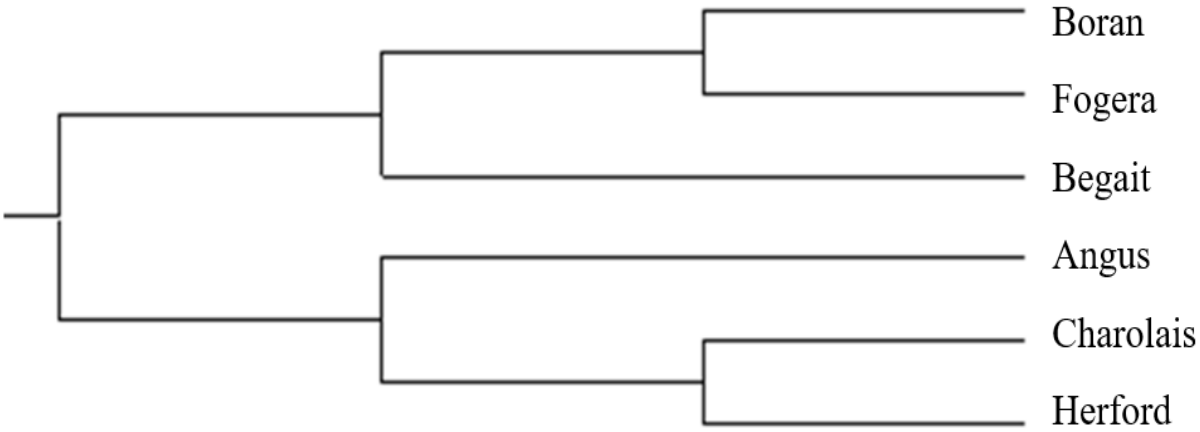


Figure 9: Phylogenetic tree revealing the genetic relationship among the six cattle breeds

4.4. Linkage disequilibrium among pairwise SNPs

Linkage disequilibrium (r^2) values were computed for all pairs of SNPs within each population/breed. Inter-SNP distances were classified into the following kb categories: 0–30, 30–60, 60–120, 120–240, 240 480, 480–960, 960–1920. For all breeds, the average genome-wide LD (r^2 and *D'*) declined with increasing in marker distance. The highest *D'* values observed in Ethiopian breeds dropped from 0.72 at a distance of 0–30 kb to 0.17 at a distance of 960–1920 kb (Figure 11). The same trend was observed in r^2 values, which dropped from 0.26 at 0–30 kb to 0.03 at 960–1920 kb (Figure 10).

In European cattle breeds, *D'* and r^2 values decayed as the distance between SNPs increased. For instance, the Charolais breed r^2 values dropped from 0.23 to 0.03 when the inter SNP distance increased from 0–30 kb to 960–1920 kb. Similarly, *D'* decayed from 0.62 in Herford to 0.35 when the inter SNP distance increased from 0–30 kb to 960–1920 kb (Figure 11). Genomic selection requires an r^2 of 0.2 to reach an accuracy of 0.85. The most practical value of r^2 for association studies is 0.25. In genome-wide association studies, r^2 values greater than 0.30 were required for sufficient power (GWAS). In the current investigation, r^2 values greater than 0.2 were obtained when the distance between adjacent markers was less than 60 kb and 30 kb in Ethiopian cattle breeds and European cattle breeds, respectively.

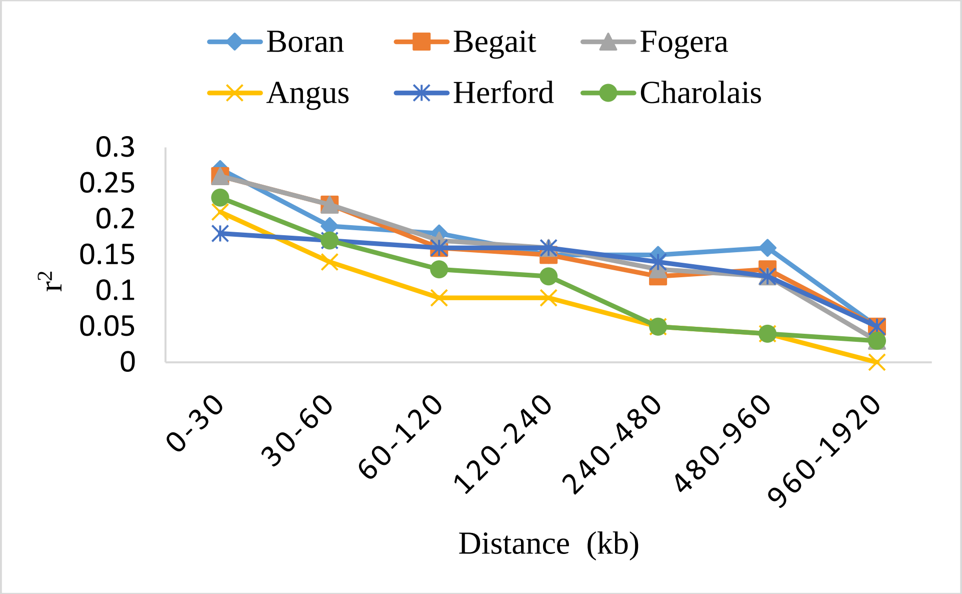


Figure 10: Average of squared correlation coefficient (r^2) value in Ethiopian indigenous cattle breeds and European beef cattle breeds

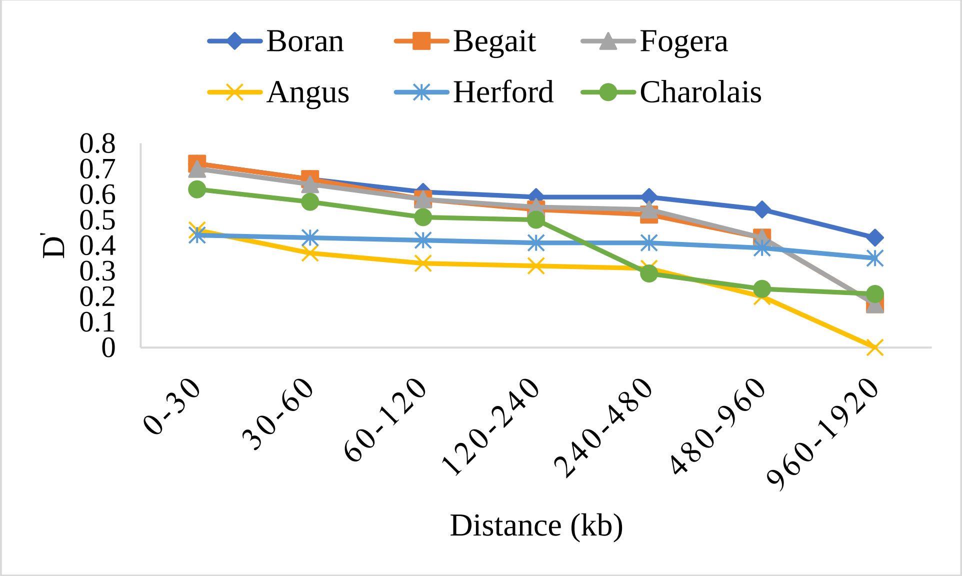


Figure 11: Average of D' value in Ethiopian indigenous breeds and European beef cattle breeds

For each breed/population, the mean r^2 and D' values are presented in Table 10. The mean r^2 values ranged from 0.17 in Angus and Hereford to o.23 in Begait. Similarly, the mean D' values varied from 067 in Boran to 0.42 in Angus.

Table 10: Overall means of linkage disequilibrium in Ethiopian cattle breeds and European beef cattle breeds

LD	Angus	Herford	Charolais	Boran	Begait	Fogera
r^2	0.17 ± 0.28	0.17± 0.28	0.18 ± 0.29	0.22 ± 0.25	0.23 ± 0.25	0.22 ± 0.25
D'	0.42 ± 0.45	0.43 ± 0.45	0.58 ± 0.43	0.67 ± 0.34	0.66 ± 0.34	0.65 ± 0.34

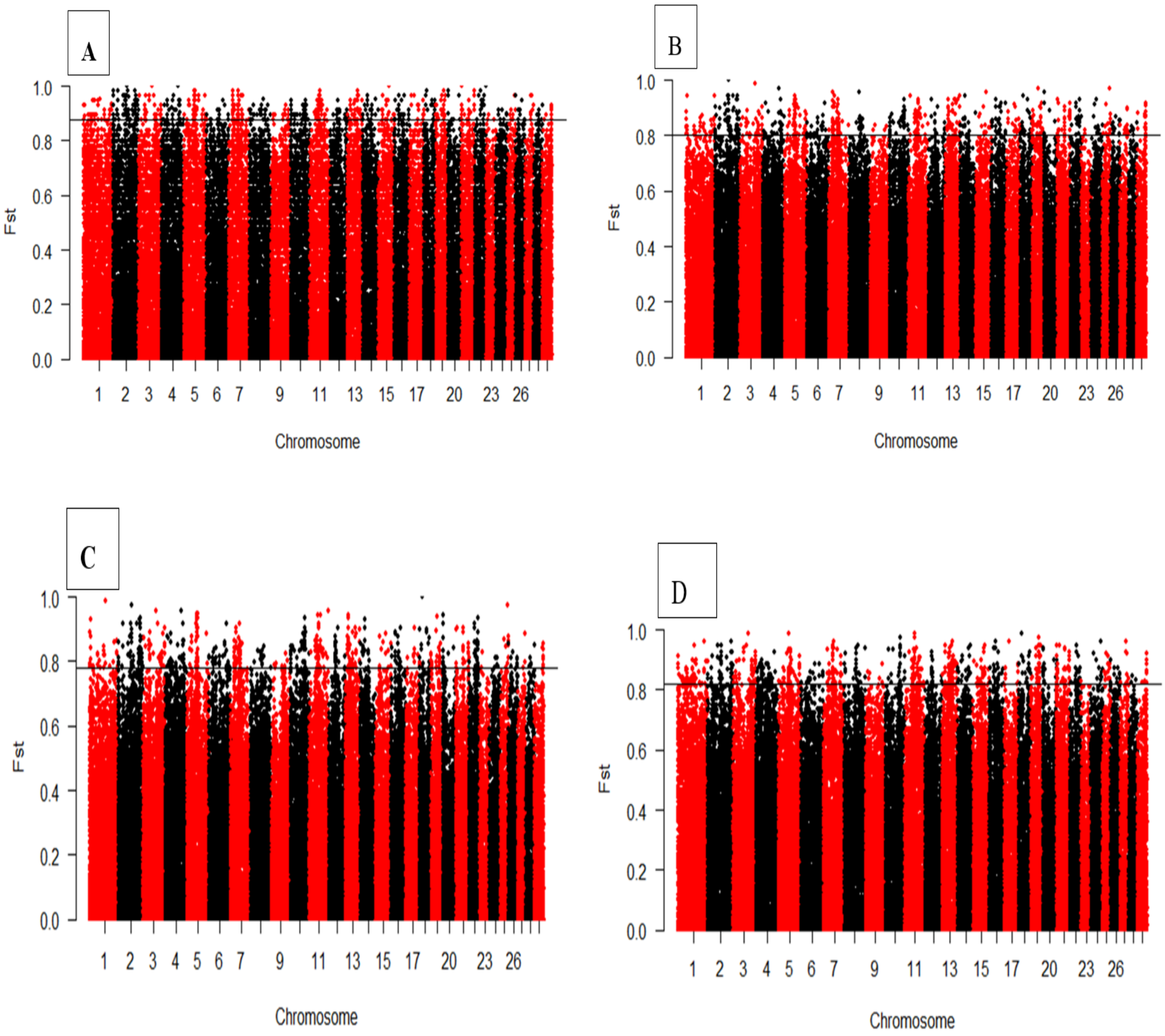
4.5. Genetic differentiation and selection signatures

In this study, the highest genetic differentiation ($FST = 0.315$) was obtained between Angus and Ethiopian cattle populations Boran and Begait. As expected, the lowest genetic differentiation was observed among Ethiopian cattle populations. Moderate genetic FST value was observed among the three European beef cattle breeds.

Table 11: Pairwise genetic differentiation among the six cattle breeds

Breed	Begait	Borana	Fogera	Angus	Charolais	Hereford
Begait	0					
Borana	0.028	0				
Fogera	0.025	0.019	0			
Angus	0.315	0.314	0.297	0		
Charolais	0.271	0.275	0.258	0.083	0	
Hereford	0.296	0.300	0.284	0.115	0.074	0

The pairwise comparison of Ethiopian indigenous breeds and European beef cattle breeds showed very high differentiation (FST value), especially Angus breed showed greater differentiation in all Ethiopian breeds as indicated in Figure 12.



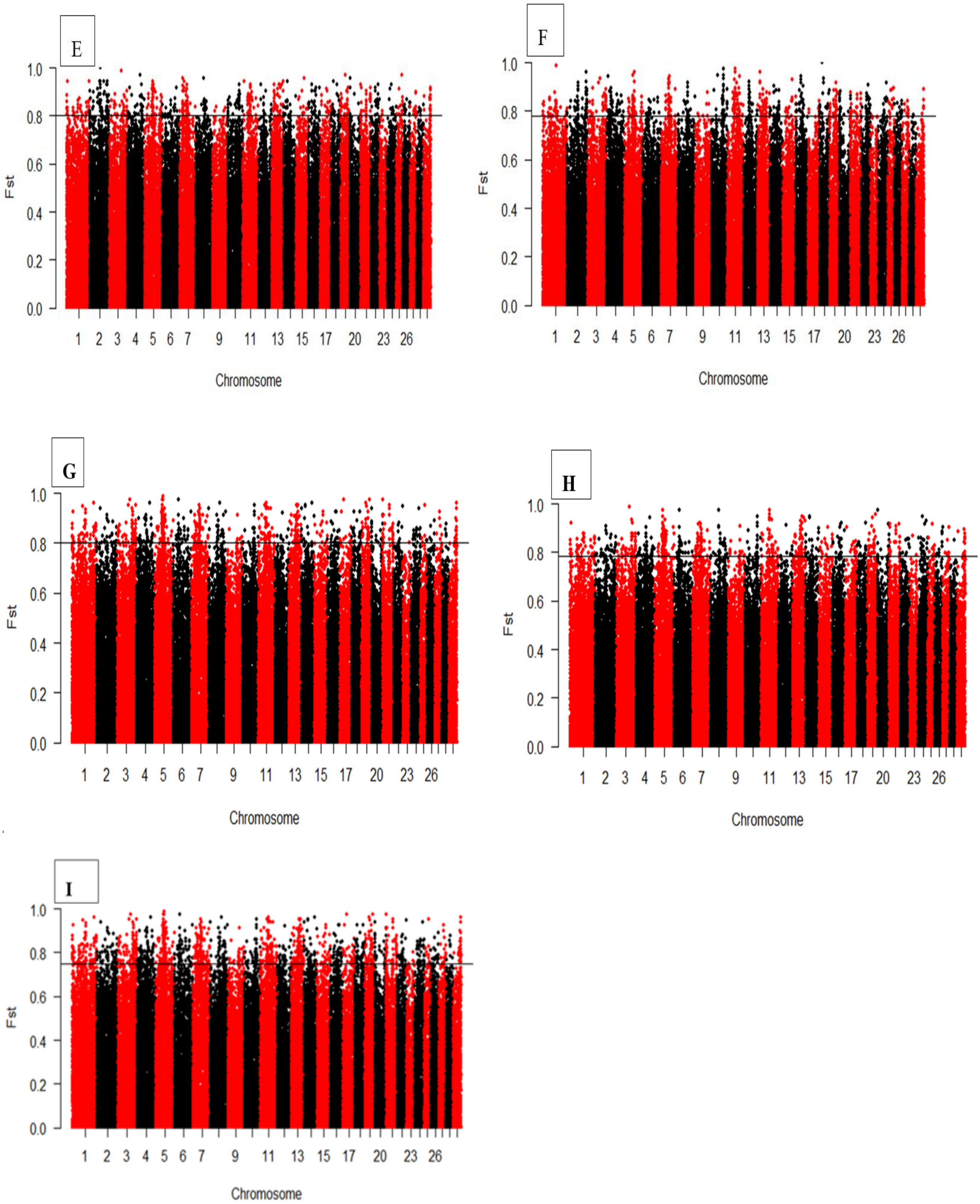


Figure 12: Manhattan plots the genome-wide distribution of FST values for pairwise comparison of Ethiopian indigenous breeds and European beef cattle breeds. (A) Begait-Angus, (B) Begait- Herford, (C) Begait-Charolais, (D) Boran-Angus, (E) Boran-Herford, (F)

Boran-Charolais, (G) Fogera-Angus (H) Fogera-Herford, (I) Fogera-Charolais. The solid dark lines indicate the top 1% FST values considered for downstream analysis

Across the nine pairwise comparisons, a total of around 672 SNPs were detected in each breed (Table 12). The largest number of SNPs (674) was detected in the comparison of Boran with Charolais.

Table 12: Number of SNPs and genes within the top 1% FST values

Breed	Boran	Begait	Fogera	Shared SNPs	
Angus	672	672	672	148	Eighty-three (83) Ethiopian indigenous breed shared SNPs were filtered
Charolais	674	672	672	144	
Herford	673	672	672	168	
Shared SNPs	387	300	361		
One hundred ninety (190) European breed shared SNPs were filtered					

4.5.1. Functional analysis of shared genes

A total of 273 SNPs were annotated and the corresponding genes were associated with biological processes and pathways relevant to meat traits and tropical adaptations.

4.5.1.1. Genes related to meat quality

Based on Gene Ontology (GO) analysis, seven genes (*CAPZB, KIF5A, DCTN2, RGS20, STIM1*, *KLHL3,* and *ITGB1BP1*) were involved in actin cytoskeleton organization and ten genes (*KLHL3, HERC3, NPEPPS, YME1L1, TNIK, KLHL3 PARK2, PPP2R5C, UBE2D3,* and *GLB1L3*) were found to be involved in protein ubiquitination, which is an important stage in protein breakdown. Further pathway analysis revealed genes associated with actin cytoskeleton regulation (*MYLK, PIP4K2C, PDGFRA*) were associated with meat tenderness.

Genes involved in the hydrolysis of phospholipids into fatty acids and phosphatidylinositol, as well as phospholipid and carbohydrates metabolism (*DGAT1, LPCAT1, PLCL2, CTNNA1,* and *B4GALNT1*), were identified as associated with meat intramuscular fat. Five genes (*ASAP3, ADAP1, APBB1IP, RTKN2,* and *STAP2*) involved in the main protein responsible for the red color of beef were also identified.

According to enrichment analysis (Table 13) genes involved in meat tenderness (*ANAPC4, GLI2, YME1L1, DCTN2, KDM1A, FAF1, PARP4, RGS20*), meat color (*ASAP3, ADAP1, APBB1IP, RTKN2, STAP2)*, and meat marbling (*DGAT1, FAR1, B4GALNT1*) were identified.

Table 13: Enrichment analysis of candidate genes

Genes	P-value	Count	Terms
ANAPC4, GLI2, YME1L1, *DCTN2, KDM1A, FAF1, PARP4, RGS20*	0.0099	8	Cell proliferation
ASAP3,ADAP1,APBB1IP,RTKN2,STAP2	0.048	5	pH
DGAT1,FAR1, B4GALNT1	0.013	3	Long-chain fatty-acyl-CoA metabolic process

4.5.2. Candidate genes related to tropical adaptation

Tropical cattle are subjected to a variety of environmental stresses. Our BP analysis revealed that the *SOD1* gene is involved in heat tolerance. *MATR3* gene is associated with fat storage in cattle. *IKBKE* a non-classical *IKK* family member is involved in the control of inflammatory reactions, immune cell activation and proliferation, and metabolic disease.

Further Pathway analysis revealed that genes involved in Vasopressin-regulated water reabsorption, notch signaling, and Rap1 signaling were identified (Table 14). *DCTN2, GNAS,* and *CREB5* genes were identified as involved in Vasopressin-regulated water reabsorption. *CTBP2, DTX3,* and *RBPJ* genes were involved in notch signaling regulation. Five genes (*IGF1, TLN1, RRAS, GNAS,* and *KDR)* were involved in the Rap1 signaling pathway. Four genes (*ASAP3, RAB22A, SH3GLB1,* and *EPS15)* were identified as involved in the endocytosis pathway. *KIT* and *KDR* genes were identified as involved in melanogenesis mechanisms that underlie the modulation of skin and hair pigmentation in animals.

Table 14: Pathway analysis of candidate genes related to adaptation to tropical conditions

KEGG pathway	P-value	Genes
Vasopressin-regulated water reabsorption	0.037	*DCTN2, GNAS, CREB5*
Notch signaling pathway	0.038	*CTBP2, DTX3, RBPJ*
Rap1 signaling pathway	0.049	*IGF1, TLN1, RRAS, GNAS, KDR*
Endocytosis	0.05	*ASAP3, RAB22A, SH3GLB1, EPS15*

The genotypic frequency of the detected genes associated with carcass quality and tropical climate adaptation (Table 15), showed higher variability in Ethiopian indigenous breeds and lower variability in European beef breeds.

Table 15: Functionally analyzed candidate genes for carcass traits and tropical adaptation genotypic frequencies

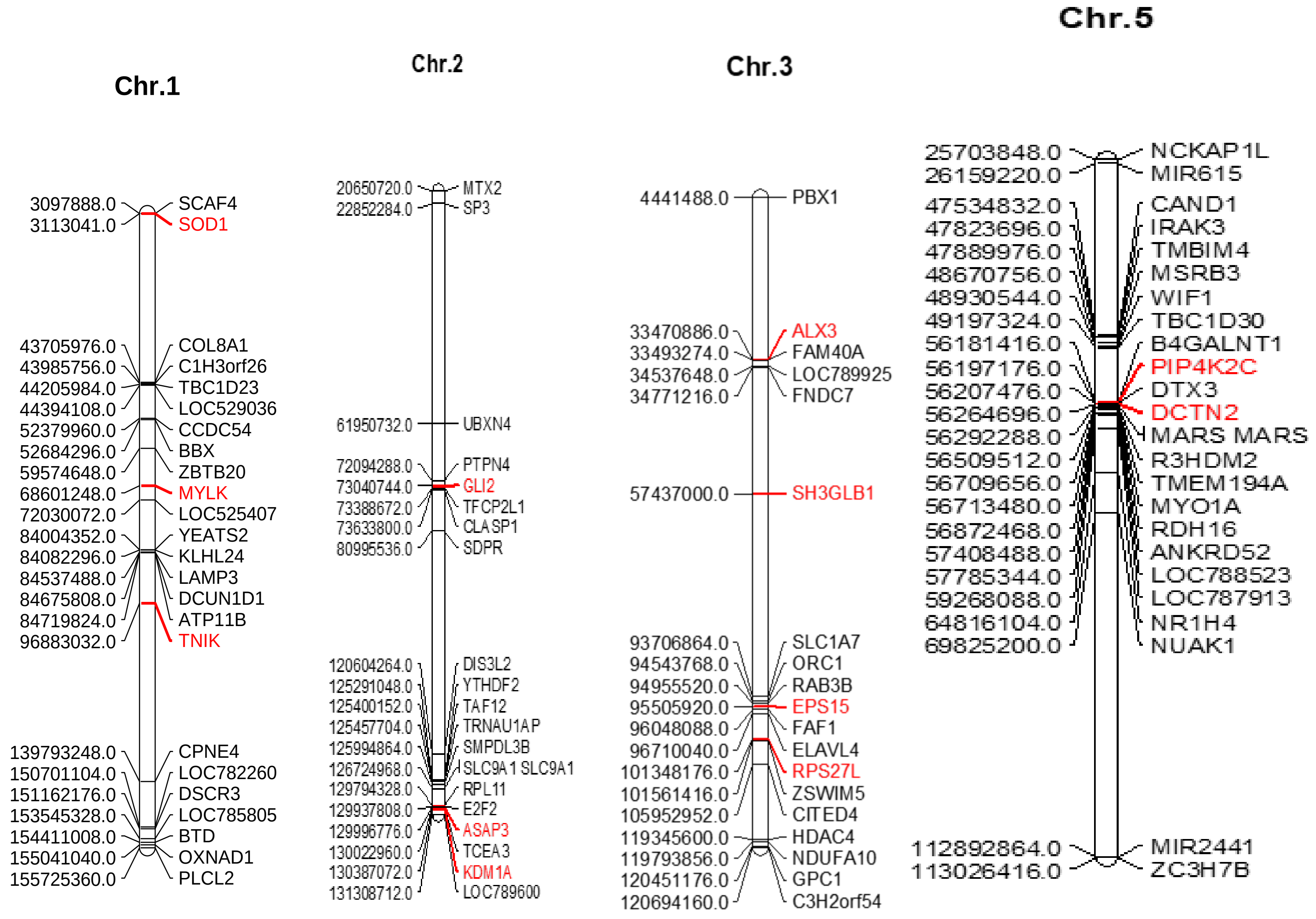
BTA	Position	Genes	SNP IDs	Genotypes	Breeds					
					Begait	Boran	Fogera	Angus	Charolais	Hereford
1	3113041	*SOD1*	BovineHD0100001037	*CC*	0.92	0.69	0.85	0	0	0
				AA	0.03	0.30	0	1.00	0.97	1.00
				AC	0 .05	0	0.15	0	0.03	0
1	68601244	*MYLK*	BovineHD0100019385	*AA*	0	0	0	0.94	0.63	0.78
				GG	0.76	0.72	0.95	0	0.05	0.05
				AG	0.24	0.28	0.05	0.05	0.32	0.17
2	129996771	*ASAP3*	BovineHD0200037782	*AA*	0.82	0.85	0.73	0	0	0
				AG	0.18	0.15	0.27	0	0.03	0.03
				GG	0	0	0	1.00	0.97	0.97
2	130411462	*KDM1A*	BovineHD0200037886	*AA*	0.68	0.64	0.48	0	0	0
				AG	0.27	0.33	0.47	0	0	0
				GG	0.05	0.03	0.05	1.00	1.00	1.00
5	56197180	*PIP4K2C*	BovineHD0500015977	*AA*	0.76	0.65	0.73	0	0	0
				AG	0.24	0.33	0.25	0	0	0
				GG	0	0	0.03	1.00	1.00	1.00
5	56275158	*DCTN2*	BovineHD0500015996	*AA*	0.87	0.72	0.80	0	0	0
				AG	0.13	0.28	0.17	0	0	0
				GG	0	0	0.03	1.00	1.00	1.00
11	88105122	*ITGB1BP1*	BovineHD1100025411	*AA*	0.03	0	02	0.98	0.95	0.89
				AG	0.18	0.24	0.20	0.02	0.05	0.11
				GG	0.79	0.76	0.78	0	0	0

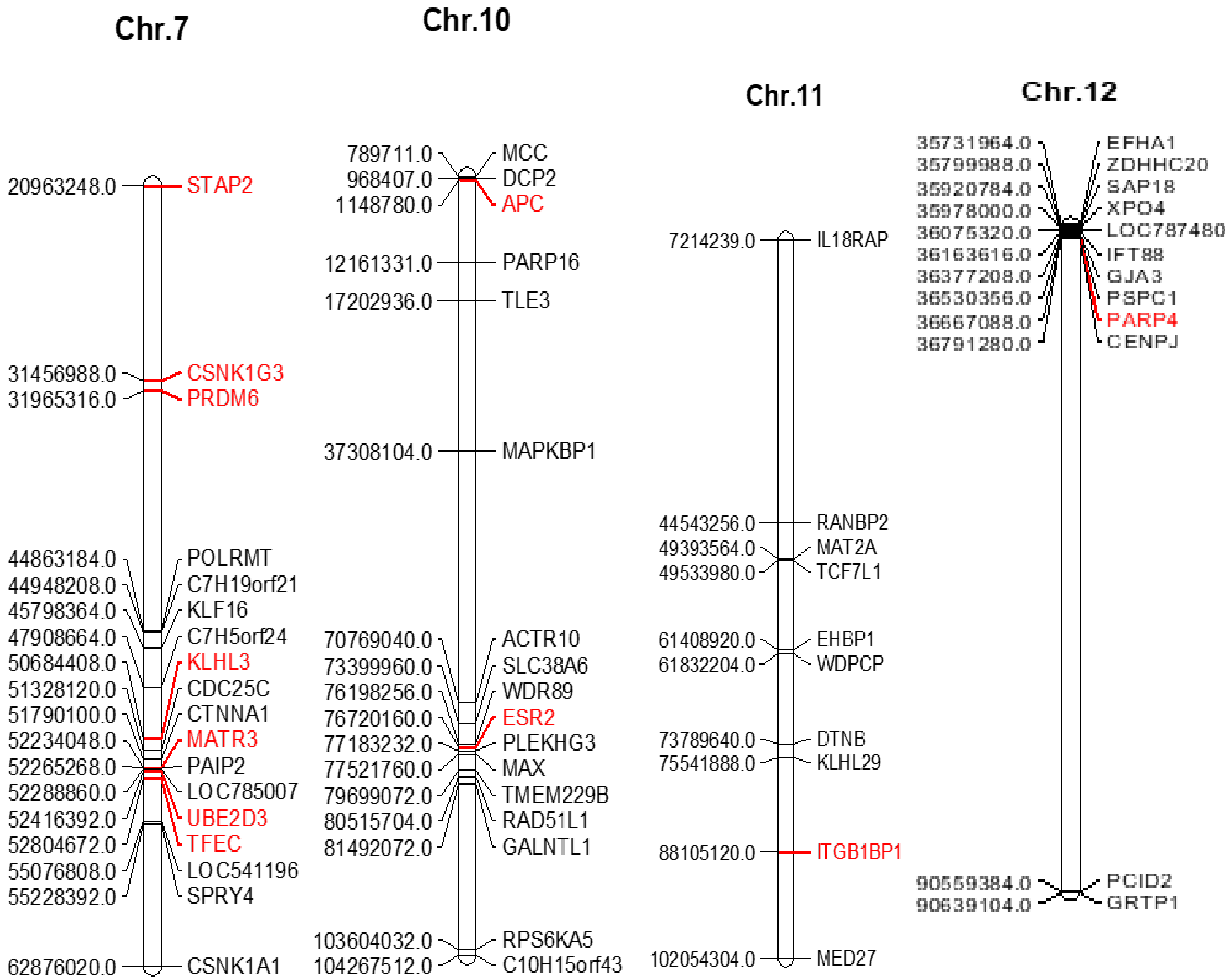
12	36667090	*PARP4*	BovineHD1200010699	*AA*	0.03	0	0.13	1.00	1.00	1.00
				AC	0.18	0.15	0.33	0	0	0
				CC	0.79	0.85	0.55	0	0	0
13	18287520	*APBB1IP*	BovineHD1300005225	*AA*	0.03	0.05	0.05	1.00	1.00	1.00
				AG	0.21	0.28	0.45	0	0	0
				GG	0.76	0.67	0.50	0	0	0
13	58051946	*GNAS*	BovineHD1300016641	*AA*	0.63	0.77	0.85	0	0	0
				AG	0.32	0.20	0.15	0	0.05	0
				GG	0.05	0.03	0	1.00	0.95	1.00
14	1801116	*DGAT1*	ARS-BFGL-NGS-4939	*AG*	0.08	0.17	0.03	0.10	0.19	0.05
				AA	0	0	0	0.90	0.81	0.86
				GG	0.92	0.83	0.97	0	0	0.09
18	56499147	*RRAS*	BovineHD1800016473	*AA*	0	0.08	0.03	1.00	0.92	1.00
				AG	0.26	0.41	0.27	0	0.03	0
				GG	0.74	0.51	0.70	0	0.05	0
19	39519992	*NPEPPS*	BovineHD1900011366	*AA*	0	0	0	1.00	0.89	1.00
				AG	0.29	0.26	0.35	0	0.11	0
				GG	0.71	0.74	0.65	0	0	0
25	42332601	*ADAP1*	BovineHD2500011865	*AA*	0.71	0.54	0.52	0	0	0
				AC	0.29	0.44	0.40	0	0	0
				CC	0	0.03	0.08	1.00	1.00	1.00

4.5.4. Identification of candidate genes QTL

Table 16: QTLs of the candidate genes (SNPs)

Genes	Position	BTA	QTL	References
SOD1	3113041	1	Heat tolerance	Zeng *et al*., 2018
IGF1	66654472	5	Carcass weight	Mullen *et al.,* 2011
ANAPC4	46492439	6	Carcass weight	Bhuiyan *et al.,* 2018
EPS15L1	95505916	7	Carcass quality traits	deLasHeras-Saldana *et al.,*2020
PARP4	36667090	12	Carcass weight	Rouleau *et al*., 2010
GNAS	58051946	13	Carcass weight	Sikora *et al*., 2011
DGAT1	1801116	14	Carcass weight	Ribeca *et al*., 2014


Chr.1
3097888.0 SCAF4
3113041.0 SOD1
43705976.0 COL8A1
43985756.0 C1H3orf26
44205984.0 TBC1D23
44394108.0 LOC529036
52379960.0 CCDC54
52684296.0 BBX
59574648.0 ZBTB20
68601248.0 MYLK
72030072.0 LOC525407
84004352.0 YEATS2
84082296.0 KLHL24
84537488.0 LAMP3
84675808.0 DCUN1D1
84719824.0 ATP11B
96883032.0 TNIK
139793248.0 CPNE4
150701104.0 LOC782260
151162176.0 DSCR3
153545328.0 LOC785805
154411008.0 BTD
155041040.0 OXNAD1
155725360.0 PLCL2
Chr.2
20650720.0 MTX2
22852284.0 SP3
61950732.0 UBXN4
72094288.0 PTPN4
73040744.0 GLI2
73388672.0 TFCP2L1
73633800.0 CLASP1
80995536.0 SDPR
120604264.0 DIS3L2
125291048.0 YTHDF2
125400152.0 TAF12
125457704.0 TRNAU1AP
125994864.0 SMPDL3B
126724968.0 SLC9A1 SLC9A1
129794328.0 RPL11
129937808.0 E2F2
129996776.0 ASAP3
130022960.0 TCEA3
130387072.0 KDM1A
131308712.0 LOC789600
Chr.3
4441488.0 PBX1
33470886.0 ALX3
33493274.0 FAM40A
34537648.0 LOC789925
34771216.0 FNDC7
57437000.0 SH3GLB1
93706864.0 SLC1A7
94543768.0 ORC1
94955520.0 RAB3B
95505920.0 EPS15
96048088.0 FAF1
96710040.0 ELAVL4
101348176.0 RPS27L
101561416.0 ZSWIM5
105952952.0 CITED4
119345600.0 HDAC4
119793856.0 NDUFA10
120451176.0 GPC1
120694160.0 C3H2orf54
Chr.5
25703848.0 NCKAP1L
26159220.0 MIR615
47534832.0 CAND1
47823696.0 IRAK3
47889976.0 TMBIM4
48670756.0 MSRB3
48930544.0 WIF1
49197324.0 TBC1D30
56181416.0 B4GALNT1
56197176.0 PIP4K2C
56207476.0 DTX3
56264696.0 DCTN2
56292288.0 MARS MARS
56509512.0 R3HDM2
56709656.0 TMEM194A
56713480.0 MYO1A
56872468.0 RDH16
57408488.0 ANKRD52
57785344.0 LOC788523
59268088.0 LOC787913
64816104.0 NR1H4
69825200.0 NUAK1
112892864.0 MIR2441
113026416.0 ZC3H7B


Chr.7
20963248.0 STAP2
31456988.0 CSNK1G3
31965316.0 PRDM6
44863184.0 POLRMT
44948208.0 C7H19orf21
45798364.0 KLF16
47908664.0 C7H5orf24
50684408.0 KLHL3
51328120.0 CDC25C
51790100.0 CTNNA1
52234048.0 MATR3
52265268.0 PAIP2
52288860.0 LOC785007
52416392.0 UBE2D3
52804672.0 TFEC
55076808.0 LOC541196
55228392.0 SPRY4
62876020.0 CSNK1A1
Chr.10
789711.0 MCC
968407.0 DCP2
1148780.0 APC
12161331.0 PARP16
17202936.0 TLE3
37308104.0 MAPKBP1
70769040.0 ACTR10
73399960.0 SLC38A6
76198256.0 WDR89
76720160.0 ESR2
77183232.0 PLEKHG3
77521760.0 MAX
79699072.0 TMEM229B
80515704.0 RAD51L1
81492072.0 GALNTL1
103604032.0 RPS6KA5
104267512.0 C10H15orf43
Chr.11
7214239.0 IL18RAP
44543256.0 RANBP2
49393564.0 MAT2A
49533980.0 TCF7L1
61408920.0 EHBP1
61832204.0 WDPCP
73789640.0 DTNB
75541888.0 KLHL29
88105120.0 ITGB1BP1
102054304.0 MED27
Chr.12
35731964.0 EFHA1
35799988.0 ZDHHC20
35920784.0 SAP18
35978000.0 XPO4
36075320.0 LOC787480
36163616.0 IFT88
36377208.0 GJA3
36530356.0 PSPC1
36667088.0 PARP4
36791280.0 CENPJ
90559384.0 PCID2
90639104.0 GRTP1

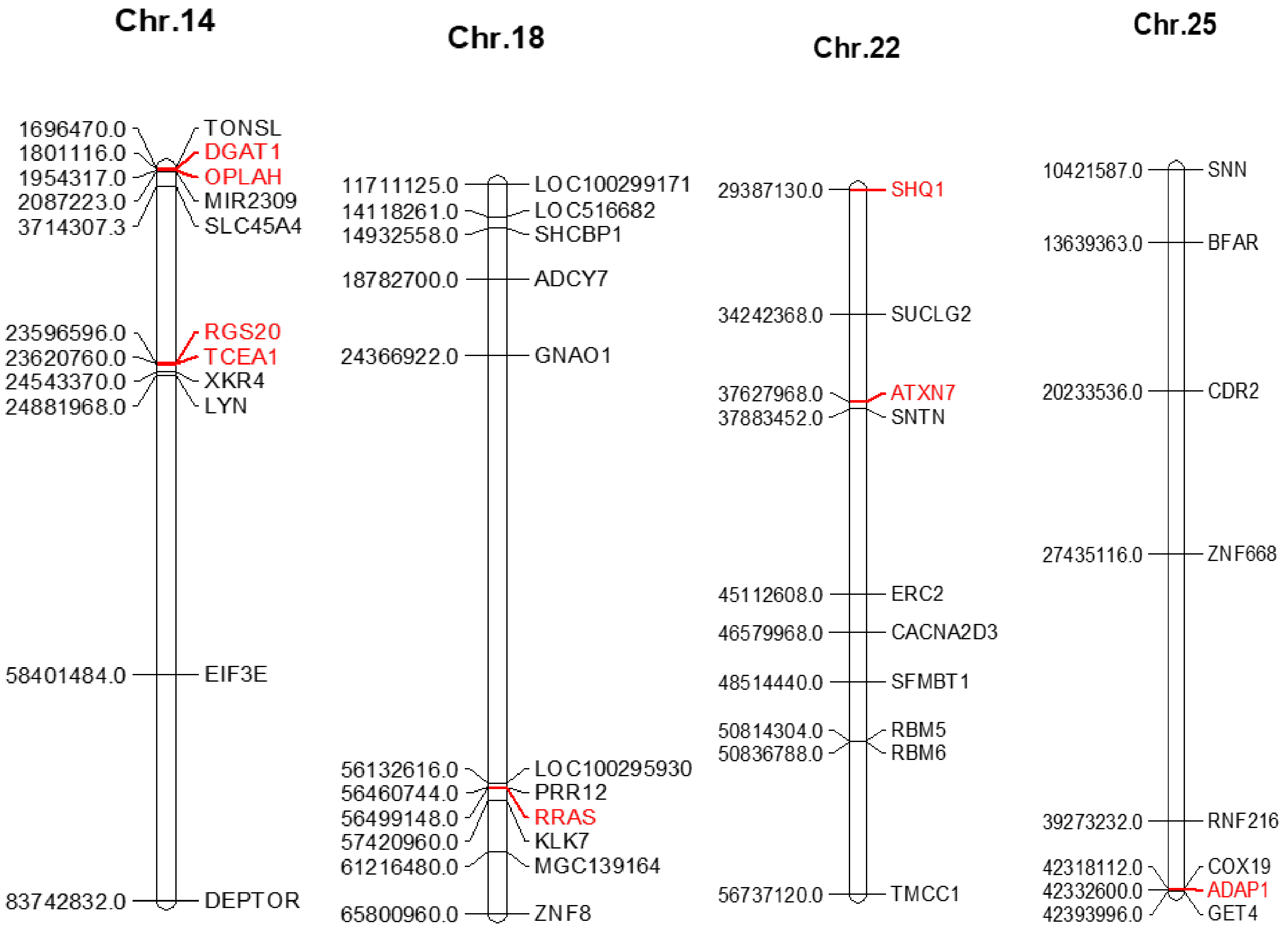

Chr.14
1696470.0 TONSL
1801116.0 DGAT1
1954317.0 OPLAH
2087223.0 MIR2309
3714307.3 SLC45A4
23596596.0 RGS20
23620760.0 TCEA1
24543370.0 XKR4
24881968.0 LYN
58401484.0 EIF3E
83742832.0 DEPTOR
Chr.18
11711125.0 LOC100299171
14118261.0 LOC516682
14932558.0 SHCBP1
18782700.0 ADCY7
24366922.0 GNAO1
56132616.0 LOC100295930
56460744.0 PRR12
56499148.0 RRAS
57420960.0 KLK7
61216480.0 MGC139164
65800960.0 ZNF8
Chr.22
29387130.0 SHQ1
34242368.0 SUCLG2
37627968.0 ATXN7
37883452.0 SNTN
45112608.0 ERC2
46579968.0 CACNA2D3
48514440.0 SFMBT1
50814304.0 RBM5
50836788.0 RBM6
56737120.0 TMCC1
Chr.25
10421587.0 SNN
13639363.0 BFAR
20233536.0 CDR2
27435116.0 ZNF668
39273232.0 RNF216
42318112.0 COX19
42332600.0 ADAP1
42393996.0 GET4

Figure 13: Functionally analyzed candidate genes associated with carcass traits and tropical adaptation physical maps showed the chromosomal location (BTA.1, BTA.2, BTA.3, BTA.5, BTA.7, BTA.10, BTA.11, BTA.12, BTA.14, BTA.18, BTA.22, BTA.25) of the related SNPs; the vertical black lines in chromosomes represent the exact relative region where SNPs were situated. The genes or markers related to carcass quality attributes and tropical adaptation were presented on the right (in red) with their QTL names

CHAPTER FIVE

5. DISCUSSIONS

5.1. Minor allele frequency and genetic diversity

Minor allele frequency refers to the frequency of the second most common allele in a population, and it affects heritability and predictive ability. Multiple studies have shown that MAF affects predictive ability (Zhu *et al.*, 2017). Most SNPs in the genome may only explain a small fraction of the genetic variance of complex traits, and that the MAF of SNPs in a panel may differ from that of causal variants with a lower MAF (Yang *et al.,* 2010). Low-frequency SNPs with significant impact sizes may have better predictive ability for quantitative characteristics (Abdollahi-Arpanahi *et al.*, 2014).

The average MAF for the Ethiopian indigenous cattle populations in this study was higher than the previous report value reported for most taurine breeds (Mustafa *et al.*, 2017; Edea *et al.*, 2012; McKay. *et al.*, 2007). The value obtained for Boran cattle was higher than the previous report value (0.21 ± 0.150) (Edea *et al.,* 2012). The average value obtained for Angus and Hereford were lower than the reported values for the same breeds (Angus = 0.27 ± 0.14; Hereford = 0.29 ± 0.14) (Zwane *et al.,* 2016). This could be due to the difference in genotypic platforms used in the previous study, which used a lower density marker (Illumina Bovine 8K and 50k SNP Bead Chip). The higher values for Ethiopian breeds can be explained by the fact that the SNP loci used in this study were discovered in indicine breeds, and their average MAF was much higher than for taurine Indicine breeds. Interestingly, the minor allele frequency revealed in this analysis corresponded to different markers density (Illumina Bovine 8K, 10K, 50K, 80K, and 700K) used in previous studies in various cattle breeds around the world, with the majority of these breeds samples not being used before or during the design of these chips (Mustafa *et al*., 2017; Edea *et al*., 2014; Uzzaman *et al*., 2014; Chen *et al*., 2010).

Inbreeding is indicated by the higher proportion of alleles (fixed) in selected cattle breeds, which is due to unintentional or intentional breeding management (Mustafa *et al*., 2017). The distribution of SNPs (MAF=0) was also investigated, and an average of 1% for Ethiopian and 23% for the European breed was revealed. The Angus breed had the highest proportion of fixed SNPs (29%), whereas all Ethiopian indigenous breeds had the lowest (1%). These results differ significantly from the reported proportion of fixed SNPs (36%) for Ethiopian cattle populations genotyped on the Illumina 50K Bead Chip (Edea *et al*., 2012). Similarly, the overall MAF in European beef breeds was lower than what was estimated for Ethiopian zebu breeds, which could be due to ascertainment bias in the BovineSNP80K assay design. Selection practices could explain the comparatively higher proportion of fixed alleles in European cattle breeds. It could be further described to ascertainment bias as the chip used in this study was mainly derived from indicine breeds.

The findings of SNP variations observed in this study were strikingly similar to some previously recorded variation in other farm animals such as sheep and goats using genome-wide SNP arrays (Kijas *et al.,* 2009). The observed polymorphism in these selected breeds could explain why the majority of the bovine sequence data used to produce the bead chip came from European cattle breeds (*B. taurus)* (Decker *et al*., 2014; Edea *et al.*, 2014). MAF has been shown to restrict the efficacy of markers in association studies due to the difficult-to-estimate effects of breed-specific alleles (Gurgul *et al*., 2013). While unique and fixed alleles could be used to explain distinct loci in a population, they could also explain why breeds have a lower percentage of informative markers (Dadi *et al.,* 2012). There has been a frequency imbalance between the two allelic groups in SNPs with low MAF, which may indicate functional significance ((Eynard *et al.*, 2015)).

Genetic drift, selection for adaptation to the local climate, or ancient divergence among founder populations could all explain the differences in allele frequencies among the breeds (Dadi *et al*., 2012). The conservation and management of animal genetic resources need an understanding of genetic diversity, genetic distinctiveness, and genetic population structure. One of the most important goals of genetic conservation and animal breeding is to understand the genetic basis of phenotypic diversity. Selection for essential breeding traits would fail if phenotypes and genotypes display no difference between individuals. As a result, maintaining genetic diversity within the global cattle population is critical to provide genetic tools to address potential societal challenges such as food security, increased competition for land, and greater climatic changes.

The overall MAF in European beef breeds was lower than Ethiopian zebu breeds, which may be due to ascertainment bias in the BovineSNP80k assay design. Heterozygosity can be considered as a measure of the amount of genetic variation within a population. This parameter indicates how much the variation exists in the population and how the variation is distributed across the alleles of analyzed markers (Nietlisbach *et al*., 2016). The observed heterozygosity (H_o) is the proportion of heterozygous individuals in population samples and expected heterozygosity (H_e) is the probability of an individual being heterozygous in any locus.

The lower observed heterozygosity in most studied genotypes reflects a low level of diversity within each breed or low levels of outcrossing. The observed and expected average heterozygosity results for the European beef breeds were 0.25, 0.26, respectively. The expected heterozygosity value was greater than the observed heterozygosity, which shows a high level of genetic homozygosity or heterozygosity deficiency. The three European beef breeds showed lower heterozygosity than the reported value based on 777 K SNP data analysis (Kelleher *et al*., 2017).

The mean observed and expected heterozygosity of the Ethiopian indigenous cattle breeds were 0.403 and 0.400, respectively. These values were higher than the previous results ($H_O = 0.314$; $H_E = 0.313$) obtained from 50K Bead Chip analysis (Edea *et al.* 2012). However, it was also lower than the values recorded from microsatellite markers (Dadi *et al*. 2008).

5.2. Principal component analysis

PC1 explains 3.51% of the variance and differentiates Ethiopian Zebu (*B. indicus*) and Zenga (Sanga x Zebu) (represented here by Boran and Fogera respectively) from Begait cattle. PC2 explains 2.82% of the variance and differentiates Boran from Fogera and Begait. The three Ethiopian indigenous cattle breeds were grouped according to their geographical distribution, insighting that these breeds have not experienced recent admixture and might have been exposed to different selective pressures and demographic effects. Such breeds could be used directly for genetic conservation and pure-bred genetic improvement programs.

The next Plot explains the variance and differentiation of the three reference European *B. turuas* from East African breeds, Boran, Begait, and Fogera. PC1 explains 70.63% of the variance and differentiates European *B. turuas* from East African breed. PC2 explains 7.27% differentiates the three European *B. turuas* (i.e., Angus, Herford, and Charolais). The Angus cattle breed was genetically distinct from Hereford and Charolais. Geneflow was observed between Charolais and Herford breeds Ethiopia's indigenous breeds showed closely cluster this may have shared a common ancestor.

5.3. Breed-specific SNPs

To assess the protection and authenticity of livestock products in global and domestic markets, product validation has become increasingly important. When an SNP had an allele that was found in only one breed, it was considered breed-specific (Ramos *et al.*, 2011). Breed-specific SNPs were only polymorphic within a breed, and one of the alleles is fixed in other breeds

(Pant *et al.,* 2012). The breed-specific SNPs detected in this study could be used for breed identification, product discrimination, and traceability.

Breed-specific SNPs that can be used to distinguish within and between Ethiopian indigenous cattle breed and European beef cattle breed SNPs identified and from Ethiopian breeds, the lowest value was scored in Fogera and the highest in Boran and this result showed the presence of variability between the Ethiopian cattle populations.

5.3.1. Biological process of candidate genes associated with Ethiopian cattle-specific SNPs

Cellular response to heat (GO:0034605), any process that causes a change in the state or activity of a cell as a result of a heat stimulus, a temperature stimulation above the ideal temperature for that organism) (Balakrishnan *et al.,* 2013). Ethiopian cattle such as Borana and Begait are evolved under arid and semi-arid environments (high ambient temperature, recurrent drought). Interestingly, Ethiopian cattle-specific SNPs were localized with genes known to be associated with heat stress. Accordingly, the candidate genes (*ST8SIA1, ANO1, C27H8orf4, FGF1, HSF1, MYOF,* and *SCARA5*) have a biological process of cellular response to heat. *HSF1* gene selected for adaptation in subtropical climate (Baena *et al*., 2018). Cellular response to forskolin (GO:1904322), as a result of a forskolin stimulation, any process that causes a change in the state or activity of a cell (in terms of motility, secretion, enzyme synthesis, gene expression, and so on). Genes implicated in the forskolin response in cells *EPHA5, GNAI1, ADCY1, ADCY3*. Pituitary gland development (GO:0021983), the progression of the pituitary gland over time from its initial formation until its mature state. The pituitary gland is an endocrine gland that secretes hormones that regulate many other glands. Genes involved in pituitary gland development were *GATA2, GLI2, TBX19, BMPR1A, KDM1A, PAX6.* Cell fate commitment (GO:0045165), cell commitment to specific cell fates, and their ability to differentiate into specific cell types.

Protein signals from a localized source within a cell (the initial one-cell zygote) or within a developing field are used to establish positional information. Genes involved in cell fate commitment were *GATA2, GATA4, TBX19, WNT1, WNT10A, WNT16, WNT3A, WNT7A, BMP2, FGFR2, GAP43, ONECUT1, ROR2, SPRY2, TRPS1*. Regulation of pH (GO:0006885), within an organism or cell, any activity involving the preservation of internal equilibrium of hydrogen ions, hence regulating the internal p^{H} and the candidate genes involved in the regulation of p^{H} were *ATP12A*, *EDNRB*, *SLC9A1*, *SLC9A9*. Cell morphogenesis (GO:0000902), the process through which a cell's size or form is determined and organized during development. Genes involved in cell morphogenesis were *CAP2, FRY, NOX4, SOX6, SS18, TBCCD1*, *CAPZB*, *COL4A3BP*, *DMRT1*, *EGFR*, *IL7R*, *MAEL*, *NRG1*, *STK4*. Energy homeostasis (GO:0097009), any procedure that involves balancing food intake (energy input) and energy expenditure. Genes involved in energy homeostasis were *ACACB*, *AMPD2*, *LEPR, MRAP2, NR4A3*. Positive regulation of Ras protein signal transduction (GO:0046579) Any procedure that triggers or increases the frequency, rate, or scope of Ras protein signaling. Genes involved in the positive regulation of Ras protein signal transduction were *KITLG, RASGRP1, SHOC2*, *IGF1*, *MMD2*, *NRG1*, *NOTCH2*. Cellular potassium ion homeostasis (GO:0030007) potassium is the most abundant cation in the intracellular fluid, and maintaining the proper distribution of potassium across the cell membrane is critical for normal cell function. Cellular potassium ion homeostasis is any procedure for maintaining an internal steady state of potassium ions at the cellular level. Genes involved in cellular potassium ion homeostasis were *ATP12A*, *ATP1A1*, *ATP1B1*, *KCNMA1*.

5.3.2. Pathway analyses of candidate genes associated with Ethiopian cattle specific SNPs

Melanogenesis is the complex process by which melanocytes produce the pigment melanin in melanosomes. Our study cattle breeds are characterized by diverse coat colors. European cattle breeds have been strongly selected for a specific color. Both breeders and researchers have been interested in the color of cattle's coats because genes that control pigmentation have economic ramifications in the event of genetic abnormalities (Seo *et al.,* 2007). The melanocortin receptor 1 is recognized to be the principal regulator of the switch between the two-coat color pigments: eumelanin (black pigment) and phaeomelanin (white pigment) in cattle (red pigment) (Gutiérrez-*Gil et al.,* 2007). Ethiopian cattle population breed-specific candidate genes, *GNAI1, GNAQ, HRAS, KITLG, WNT1, WNT10A, WNT16, WNT3A, WNT7A, ADCY1, ADCY3, ADCY8, CREB3L2, CREB3L4, CREB3, CAMK2A, CAMK2B, CAMK2D, EDNRB, FZD3, FZD8, LEF1, MAPK1, PLCB1, PLCB2, PLCB4, TCF7* were identified as involved in melanogenesis path way. KITLG gene selected for roan hear pigment (Talenti *et al.,* 2018).

Insulin is a metabolically active hormone that contributes to growth, development, and nutritional homeostasis through a variety of anabolic mechanisms. Food restriction in cattle is often accompanied by a decrease in basal insulin concentrations, followed by a rise in systemic insulin concentrations following re-feeding and compensatory growth, reflecting the availability of dietary substrate and hepatic gluconeogenesis (Keogh *et al*., 2015). The candidate genes, (*ADCYAP1R1, ATP1A1, ATP1B1, GNAQ, RAPGEF4, ADCY1, ADCY3, ADCY8, CREB3L2, CREB3L4, CREB3, CREB5, CACNA1D, CAMK2A, CAMK2B, CAMK2D, GLP1R, PLCB1, PLCB2, PLCB4, KCNMA1, KCNMB1, KCNN1, KCNN3, KCNU1, and SNAP25)* were involved in insulin secretion pathway.

TLR gene family members play a key role in the identification of pathogen-associated molecular patterns from bacteria, viruses, protozoa, and fungus in mammals. Using a radiation hybrid panel, cattle TLR genes were recently localized to chromosomes. The nucleotide sequences of the TLR2, TLR4, and TLR6 genes in cattle were searched for novel SNPs that could be exploited in disease resistance research (Mariotti *et al.*, 2009). Genes, *AKT3, FOS, RELA, TBK1, TRAF6, CTSK, IKBKB, IL12A, LBP, MAPK1, MAPK10, MAPK9, MAP2K6, MAP3K8, PIK3CB, PIK3R2, RIPK1, SPP1* were identified as novel SNPs involved in the Toll-like receptor signaling pathway. Platelet activation signaling is important for platelet function in hemostasis. Genes involved in Platelet activation were *AKT3, GNAI1, GNAQ, LYN, RASGRP1, ADCY1, ADCY3, ADCY8, APBB1IP, COL1A2, COL11A1, GUCY1A2, ITPR2, ITGB1, MAPK1, MYLK4, MYL12B, PIK3CB, PIK3R2, PLCB1, PLCB2, PLCB4, PRKG1, PPP1CB, P2RY1, TLN1, TLN2, TBXAS1.* Calcium signaling is a key early aspect in immune cell activation, the increased demand for calcium in periparturient cattle could negatively affect immune cell intracellular calcium reserves. This decrease in intracellular calcium reserves in immune cells may reduce intracellular calcium release in response to an activating stimulation, contributing to the immunological suppression (*Kimura et al*., 2006). The genes *ATP2B1, ATP2A2, GNAL, GNAQ, ADCY1, ADCY3, ADCY8, ADRA1D, ADRB3, AGTR1, AVPR1A, CACNA1B, CACNA1D, CACNA1H, CAMK2A, CAMK2B, CAMK2D, DRD5, EDNRB, EGFR, ERBB3, ITPR2, LHCGR, MYLK4, PLCB1, PLCB2, PLCB4, PLCD3, PLCE1, PLCZ1, PHKB, PPP3CA, PPP3CC, PTK2B, P2RX3, SLC8A1, SLC8A3, VDAC3* were identified as involved in the Calcium signaling pathway. Rap1 is a tiny cytosolic protein that functions as a signal transduction switch. Rap1 inhibits Ras signaling by including an effector domain that is similar to that of Ras. Rap1 has been reported in various animal species to fulfill different roles as an evolutionarily conserved protein (Dong *et al.*, 2012). The identified genes involved in the Rap1 signaling pathway were *AKT3, EPHA2, GNAI1, GNAQ, HRAS, KITLG, MET, RAPGEF4,*

RAPGEF5, TIAM1, ADCY1, ADCY3, ADCY8, APBB1IP, EFNA3, EGFR, FGF1, FGF18, FGFR1, FGFR2, FGFR3, FLT1, IGF1, ITGB1, ITGB2, MAGI3, MAPK1, MAP2K6, PARD3, PIK3CB, PIK3R2, PLCB1, PLCB2, PLCB4, PLCE1, PDGFD, P2RY1, SIPA1L1, SKAP1, TLN1, TLN2. Endocytosis is a cell's method for removing ligands, nutrients, plasma membrane proteins, and lipids from the cell surface and transporting them into the cell interior. *RF1, ARFGEF1, ARAP2, CBLB, CBL, GRK5, GRK7, HRAS, RAB10, RAB11FIP3, RAB22A, SH3GL2, TRAF6, AMPH, BIN1, CAPZA2, CAPZB, CHMP4B, CCDC53, DNM3, EGFR, FGFR2, FGFR3, HSPA2, IL2RA, KIF5A, KIF5B, LDLRAP1, PARD3, PIP5K1B, SNX6, TFRC, TGFBR1, VPS36, VTA1* genes were detected as involved in Endocytosis pathway.

5.3.3. Functional analysis of annotated European beef cattle specific SNPs

The identification of functional variants in European beef cattle, such as missense variants and variants inside upstream and downstream genic regions, would allow these variations to be tested for their impacts on complex characteristics. Some of the European cattle-specific SNPs corresponding genes were: - *LOC618554* gene associated with the olfactory transduction pathway. As previously mentioned, these genes have been identified as significant genetic factors affecting mammalian evolution and adaptation. Sensing changes in the environment is vital for survival. Animals from invertebrates to vertebrates employ both visual and olfactory inputs to direct survival behaviors including detection of food sources, finding mates, and predator avoidance (Fernandes *et al*., 2020; Abbas and Vinberg, F. 2021). The *C1H21orf62* gene is a poll locus mutant that is a candidate for longhorn growth. The poll locus is responsible for the lack of horns (Liu *et al*., 2013). At both taurine breeds and taurine–indicine crosses, the POLL locus has been localized to the centromeric region of bovine chromosome 1 (BTA1) in an interval of around 1 Mb. In pure-bred zebu cattle, it has not yet been mapped (Mariasegaram *et al*., 2012). *DCBLD2* has been linked to the epidermal growth factor receptor, a tumor suppressor, vascular repair and angiogenesis, as well as glucose uptake and thrombus formation (He *et al*., 2020). The *GALNTL6* gene is linked to growth (Seabury *et al*., 2017). *HACE1* is linked to host defenses against pathogens. The regulation of apoptosis, phagocytosis, as well as the development of reactive oxygen species and inflammatory mediators (Torrino *et al*.,2011). *MGRN1* is an essential component of the homeostasis control system in neurons, acting on two levels: in the cytoplasm, directly challenging protein aggregates and toxic stress, and in the nucleus, achieved at the expense of its ‘‘cytoplasmic" feature, through localization inactive chromatin regions to potentiate the cellular response to proteotoxicity (Benvegnu *et al*., 2017). *MIR34A* is a possible immune cytokine regulator that is expressed in heat stress. In chickens, *MIR34A* targets the genes cytokine–cytokine receptor interleukin 2 (IL-2) and interleukin 12 (IL-12) (Li *et al*., 2020). Ciliaogenesis and Hedgehog signaling pathways are regulated by *RFX3*, which are linked to ciliopathies, which are developmental and degenerative disorders (Chen *et al*., 2018). The *SCN2A* gene controls the voltage-gated sodium channel and keeps the cell's physiology in check (Lena and Mantegazza, 2019). The *Rad54B* gene regulates DNA damage and repair, and its levels of activation are regulated in response to genotoxic stresses (Yasuhara *et al*., 2014).

5.4. Genetic distance and signatures of selection

Pairwise comparison of genetic distance among the six populations ranged from 0.019-0.315. Great differentiation (0.315) was observed between Angus (*Turuas*) and Begait (*Indicus x Turuas*) populations, which is expected. This could be because of parental history and low genetic material exchange between the populations. Relatively low genetic distances were also observed between Ethiopian cattle populations (0.019) between Fogera (*Zebu X Sanga*) and Boran (*Zebu*).

These low values genetic distance indicates may be due to gene flow and shared ancestry. Generally, the level of genetic differentiation between populations increases with increasing geographic distance (Deng *et al*., 2020).

5.5. Biological process and pathway analysis of candidate genes related to carcass

Following domestication, cattle have been subjected to different selective pressures and distributed throughout the world covering various agro-ecologies and production environments. The European cattle breeds included in this study have been highly selected and improved for beef traits, while Ethiopian cattle breeds have been less selected artificially for production traits. Genome-wide analysis of these breeds can aid to better understand the impacts of selection and differences in genomic structure.

Results from the whole genome scan revealed several positively selected genes involved in different biological and cellular functions including those affecting meat quality characteristics. Meat quality is a multifactorial and complex character that is determined by several factors at different levels, from molecular to mechanical.

5.5.1. Genes related to meat tenderness

At the molecular level, several cellular pathways have been involved in meat quality characteristics, including muscle growth, glycolysis, muscle contraction, stress reaction, cell cycle, proteolysis, protein ubiquitination, and apoptosis. According to GO keywords, actin cytoskeleton organization represented by eight genes (*CAPZB, DCTN2, KIF5A, DCTN2, RGS20, STIM1, KLHL3*, and *ITGB1BP1*) were linked to meat tenderness (Guillemin *et al*., 2011, Gao *et al*., 2011) and elven genes (*KLHL3, HERC3, NPEPPS, YME1L1, TNIK, KLHL3 PARK2, PPP2R5C, UBE2D3, GLB1L3,* and *SOD1*) were found to be involved in protein ubiquitination. Ubiquitination is an important stage in protein breakdown (Jiang *et al*.,2010).

The ubiquitination pathway affects muscle qualities that are important for postmortem meat quality, such as softness (Hamill *et al.,* 2012). Negative regulation of actin filament depolymerization and negative regulation of protein complex disassembly are GO keywords that describe how adipocytes are controlled (Gao *et al.,* 2011).

5.5.2. Genes related to meat intramuscular fat

Intramuscular fat (IMF), a heritable meat quality trait, has an impact on taste, juiciness, appearance, and meat tenderness. Genes *DGAT1, LPCAT1, PLCL2,* and *B4GALNT1* are involved in the hydrolysis of phospholipids into fatty acids and phosphatidylinositol, as well as phospholipid and carbohydrates metabolism intramuscular fat, according to the pathway analysis (Roux *et al.,* 2015). The metabolism of lipids has been linked to carbohydrates. Glucose levels influence the formation of fatty acids in the liver as well as the quantity of cholesterol or lipid in the blood. The *CTNNA1* gene has been linked to the degree of myostatin expression in the skeletal muscle of Holstein-Friesian bulls (Sadkowski *et al.*, 2008). Myostatin is a critical protein that regulates skeletal muscle growth and is thought to be one of the most critical elements in cattle meat productivity.

5.5.3. Genes related to meat color

The color of the meat and its ability to hold water are two consistency measures that are considered indicators of freshness and wholesomeness (Joo *et al.,* 2013). These traits are linked to variations in glycolysis rate and muscle temperature drop after death. The main protein responsible for the red color of beef is a globular single-chain protein found in the sarcoplasm and represented by five candidate genes association (*ASAP3, ADAP1, APBB1IP, RTKN2,* and *STAP2*). myoglobin serves as a secondary source of oxygen and assists in the delivery of oxygen within muscles (Joo *et al*., 2013). Meat pigmentation is associated with Myoglobin.

The darker the meat, the higher the concentration of myoglobin. Exercise, the animal's nutrition, genetics, and environmental conditions all have an impact on myoglobin content. The brilliant red hue of red meat, which is related to a high level of oxymyoglobin, is a positive indicator of quality, whereas the myoglobin concentration in brown meat is a negative indicator (Joo *et al.*, 2013). Muscle glycogen (stored energy) live animal p^H =7.1 a conversion to lactic acid adequate levels will result in a p^H level lowered. The more glycogen there is, the more lactic acid will be produced the lower p^H the darker color meat.

5.5.4. Enrichment and biological process analysis using a gene-to-gene similarity matrix

Based on mutual functional annotation, the DAVID 6.8functional clustering annotation method classifies closely related genes into functionally related groupings. Each functional gene cluster should include a list of shared 'consensus words, a display of enriched terms. Meat tenderness is linked to actin cytoskeleton organization, actin filament-based processes, and protein ubiquitination, whereas adipocyte regulation is linked to the cellular component organization, negative regulation of actin filament depolymerization, and negative control of protein complex disassembly. Meat tenderness is improved by the GO term biological process involved in cell growth. Nine genes (*APC, GLI2, YME1L1, DCTN2, KDM1A, FAF1, PARP4, ORC1*, and *RGS20*) are involved in cell proliferation and have a beneficial impact on meat tenderness (Chang K. 2007).

5.5.5. Candidate genes related to tropical adaptation

Tropical cattle are subjected to a variety of environmental stresses, including hot and humid weather, limited feed and water supplies, diseases, and parasites (Porto *et al.*, 2014).

The *SOD1* gene is involved in heat tolerance (Zeng *et al.*, 2018). Heat stress is the most common cause of oxidative stress, as it causes mitochondrial oxidative stress and cell malfunction, which leads to cell death and damage.

Cell survival in stressful situations necessitates rapid response mechanisms and, as a result, effective resumption of cell functioning when stress has been alleviated. When cells are exposed to heat stress, molecules are produced that are ready to mediate cell death and survival signals, as well as assist the cell's tolerance and/or recovery from damage (Zeng *et al*., 2018). The *SOD1* gene is identified as involved in heat tolerance in tropical breeds and the genotypic frequency result shows AA genotype fixed in European beef breeds, whereas the CC genotype was the most frequent in Ethiopian cattle populations. It was determined that the *MATR3* gene is associated with fat storage in cattle. This gene regulates insulin sensitivity and obesity susceptibility (Akakabe *et al*., 2013). Furthermore, the protein encoded by this gene is primarily found in endothelial cells and blood vessels. Angiogenesis is a physiological process in which pre-existing vessels give rise to new blood vessels. Endothelial cells play an important role in this process. Vasodilation is the dilation of blood vessels to dissipate heat to the environment.

IKBKE is involved in several signaling pathways, including the activation of pro-inflammatory signaling pathways by Toll-like receptor (TLR) at the beginning of immune responses against pathogens. Tri-acyl lipopeptides from bacteria or mycobacteria are ligands for *TLR 1,* and Di-acyl lipopeptides from mycoplasma are ligands for *TLR 6,* but the ligand for *TLR10* is unknown (Yin *et al.*, 2020).

5.5.6. Pathway analysis candidate gene related to adaptation to tropical conditions

Pathway analysis identified *DCTN2, GNAS,* and *CREB5* genes as involved in Vasopressin-regulated water reabsorption. Terrestrial animals have evolved a delicate and diverse system to maintain their water homeostasis, thanks to vasopressin-regulated water reabsorption. The antidiuretic hormone vasopressin is released from the pituitary in situations of hypernatremia or hypovolemia and binds to its type-2 receptor in renal main cells (Fukuoka *et al*., 2020). Notch signaling is regulated by the genes *CTBP2, DTX3*, and *RBPJ.*

Through the regulation of a range of cell fate determination of vascular endothelial cells, regulation of arterial differentiation, and angiogenesis, involvement in Notch signaling induced by classical Notch ligands is important for tissue homeostasis. Angiogenesis is the physiological process by which pre-existing blood vessels give rise to new blood vessels (Akil *et al.*, 2021). The genes *IGF1*, *TLN1, RRAS, GNAS*, and *KDR* have a substantial effect on the Rap1 signaling pathway. These pathways are involved in a variety of key cellular functions, including cell adhesion and cell junction information and regulation, cell migration, polarization, and cell proliferation and survival (Van Hooren *et al.,* 2012).

The genes *ASAP3, RAB22A, SH3GLB1*, and *EPS15* are all involved in the endocytosis pathway. Endocytosis is the mechanism by which cells transport items into the cell that are too big to pass through the cell membrane's lipid bilayer. Multiple kinds of plasma membrane invaginations regulate this pathway in mammalian cells, each with a different biological function, composition, and cargo recruitment. A variety of stressful situations, such as changes in osmolality, oxygen, or food delivery, pose a threat to cellular viability. As a result, cells have evolved sophisticated stress mechanisms to deal with these difficulties. Some of these stress responses, such as the heat shock response, are well understood (Lopez *et al.,* 2020).

5.6. Linkage disequilibrium

D' and r^2 are two alternative LD measures. D' represents historical recombination via allelic association, whereas r^2 measures the squared correlation coefficient between locus allele frequencies (Bohmanova *et al*., 2010). These measurements have a range of 0 to 1. D' = 1 denotes the absence of recombination between the two loci due to the presence of one of the polymorphisms, whereas D' < 1 denotes the presence of historical recombination between the loci. As a result, rather than being a true estimate of LD, D' is a signal of missing haplotypes.

For association studies, dealing with the r^2 value is recommended since there is a simple inverse relationship between r^2 and the sample size required to detect the association between QTL and SNP (Bohmanova *et al.*, 2010). The r^2 value indicates the degree of correlation between the two loci; it equals 1 only if two haplotypes are present. A recent study found that in all populations, the amount of LD available for association analysis does not exceed 960 kb.

The overall average of r^2 estimates of LD in Ethiopian cattle populations was lower than the previous SNP-based studies. The overall r^2 values obtained for European beef breeds were far lower than those obtained for Angus cattle *(*r^2 = 0.25) Porto-Neto *et al.*, (2014) based on 50. However, the average r^2 values for Ethiopian cattle were higher than those reported for the Nellore cattle breed (r^2 = 0.17; D' = 0.52) (Espigolan *et al.*, 2013). According to Meuwissen *et al*. (2016), genomic selection requires an r^2 of 0.20 to reach an accuracy of 0.85 for genomic breeding values. The most practical value of r^2 for association studies is 0.25. Brien *et al.,* (2014) showed that in genome-wide association analyses, r^2 values greater than 0.3 were necessary to provide appropriate power (GWAS). Hence, 80K chip is more informative for GWAS and genomic selection in Ethiopian cattle populations.

6. CONCLUSION AND RECOMMENDATION

The levels of genetic variation for SNPs on the Bovines GGP-80K assays obtained in this study indicate that these assays have utility for genetic studies in Ethiopian indigenous cattle breeds. The higher breed-specific SNPs detected in Ethiopian indigenous breeds show the presence of high genetic variability. This genetic variation in Ethiopian cattle breeds is used as a potential source for future breeding programs. The higher LD in Ethiopian indigenous cattle populations shows that the chips are biased to Ethiopian breeds. The genetic differentiation between Ethiopian and European beef breeds coincides with existing knowledge, that European beef breeds considered in this study have been strongly artificially selected, while Ethiopian indigenous breeds are naturally selected. The identified candidate genes have to be confirmed through GWAS, sequencing, and omics studies to include them in cattle genetic improvement programs. Most estimates of genetic variability are biased to the reference genome. Hence a reference genome sequence database has to be developed for indigenous breeds.

7. REFERENCES

Abdollahi-Arpanahi, R., Nejati-Javaremi, A., Pakdel, A., Moradi Shahrbabak, M., Morota, G., Valente, D., Kranis, A., Rosa, J., Gianola, D. (2014). Effect of allele frequencies, effect sizes, and several markers on prediction of quantitative traits in chickens. *J. Anim. Genet.***131**: 123–133.

Abbas, F., and Vinberg, F. (2021). Transduction and Adaptation Mechanisms in the Cilium or Microvilli of Photoreceptors and Olfactory Receptors from Insects to Humans. *Fronter,* **15**: 84-92.

Akakabe, Y., Koide, M., Kitamura, Y., Matsuo, K., Ueyama, T., Matoba, S., Ikeda, K. (2013). Ecscr regulates insulin sensitivity and predisposition to obesity by modulating endothelial cell functions. *Nature,* **4**:1-13.

Akil, A., Gutiérrez-García, A. K., Guenter, R., Rose, J. B., Beck, A. W., Chen, H., and Ren, B. (2021). Notch signaling in vascular endothelial cells, angiogenesis, and tumor progression: An update and perspective. *Fronter*, **9**:148-177.

Albers, P. K., and McVean, G. (2018). Dating genomic variants and shared ancestry in population-scale sequencing data. *Mol. Biol.* **8:** 1–26.

Alemneh Tewodros and Mebrate Getabalew, (2019). Beef Cattle Production Systems, Challenges, and Opportunities in Ethiopia. *JOJPH.* **5**: 265-273.

Allais, S., Levéziel, H., Payet-Duprat, N., Hocquette, J.F., Lepetit, J., Rousset, S., Denoyelle, C., Bernard-Capel, C., Journaux, L. and Bonnot, A. (2010). The two mutations, Q204X and nt821, of the myostatin gene affect carcass and meat quality in young heterozygous bulls of French beef breeds. *J. Anim. Sci.* **88**: 446-454.

Angres, S., and Reed, C.A. (2014). Origin and descent of domestic mammals. *Nature*, **54**: 3-7.

Armando, C., Aurora, G., Dorado, D. (2013). Allelic diversity and its implications for the rate of adaptation. *Genetics*, **195**: 1373–1384.

Ayalew Addis, Chanie Demise, Lamesegn Desalegn, (2018). Review on productive and reproductive performance of indigenous dairy cattle breeds under farmer's management practices in Ethiopia. *J. Anim. Sci.* **8**: 169-174.

Bahbahani, H., Clifford, H., Wragg, D., Mbole-Kariuki, M. N., Van Tassell, C., Sonstegard, T., Hanotte, O. (2015). Signatures of positive selection in East African Shorthorn Zebu: A genome-wide single nucleotide polymorphism analysis. *Science*, **5**: 1-13.

Baena, M. M., Tizioto, P. C., Meirelles, S. L. C., Regitano, L. C. D. A. (2018). HSF1 and HSPA6 as functional candidate genes associated with heat tolerance in Angus cattle. *J. Anim. Sci*. **47**: e20160390.

Bahbahani, H., Salim, B., Almathen, F., Al Enezi, F., Mwacharo, J. M., and Hanotte, O. (2018). Signatures of positive selection in African Butana and Kenana dairy zebu cattle. *PloS*. **13**: e0190446.

Bejarano, D., Martínez, R., Manrique, C., Parra, L. M., Martínez Rocha, J. F., Gómez, Y., Gallego, J. (2018). Linkage disequilibrium levels and allele frequency distribution in blanco orejinegro and romosinuano creole cattle using medium density snp chip data. *Genet. Mol. Biol*. **41**:426–433.

Belhadj Slimen, I., Najar, T., Ghram, A., Abdrrabba, M. (2016). Heat stress effects on livestock: molecular, cellular and metabolic aspects, a review. *J. Anim. Nutr.* **100**: 401-412.

Berthier, D., Peylhard, M., Dayo, G. K., Flori, L., Sylla, S., Bolly, S., Thevenon, S. (2015). A comparison of phenotypic traits related to trypanotolerance in five West African cattle breeds highlights the value of shorthorn taurine breeds. *PLoS*. **10**: e0126498.

Bhuiyan, M. S., Lim, D., Park, M., Lee, S., Kim, Y., Gondro, C., Lee, S. (2018). Functional partitioning of genomic variance and genome-wide association study for carcass traits in Korean Hanwoo cattle using imputed sequence level SNP data. *Front. Genet.* **9**:217-222.

Birhanu Addis-fiker, (2019). Ethiopian Meat production trends consumption and meat quality parameters. *Int. J. Agric*. *Sci.* **3**:267-274.

Bohmanova, J., Sargolzaei, M., Schenkel, F.S. (2010). Characteristics of linkage disequilibrium in North American Holsteins. *Genomics,* **11**:421–432.

Bovine HapMap Consortium. (2009). Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*. **324**: 528-532.

Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., Cook, D. A. (2014). Standards for reporting qualitative research: a synthesis of recommendations. *Acad. Med*. **89**:1245-1251.

Caballero, A., Rodriguez-Ramilo, S. (2010). A new method for the partition of allelic diversity within and between subpopulations. *Genetics*, **11**: 2219–2229.

Carroll, E. L., Bruford, M. W., DeWoody, J. A., Leroy, G., Strand, A., Waits, L., and Wang, J. (2018). Genetic and genomic monitoring with minimally invasive sampling methods. *Evol. Appl*. **11**:1094–1119.

Chang K. (2007). Key signaling factors and pathways in the molecular determination of skeletal muscle phenotype. *Anim*. *Genet*. **1**:681–98.

Chen, J., Althagafi, A., and Hoehndorf, R. (2020). Predicting candidate genes from phenotypes, functions, and anatomical site of expression. *Mol. Biol*. **6**:1–7.

Central Statistical Agency, (2020/21). Report on livestock and livestock characteristics.

Czech, B., Frąszczak, M., Mielczarek, M., Szyda, J. (2018). Identification and annotation of breed-specific single nucleotide polymorphisms in Bos taurus genomes. *PloS*. **13**: e0198419.

Dadi Hailu, Kim, J. J., Yoon, D., Kim, K. S. (2012). Evaluation of single nucleotide polymorphisms (SNPs) genotyped by the Illumina Bovine SNP50K in cattle focusing on Hanwoo breed. *Asian-Australasian j. anim. sci.* **25**: p,28.

Dadi Hailu, Tibbo, M., Takahashi, Y., Nomura, K., Hanada, H., Amano, T. (2008). Microsatellite analysis reveals high genetic diversity but low genetic structure in Ethiopian indigenous cattle populations. *Anim.Genet.* **39**: 425-431.

Daetwyler, H. D., A. Capitan, H. Pausch, P. Stothard, R. van Binsbergen, R. F. Brondum, X. Liao, A. Djari, S. C. Rodriguez, C. Grohs, D. Esquerre, O. Bouchez, M. Rossignol, C. Klopp, D. Rocha, S. Fritz, A. Egging, P.J. Bowman, D. Coote, A.J. Chamberlain, C. Anderson, C.P. VanTassell, I. Hulsegge, M.E. Goddard, B. Guldbrandtsen, M.S. Lund, RF. Veerkam, D.A. Boichard, R. Fries and B.J. Hayes. 2014. Whole-genome sequencing of 234 bulls facilitates the mapping of monogenic and complex traits in cattle. *Nat. Genet.* **46**: 858-865.

Dong, W., Yang, Z., Yang, F., Wang, J., Zhuang, Y., Xu, C., Liu, D. (2012). Suppression of Rap1 impairs cardiac myofibrils and conduction system in zebrafish. *PloS*. **11**: e50960.

Daetwyler, H. D., Kemper, K. E., Van der Werf, J. H. J., Hayes, B. J. (2012). Components of the accuracy of genomic prediction in a multi-breed sheep population. *J.anim. sci.* **90**:3375-3384.

DAGRIS, (2007). Domestic Animal Genetic Resources Information System (edited by S. Kempo, Y. Mamo, B. Astrat and Tadele Dessie). *ILRI. http://dagris.ilri.cgiar.org.*

De Las Heras-Saldana, S., Lopez, B. I., Moghaddar, N., Park, W., Park, J. E., Chung, K. Y., van der Werf, J. H. (2020). Use of gene expression and whole-genome sequence information to improve the accuracy of genomic prediction for carcass traits in Hanwoo cattle. *Genet. Sel. Evol.* **52**:1-16.

Decker, J. E., McKay, S. D., Rolf, M. M., Kim, J. W., Molina Alcalá, A., Sonstegard, T. S., Taylor, J. F. (2014). Worldwide Patterns of Ancestry, Divergence, and Admixture in Domesticated Cattle. *Genetics,* **10**: e1004254.

Deng, Y., Liu, T., Xie, Y., Wei, Y., Xie, Z., Shi, Y., Deng, X. (2020). High Genetic Diversity and Low Differentiation in Michelia shiluensis, an Endangered Magnolia Species in South China. *Fronter,* **11**: P, 469.

Derks, M. F., Megens, H. J., Bosse, M., Visscher, J., Peeters, K., Bink, M. C., Groenen, M. A. (2018). A survey of functional genomic variation in domesticated chickens. *Genet. Sel. Evol.* **50**:1-14.

Edea Zewdu, Bhuiyan, M.S.A., Dessie Tadelle, Rothschild, M.F., Dadi Hailu. and Kim, K.S. (2015). Genome-wide genetic diversity, population structure, and admixture analysis in African and Asian cattle breeds. *Anim. Sci.* **9**:218-226.

Edea Zewdu, Dadi Hailu, Dessie Tadelle, Kim, K. S. (2018). Revealing of Selection Signatures Underlying Adaptive and Production Traits in Hanwoo and Ethiopian Cattle Breeds. *J. Anim. Biol. Genet.* **3**:1-13.

Edea Zewdu, Dadi Hailu, Kim, S. W., Park, J. H., Shin, G. H., Dessie, T., Kim, K. S. (2014). Linkage disequilibrium and genomic scan to detect selective loci in cattle populations adapted to different ecological conditions in Ethiopia. *J. Anim. Genet*. **131**:358-366.

Edea Zewdu, Dadi Hailu, Kim, S., Dessie Tadele, Lee, T., Kim, H., Kim, K. (2013). Genetic diversity, population structure, and relationships in indigenous cattle populations of Ethiopia and Korean Hanwoo breeds using SNP markers. *Front.Genet.* **4**: 1–9.

Edea Zewdu, Dadi Hailu, Kim, S.W., Dessie Tadele, and Kim, K.S. (2012). Comparison of SNP variation and distribution in indigenous Ethiopian and Korean Cattle (Hanwoo) populations. *Genomics,* **10**: 200-205.

Edea Zewdu, Jeoung, Y. H., Shin, S. S., Ku, J., Seo, S., Kim, I. H., Kim, K. S. (2018). Genome-wide association study of carcass weight in commercial Hanwoo cattle. J. *Anim. Sci*. **31**:321-327.

Edea Zewdu, Jung, K. S., Shin, S., and Yoo, S. (2020). Signatures of positive selection underlying beef production traits in Korean cattle breeds. *J. Anim. Sci.* **62**:293–305.

Elsik, C.G., Tellam, R.L., Worley, K.C. (2009). The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science,* **324**:522-528.

Espigolan, R., Baldi, F., Boligon, A., Souza, P., Gordo, G.M., Tonussi, L., Cardoso, F., Oliveira, N., Tonhati, H., Sargolzaei, M. (2013). Study of whole genome linkage disequilibrium in Nellore cattle. *Genomics*, **14**:305.

Ethiopian Biodiversity Institute (2016). Ethiopian National Strategy and Plan of Action for Conservation.

Eynard, S. E., Windig, J. J., Leroy, G., van Binsbergen, R., Calus, M. P. (2015). The effect of rare alleles on estimated genomic relationships from whole genome sequence data. *Genetics*, **16**: 1-12.

Fariello, M. I., Servin, B., Tosser-Klopp, G., Rupp, R., Moreno, C., San Cristobal, M., International Sheep Genomics Consortium. (2014). Selection signatures in worldwide sheep populations. *PloS* **9**: p, e103813.

Fernandes Júnior, G. A., de Oliveira, H. N., Carvalheiro, R., Cardoso, D. F., Fonseca, L. F. S., Ventura, R. V., and de Albuquerque, L. G. (2020). Whole-genome sequencing provides new insights into genetic mechanisms of tropical adaptation in Nellore (Bos primigenius indicus). *Science*, **10**: 1–7.

Fijarczyk, A., and Babik, W. (2015). Detecting balancing selection in genomes: limits and prospects. *Mol. Ecol.* **24**: 3529-3545.

Flannery, M. A. (2018). Nature's Prophet: Alfred Russel Wallace and His Evolution from Natural Selection to Natural Theology. University of Alabama Press.

Fukuoka, H., Tachibana, K., Shinoda, Y., Minamisaka, T., Inui, H., Ueno, K., Hoshida, S. (2020). Tolvaptan-induced hypernatremia is related to low serum potassium level accompanying high blood pressure in patients with acute decompensated heart failure. *Rese. Squ.* **20**: 1-7.

Gao Y, Zhang Y, Jiang H, Xiao S, Wang S, Ma Q. (2011). Detection of differentially expressed genes in the longissimus dorsi of Northeastern indigenous and large white pigs. *Genet. Mol.* **10:** 779–91.

Gao, Y., Zhang, R., Hu, X., and Li, N. (2007). Application of genomic technologies to the improvement of meat quality of farm animals. *Science,* **77**: 36-45.

Getachew, T., Haile, A., Tibbo, M., Sharma, A. K., Solkner, J., and Wulzinger, M., 2010. Herd management and breeding practices sheep owners in a mixed crop-livestock and a pastoral system of Ethiopia. *J. Agri. Rese.* **5**: 685-691.

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

Gomes, M. S., Chagas, P., Padilha, D. M. P., Caramori, P., Hugo, F. N., Schwanke, C. H. A., Hilgert, J. B. (2012). Association between self-reported oral health, tooth loss, and atherosclerotic burden. *Braz. Oral Res.* **26**:436-442.

Greenbaum, G, Templeton A.R., Zarmi Y, Bar-David S. (2014). Allelic Richness following Population Founding Events – A Stochastic Modeling Framework Incorporating Gene Flow and Genetic drift. *PLoS*. **9**: e115203. https://doi.org/10.1371/journalpone.0115203

Greenwood, P., Cafe, L., McIntyre, B., Geesink, G., Thompson, J., Polkinghorne, R., Pethick, D., and Robinson, D. (2013). Molecular value predictions: Associations with beef quality, carcass, production, behavior, and efficiency phenotypes in Brahman cattle. *J. Anim. Sci.* **91**:5912-5925.

Groeneveld, L. F., Lenstra, J. A., Eding, H., Toro, M. A., Scherf, B., Pilling, D., Weigend, S. (2010). Genetic diversity in farm animals - A review. *Anim. Genet.* **41**: 6–31.

Gutiérrez-Gil, B., Wiener, P., Williams, J. L. (2007). Genetic effects on coat colour in cattle: dilution of eumelanin and phaeomelanin pigments in an F2-Backcross Charolais× Holstein population. *Genetics*, **8**: 1-12.

Guillemin N, Bonnet M, Jurie C, Picard B. (2011). Functional analysis of beef tenderness. *Elsevier,* **75**:352–65

Gurgul, A., Jasielczuk, I., Ropka-Molik, K., Semik-Gurgul, E., Pawlina-Tyszko, K., Szmatoła, T., Krupiński, J. (2018). A genome-wide detection of selection signatures in conserved and commercial pig breeds maintained in Poland. G*enetics*, **19**: 1-17.

Hagos Berhane, (2017). Ethiopian Cattle Genetic Resource and Unique Characteristics under a Rapidly Changing Production Environment-A Review. *Int.J.Sol.Res.* **6**: 1959–1968.

Haile, A., Joshi, B. K., Ayalew, W., Tegegne, A., Singh, A. (2011). Genetic evaluation of Ethiopian Boran cattle and their crosses with Holstein Friesian for growth performance in central Ethiopia. *J. Anim. Genet.* **128**: 133-140.

Hamill, R.M., Mc.Bryan, J, McGee C, Mullen AM, Sweeney T, Talbot A. (2012). Functional analysis of muscle gene expression profiles associated with tenderness and intramuscular fat content in pork. *Science,* **92**:440–50.

Hanotte, O., Bradley, D. G., Ochieng, J. W., Verjee, Y., Hill, E. W., and Rege, J. E. O. (2003). African pastoralism: Genetic imprints of origins and migrations. *Science,* **296**: 336–339.

He,J., Huang, H., Du, Y., Peng, D., Zhou, Y., Li, Y., and Wang, H. (2020). Association of DCBLD2 upregulation with tumor progression and poor survival in colorectal cancer. *Cell. Oncol.***21**: 2822–2828.

Hu., Scheben, A., Edwards, D. (2018). Advances in integrating genomics and bioinformatics in the breeding program. *Agriculture,* **8**:56-69.

Hui, T. Y. J., and Burt, A. (2020). Estimating linkage disequilibrium from genotypes under Hardy-Weinberg equilibrium. *Genetics,* **21**: 1–11.

Ijarczyk A. and Babik W. (2015). Detecting balancing selection in genomes: limits and prospects. *Mol. Ecol.* **24**: 3529-3545.

Jacobs, G. S., Sluckin, T. J., and Kivisild, T. (2016). Refining the use of linkage disequilibrium as a robust signature of selective sweeps. *Genetics,* **203**: 1807–1825.

Jian, X., Boerwinkle, E., Liu, X. (2014). In silico prediction of splice-altering single-nucleotide variants in the human genome. *Nucleic Acids Res.* **42**:13534-13544.

Jiang C, Shi P, Li S, Dong R, Tian J, Wei J. (2010). Gene expression profiling of skeletal muscle of nursing piglets. *Int. J. Biol. Sci.* **6**: 627-633.

Joo, S., Kim, G., Hwang, Y., Ryu, Y. (2013). Control of fresh meat quality through manipulation of muscle fiber characteristics. *Meat Sci*. **95**: 828–36.

Kelleher, M. M., Berry, D. P., Kearney, J. F., McParland, S., Buckley, F., Purfield, D. C. (2017). Inference of population structure of purebred dairy and beef cattle using high-density genotype data. *Anim. Genet*. **11:** 15–23.

Keogh, K., Kenny, D. A., Kelly, A. K., Waters, S. M. (2015). Insulin secretion and signaling in response to dietary restriction and subsequent re-alimentation in cattle. *Physiol. Genomics,* **47**: 344-354.

Khan, A., Dou, J., Wang, Y., Jiang, X., Khan, M. Z., Luo, H., Zhu, H. (2020). Evaluation of heat stress effects on cellular and transcriptional adaptation of bovine granulosa cells. *J. anim. sci. biotechol.* **11**: 1-20.

Kimura, K. A. Y. O. K. O., Reinhardt, T. A., Goff, J. P. (2006). Parturition and hypocalcemia blunts calcium signals in immune cells of dairy cattle. *J. Sci.* **89**:2588-2595.

Kim, J., Hanotte, O., Mwai, O. A., Dessie, T., Bashir, S., Diallo, B., Kim, H. (2017). The genome landscape of indigenous African cattle. *Geno. biol*. **18**: 1-14.

Kim, Y. B., Komor, A. C., Levy, J. M., Packer, M. S., Zhao, K. T., Liu, D. R. (2017). Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol*. **35**: 371-376.

Lee, H. J., Kim, J., Lee, T., Son, J. K., Yoon, H. B., Baek, K. S.,Kim, H. (2014). Deciphering the genetic blueprint behind Holstein milk proteins and production. *Geno. Biol. Evol*. **6**: 1366-1374.

Lee, Y., Oh, D., Lee, J., La, B., and Yeo, J. (2013). Novel single nucleotide polymorphisms of the bovine SREBP1 gene are associated with the fatty acid composition and marbling score in commercial Korean cattle (Hanwoo). *Mol. Biol.* **40**:247-254.

Lena, I., Mantegazza, M. (2019). Haploinsufficiency in Scn2a knock-out mice causes an autistic-like phenotype attenuated with age. *Science.* **9**: 1-14.

Lenis Sanin, Y., Zuluaga Cabrera, A. M., Tarazona Morales, A. M. (2016). Adaptive responses to thermal stress in mammals. *Rev. Med. Vet.* **31**: 121-135.

Lenstra, J. A., Ajmone-marsan, P., Beja-pereira, A., Bollongino, R., Bradley, D. G., Colli, L., Science, F. (2014). Meta-Analysis of Mitochondrial DNA Reveals Several Population Bottlenecks during Worldwide Migrations of Cattle. *Diversity*, **6**:178–187.

Liu, K and Muse, S.V., 2005. Power Marker integrated analysis environment for genetic marker data. *Bioinformatics*, **21**: 2128–2129.

Li, T., Song, Y., Bao, X., and Zhang, J. (2020). Mediation of miR-34a/miR-449c for Immune Cytokines in Acute Cold/Heat-Stressed Broiler Chicken. *Anim. Sci.* **10**:2160- 2168.

Li, W.F., Li, J.Y., Gao, X., Xu, S.Z., and Yue, W.B. (2013). Association analysis of PRKAG3 gene variants with carcass and meat quality traits in beef cattle. *Afr. J. Biotechnol.* **11**:1855-1861.

Liu, B. J., Li, Y. L., Zhang, B. D., and Liu, J. X. (2020). Genome-Wide Discovery of Single-Nucleotide Polymorphisms and Their Application in Population Genetic Studies in the Endangered Japanese Eel (Anguilla japonica). *Science,* **6**:1–11.

Lashmar, S. F., Muchadeyi, F. C., Visser, C. (2019). Genotype imputation as a cost-saving genomic strategy for South African Sanga cattle: A review. *S. Afr. J. Anim. Sci*. **49**: 262-280.

Liu, W. B., Liu, J., Liang, C. N., Guo, X., Bao, P. J., Chu, M., Yan, P. (2014). Associations of single nucleotide polymorphisms in candidate genes with the polled trait in D atong domestic yaks. *Anim. Genet.* **45**:138-141.

Lopez, T., Haucke, V. and Maritzen, T. (2020). Endocytosis in the adaptation to cellular stress. *Cell Stress,* **4:**217- 230.

Lu, L., Xing, D., Ren, N. (2012). Pyrosequencing reveals highly diverse microbial communities in microbial electrolysis cells involved in enhanced H2 production from waste activated sludge. *Water res*. 46: 2425-2434.

MacHugh, D. E., Larson, G., Orlando, L. (2017). Taming the past: ancient DNA and the study of animal domestication. *Anim. Biosci*. **5**: 329-351.

Magnabosco, C., Ryan, K., Lau, M. C., Kuloyo, O., Lollar, B. S., Kieft, T. L., Onstott, T. C. (2016). A metagenomic window into carbon metabolism at 3 km depth in Precambrian continental crust. *Int. J. Biochem. Cell Biol.* **10**:730-741.

Makina, S. O., Muchadeyi, F. C., van Marle-Köster, E., Taylor, J. F., Makgahlela, M. L., Maiwashe, A. (2015). Genome-wide scan for selection signatures in six cattle breeds in South Africa. *Genet. Sel. Evol.* **47**: 1-14.

Mariotti, M., Williams, J. L., Dunner, S., Valentini, A., Pariset, L. (2009). Polymorphisms within the toll-like receptor (TLR)-2, -4, and-6 genes in cattle. *Diversity*, **1**: 7-18.

Makina, S. O., Whitacre, L. K., Decker, J. E., Taylor, J. F., MacNeil, M. D., Scholtz, M. M., Maiwashe, A. (2016). Insight into the genetic composition of South African Sanga cattle using SNP data from cattle breeds worldwide. *Genet.Sel. Evol.* **48**:1–7.

Mariasegaram, M., Harrison, B. E., Bolton, J. A., Tier, B., Henshall, J. M., Barendse, W., and Prayaga, K. C. (2012). Fine-mapping the POLL locus in Brahman cattle yields the diagnostic marker CSAFG29. *Anim. Genet*. **43**: 683-688.

Matukumalli, L. K., Lawley, C. T., Schnabel, R. D., Taylor, J. F., Allan, M. F., Heaton, M. P., Van Tassell, C. P. (2009). Development and characterization of a high-density SNP genotyping assay for cattle. *PloS*. **4**: e5350.

McKay, S. D., Schnabel, R. D., Murdoch, B. M., Matukumalli, L. K., Aerts, J., Coppieters, W., Moore, S. S. (2007). Whole genome linkage disequilibrium maps in cattle. *Genetics,* **8**: 1-12.

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E. (2016). Prediction of total genetic value using genome-wide dense marker maps. *Genetics,* **157**: 1819–1829.

Moreira, H. L., Savegnago, R. P., de Freitas, L. A., Lobo, R. B., Bezerra, L. A. F., de Paz, C. C. P. (2019). Breeding goals and economic values for Nellore cattle in a full-cycle production system. *Anim. Sci*. **41**: 1–8.

Mullen, M. P., Berry, D. P., Howard, D. J., Diskin, M. G., Lynch, C. O., Giblin, L., Waters, S. M. (2011). Single nucleotide polymorphisms in the insulin-like growth factor 1 (IGF-1) gene are associated with performance in Holstein-Friesian dairy cattle. *Front. genet.* **2**: 3-17.

Mustafa, H., Ahmad, N., Heather, H. J., Eui-Soo, K., Khan, W. A., Ajmal, A., Sonstegard, T. S. (2018). A whole genome study of linkage disequilibrium in Sahiwal cattle. *S. Afr. J. Anim. Sci.* **48**:353-360.

Mustafa, H., Eiusoo, K., Heather, H., Ajmal, A., Riley, D., Pasha, N., Sonstegard, T. (2017). Genome-wide SNPs analysis of indigenous zebu breeds in Pakistan. *Biotechnol. Anim.* **33**:13–25.

Mwai, O., Hanotte, O., Kwon, Y., and Cho, S. (2015). - Invited Review - African Indigenous Cattle: Unique Genetic Resources in a Rapidly Changing World. *J. Anim. Sci.* **28***:* 911–921.

Neves, H. H., Carvalheiro, R., O'brien, A. M. P., Utsunomiya, Y. T., Do Carmo, A. S., Schenkel, F. S., Garcia, J. F. (2014). Accuracy of genomic predictions in Bos indicus (Nellore) cattle. *Genet. Sel. Evol.* **46**: 1-13.

Nietlisbach, P., Keller, L. F., and Postma, E. (2016). Genetic variance components and heritability of multiallelic heterozygosity under inbreeding. *Heredity*, **116**:1–11.

Njisane, Y. Z., Mukumbo, F. E., Muchenje, V. (2020). An outlook on livestock welfare conditions in African Communities A review. *Asia. Aust. J. Anim. Sci.* **33**:867–878.

Okeyo-Owuor, T., White, B. S., Chatrikhi, R., Mohan, D. R., Kim, S., Griffith, M., Graubert, T. A. (2015). U2AF1 mutations alter sequence specificity of pre-mRNA binding and splicing. *Leukumia,* **29**:909-917.

Pant, S. D., Schenkel, F. S., Verschoor, C. P., and Karrow, N. A. (2012). Use of breed-specific single nucleotide polymorphisms to discriminate between Holstein and Jersey dairy cattle breeds. *Anim. Biotechnol.* **23**:1–10.

Pérez-Enciso, M., Rincón, J. C., Legarra, A. (2015). Sequence-vs. chip-assisted genomic selection: accurate biological information is advised. *Genet. Sel. Evol*. **47**: 1-14.

Pitt, D., Sevane, N., Nicolazzi, E. L., MacHugh, D. E., Park, S. D. E., Colli, L., Orozco-ter Wengel, P. (2019). *Evol. Appl.* **12**: 123–136.

Porto-Neto, L.R., Kijas, J.W., Reverter, A. (2014). The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. *Genet. Sel. Evol*. **46**:22-37.

Purcell Neale, B., Todd-Brown, K. (2007). PLINK a toolset for whole-genome association and population-based linkage analysis. *Anim. J. Genet*. **81**:559-562.

Purfield, D. C., McClure, M., and Berry, D. P. (2016). Justification for setting the individual animal genotype call rate threshold at eighty-five percent. *J. Anim.Sci.* **94**:4558–4569.

Qanbari, S., Simianer, H. (2014). Mapping signatures of positive selection in the genome of livestock. *Livest. Sci.* **166**:133-143.

Qanbari, S., Pausch, H., Jansen, S., Somel, M., Strom, T. M., Fries, R., Simianer, H. (2014). Classic selective sweeps revealed by massive sequencing in cattle. *Genetics,* **10**: e1004148.

Qanbari, S., Pimentel, E. C. G., Tetens, J., Thaller, G., Lichtner, P., Sharifi, A. R., and Simianer, H. (2010). The pattern of linkage disequilibrium in German Holstein cattle. *Anim. Genet.* **41**: 346-356.

Ramos, A. M., Megens, H. J., Crooijmans, R. P. M. A., Schook, L. B., and Groenen, M. A. M. (2011). Identification of high utility SNPs for population assignment and traceability purposes in the pig using high-throughput sequencing. *Anim. Genet.* **42**:613–620.

Randhawa I.A., Khatkar M.S., Thomson P.C. and Raadsma H.W. (2016). A Meta-Assembly of Selection Signatures in Cattle. *PloS.* **11**: e0153013.

Rege, J. (1999). The state of African cattle genetic resources I. Classification framework and identification of threatened and extinct breeds. *Anim. Genet.* **25**, 1-25.

Ribeca, C., Bonfatti, V., Cecchinato, A., Albera, A., Gallo, L., Carnier, P. (2014). Effect of polymorphisms in candidate genes on carcass and meat quality traits in double-muscled Piemontese cattle. *Science,* **96**: 1376-83.

Rincón, A. D., Prevosti, F. J., and Parra, G. E. (2011). New saber-toothed cat records (Felidae: Machairodontinae) for the Pleistocene of Venezuela, and the Great American Biotic Interchange. *J. Vertebr. Paleontol.* **31**: 468–478.

Rouleau, M. Patel, A., Hendzel, M.J. (2010). PARP inhibition: PARP1 and beyond. *Nature,* **10**:293–301.

Rodriguez-Peña, R. A., Johnson, R. L., Johnson, L. A., Anderson, C. D., Ricks, N. J., Farley, K. M., Stevens, M. R. (2018). Investigating the genetic diversity and differentiation patterns in the Penstemon scariosus species complex under different sample sizes using AFLPs and SSRs. *Genetics*, **19**:1335–1348.

Rothammer, S., Seichter, D., Förster, M., Medugorac, I. (2013). A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. *Genomics,* **14**: 1-17.

Roux P-F, Boitard S, Blum Y, Parks B, Montagner A, Mouisel E. (2015). Combined QTL and selective swe1ep mappings with coding SNP annotation and cis-eQTL analysis revealed PARK2 and JAG2 as new candidate genes for adiposity regulation. *Genetics,* **5**:517–29.

Sadkowski, T., Jank, M., Zwierzchowski, L., Siadkowska, E., Oprzdek, J., Motyl, T. (2008). Gene expression profiling in skeletal muscle of Holstein-Friesian bulls with single nucleotide polymorphism in the myostatin gene 5’-flanking regions. *J.Genet.***49**:237–50.

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

Seabury, C. M., Oldeschulte, D. L., Saatchi, M., Beever, J. E., Decker, J. E., Halley, Y. A., Taylor, J. F. (2017). Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *Genomics,* **18**:376-386.

Sendeku Asemu Tesfa, Kumar, D., Abegaz Solomon, Mekuriaw Getenet, (2016). Evaluations of Reproductive Performances of Fogera Cattle Breed in Selected Districts of Amhara Region, Ethiopia. *Int. J. Pharma. Med. Biol. Sci.* **5**:52-57.

Shor_shimoni, E., Shabtay, A., Agmon, R., Cohenzinder, M. (2017). Detection of Allelic and Genotypic Frequencies of Polymorphisms Associated with Meat Quality in the Mediterranean Baladi Cattle. *Open Agric. J.* **11**:1-10.

Sikora, K. M., Magee, D. A., Berkowicz, E. W., Berry, D. P., Howard, D. J., Mullen, M. P., Spillane, C. (2011). DNA sequence polymorphisms within the bovine guanine nucleotide-binding protein Gs subunit alpha (Gsα)-encoding (GNAS) genomic imprinting domain are associated with performance traits. *Genetics,* **12**: 1-16.

Stafuzza, N. B., Zerlotini, A., Lobo, F. P., Yamagishi, M. E. B., Chud, T. C. S., Caetano, A. R., Da Silva, M. V. G. B. (2017). Single nucleotide variants and InDels were identified from whole-genome re-sequencing of Guzerat, Gyr, Girolando, and Holstein cattle breeds. *PloS.* **12**:1–15.

Stranger, B. E., Stahl, E. A., and Raj, T. (2011). Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics, **187:*** 367–383.

Tarekegn Getenet Mekuriaw, Ji, X., Bai, X., Liu, B., Zhang, W., Birungi, J., Tesfaye Kassahun, (2018). Variations in mitochondrial cytochrome b region among Ethiopian indigenous cattle populations assert *B. taurus* maternal origin and historical dynamics. *Ecol. Evol.* **31**: 1393–1400.

Talenti, A., Bertolini, F., Williams, J., Moaeen-ud-Din, M., Frattini, S., Coizet, B., Italian Goat Consortium. (2018). Genomic analysis suggests KITLG is responsible for a roan pattern in two Pakistani goat breeds. *J. Hered.* **109**:315-319.

Teneva, A., Dimitrov, K., Caro Petrović, V., Petrović, M. P., Dimitrova, I., Tyufekchiev, N., Petrov, N. (2013). Molecular genetics and SSR markers as a new practice in farm animal genomic analysis for breeding and control of disease disorders. *Biotechnol. Anim. Husb.* **29**: 405-429.

Theunissen, B., and Lenstra, J. A. (2015). Conservation of cattle genetic resources: The role of breeds Conservation of cattle genetic resources: the role of breeds. *Cambridge University Press,* **153**: 152–162.

Torrino, S., Visvikis, O., Doye, A., Boyer, L., Stefani, C., Munro, P., Lemichez, E. (2011). The E3 ubiquitin-ligase HACE1 catalyzes the ubiquitylation of active Rac1. *Cell*, **21**:959-965.

Utsunomiya Y.T., Perez O'Brien A.M., Sonstegard T.S., Solkner J. and Garcia J.F. (2015). Genomic data as the "hitchhiker's guide" to cattle adaptation: tracking the milestones of past selection in the bovine genome. *Front. Genet.* **6**: 36-42.

Uzzaman, M. R., Zewdu Edea, M., Bhuiyan, S. A., Walker, J., Bhuiyan, A. K. F. H., and Kim, K. S. (2014). Genome-wide single nucleotide polymorphism analyses reveal genetic diversity and structure of wild and domestic cattle in Bangladesh. *Asian Australas J. Anim. Sci.* **27**:1381.

Van Den Berg, I., Xiang, R., Jenko, J., Pausch, H., Boussaha, M., Schrooten, C., Goddard, M. E. (2020). Meta-analysis for milk fat and protein percentage using imputed sequence variant genotypes in 94,321 cattle from eight cattle breeds. *Genet. Sel. Evol*. **52**:1-16.

van Hooren, K. W., van Agtmaal, E. L., Fernandez-Borja, M., van Mourik, J. A., Voorberg, J., and Bierings, R. (2012). The Epac-Rap1 signaling pathway controls cAMP-mediated exocytosis of Weibel-Palade bodies in endothelial cells. *J. Biol.Chem.* **287**:24713-24720.

Voorrips, R. E. (2002). Map Chart: software for the graphical presentation of linkage maps and QTLs. *J. Hered*. **93**: 77-78.

Vitti, J.J., Grossman, S.R. Sabeti, P.C. (2013). Detecting natural selection in genomic data. *Anim. Genet.* **47**: 97-120.

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

Wientjes, Y. C. J., Veerkamp, R. F., Calus, M. P. L. (2013). Effect of linkage disequilibrium, haplotypes, and family relations on the accuracy of genomic prediction. In Book of Abstracts of the 63rd Annual Meeting of the European Association for Animal Production. *Science,* **43**: 306-306.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, **3**:395-420.

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat.Genet.***42**:565–569.

Yaro, M., Munyard, K. A., Stear, M. J., Groth, D. M. (2016). Combatting African animal trypanosomiasis (AAT) in livestock: the potential role of trypanotolerance. *J. Biosci.* **225**: 43-52.

Yasuhara, T., Suzuki, T., Katsura, M., Miyagawa, K. (2014). Rad54B serves as a scaffold in the DNA damage response that limits checkpoint strength. *Nature,* **5**: 1-11.

Yin, M., Wang, X., and Lu, J. (2020). Advances in IKBKE as a potential target for cancer therapy. *Cancer Med*. **9**:247-258.

Yu, Q. P., Feng, D. Y., Xiao, J., Wu, F., He, X. J., Xia, M. H., Liu, Y. H. (2017). Studies on meat color myoglobin content enzyme activities and genes associated with the oxidative potential of pigs slaughtered at different growth stages. *J. Anim. Sci.* **30**:1739–1750.

Yuan, Z., Li, J., Li, J., Gao, X., Gao, H., and Xu, S. (2013). Effects of DGAT1 gene on meat and carcass fatness quality in Chinese commercial cattle. *Mol. Biol. Rep.* **40**: 1947-1954.

Zeng, L., Chen, N., Ning, Q., Yao, Y., Chen, H., Dang, R., Zhang, H. (2018). CPRLH and SOD1 gene variations associated with heat tolerance in Chinese cattle. *Anim. Genet.* **49**:447-451.

Zhao, F., McParland, S., Kearney, F., Du, L., and Berry, D. P. (2015). Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genet. Sel. Evo.* **47**:1–12.

Zhu, B., Zhang, J. Jing, Niu, H., Guan, L., Guo, P., Xu, L. Yang, Li, J. Ya. (2017). Effects Of Marker Density And Minor Allele Frequency On Genomic Prediction For Growth Traits In Chinese Simmental Beef Cattle. *Int. J. Agric. Biol.* **16**:911–920.

Zwane, A. A., Maiwashe, A., Makgahlela, M. L., Choudhury, A., Taylor, J. F., and van Marle-Koster, E. (2016). Genome-wide identification of breed-informative single-nucleotide polymorphisms in three South African indigenous cattle breeds. *S. Afr. J. Anim. Sci.* **46**: 302–312.

Zwane, A. A., Schnabel, R. D., Hoff, J., Choudhury, A., Makgahlela, M. L., Maiwashe, A., Taylor, J. F. (2019). Genome-wide SNP discovery in indigenous cattle breeds of South Africa. *Front. Genet.* **10:** 1–16.