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

Dejenie Mengistie (dejeniebiot2006@gmail.com)
National Agricultural Biotechnology Research center
Zewdu Edea
Ethiopian Biotechnology Institute
Tesfaye sisay Tesema
Addis Ababa University Institute of Biotechnology
Genet Dejene
Ethiopian Biotechnology Institute
Tadelle Dessie
International Livestock Research Institute
Jeilu Jemal
National Agricultural Biotechnology Research center
Kwan Suk Kim
Chungbuk National University
Hailu Dadi
Ethiopian Biotechnology Institute

Research Article

Keywords: Linkage disequilibrium, SNP array, Selection signature

Posted Date: April 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1554212/v1

Abstract

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 and linkage disequilibrium Ethiopian cattle populations were genotyped with an 80K SNP array. European beef cattle breeds were also used for comparison purposes. Across Ethiopian cattle populations, the mean observed and expected heterozygosity were 0.403 and 0.400, respectively. Similarly, for European cattle, observed and expected heterozygosity were 0.25 and 0.26 respectively. PCA and NJ-tree revealed a separation of Ethiopian cattle breeds from European beef breeds. Similarly, NJ-tree grouped the study cattle according to their breed group with close clustering of Ethiopian cattle populations. The average r^2 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. 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. Furthermore, some of the candidate genes are associated with tropical adaptation to heat tolerance and resistance to disease. Significant SNP variation found in this study suggests that these markers could be useful for genetic research in Ethiopian cattle breeds.

1. Introduction

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 indigenous 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 genomic 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 breeds 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, and linkage disequilibrium (LD) in three Ethiopian indigenous cattle populations (Boran, Fogera, Begiat) and for comparison purpose European beef cattle breeds (Angus, Herford, Charolais) were considered.

2. Materials And Methods

2.1. Study breeds, sample collection, and DNA extraction

DNA samples were collected from three Ethiopian indigenous cattle populations that were kept in government ranches: Begait (n = 40) from Humera ranch, Boran (n = 40) from Dida Tiyura ranch, and Fogera (n = 43) from Andassa ranch. 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 from online database.

2.2. 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 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.

2.3. Data analysis

To examine the basic indices of within-breed genetic variability, genetic diversity (observed and expected heterozygosity) was estimated using PLINK (Purcell *et al.*, 2007). Principal component analysis (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).

Linkage disequilibrium was calculated between pairs of SNPs within a chromosome using SNP & VARIATION SUITE (Golden Helix, Inc; www.goldenhelix.com). The r^2 measure, 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:

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

3. Results

3.1. 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 1. The overall mean of observed and expected heterozygosity for Ethiopian breeds were 0.40, and 0.40 respectively; while 0.25 and 0.26 values were for European beef breeds (Table 1).

Table 1: 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

3.2. Population structure and relationship

To assess the population structure among the six cattle breeds/populations, a PCA 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 (Fig. 1). As portrayed in Fig. 2, PC1 explained 70.63% of the total variation and unequivocally separated the Ethiopian indigenous breeds from European

beef breeds (Fig. 2). Similarly, the phylogenetic analysis reiterated the PCA results. The three Ethiopian cattle breeds shared the same clade, while the European cattle breeds were found in separate clades (Fig. 3). As depicted in Fig. 4 the two European cattle breeds (Hereford and Charolais) were closely clustered as did Boran and Fogera.

Phylogenetic tree analysis (Fig. 4) strongly supports the PCA result, and clustering analysis a clear divergence among breeds of the Ethiopian breeds and the European beef cattle breeds.

3.3. 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 (Fig. 6). 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 (Fig. 5).

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 (Fig. 6). 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.

For each breed/population, the mean r^2 and D' values are presented in Table 2. The mean r^2 values ranged from 0.17 in Angus and Hereford to 0.23 in Begait. Similarly, the mean D' values varied from 067 in Boran to 0.42 in Angus.

Table 2
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

3.4. Genetic differentiation and selection signatures

The highest genetic differentiation (*FST = 0.315*) was observed 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 3
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

Pairwise comparison genetic differentiation between Ethiopian indigenous population and European beef cattle breeds showed very high differentiation (FST value), especially Angus breed showed greater differentiation in all Ethiopian breeds as indicated in Fig. 7.

3.4.1. Functional analysis of commonly differentiated genes

Commonly differentiated SNPs were annotated and the corresponding genes were associated with biological processes and pathways relevant to meat traits and tropical adaptations.

3.4.2. *Genes related to meat quality trait in European beef breeds*

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 4) 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 4
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

3.4.3. Commonly differentiated candidate genes related to tropical adaptation in Ethiopian cattle populations

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 5
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 6), showed higher variability in Ethiopian indigenous breeds and lower variability in European beef breeds.

Table 6
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

3.4.4. Identification of candidate genes QTL

Table 7
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

4. Discussions

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, and 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 the 50K Bead Chip analysis (Edea *et al.* 2012). However, it was also lower than the values recorded from microsatellite markers (Dadi *et al.* 2008).

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

4.2. Genetic distance and signatures of selection

Pairwise comparison of genetic distance among the six populations ranged from 0.019 to 0.315. Great differentiation (0.315) was observed between Angus (*Taurus*) and Begait (*Zebu*) 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 (Zenga) and Boran(zebu).

These low values of genetic distance indicate 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).

4.3. 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 in better understanding 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.

4.4. 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, genes involved in actin cytoskeleton organization were determined as meat tenderness (Guillemin *et al*., 2011, Gao *et al*., 2011), and similar genes 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).

4.5. 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 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.

4.6. 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 candidate genes association. 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 the color meat.

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

Based on mutual functional annotation, the DAVID 6.8 functional 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 (Chang K. 2007).

4.8. 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).

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

Pathway analysis identified 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).

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). 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).

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).

4.9. 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). 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). The LD value in Ethiopian indigenous cattle populations is higher than in European breeds, this is due to the chips designed from *B. indicus* and it is biased to Ethiopian breeds, since Ethiopian cattle populations (i.e., Boran, Begait, Fogera) are Zebu, while European breeds (i.e., Angus, Herford, Charolais) are *B. taurus*. 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, the 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 genetic distance of the indigenous Ethiopian breeds from European beef breeds coincides with existing knowledge, that European beef breeds considered under this study are artificially selected and are taurine, while Ethiopian indigenous breeds are naturally selected. The higher LD in Ethiopian indigenous cattle populations shows that the chips are biased to Ethiopian breeds. The identification of selection signatures helps us to understand domestication, breed formation, population structure, adaptive evolution, and selection consequences better. 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. To prevent the false-positive result the identified candidate genes have to be confirmed through GWAS, sequencing, and omics studies to include them in cattle genetic improvement programs. Future developments of structured breeding programs and the use of genomic selection in the development sector will allow the livestock industry to breed more productive and resilient cattle populations, which will ultimately assist with alleviating poverty and reducing hunger.

Declarations

Acknowledgments

It gives great honor to thank CHUNGBUK NATIONAL UNIVERSITY for providing us SNP dada.

References

1. 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.
2. 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.
3. 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.
4. 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.
5. Bohmanova, J., Sargolzaei, M., Schenkel, F.S. (2010). Characteristics of linkage disequilibrium in North American Holsteins. *Genomics,* **11**:421–432.
6. Chang K. (2007). Key signaling factors and pathways in the molecular determination of skeletal muscle phenotype. *Anim. Genet.* **1**:681–98.
7. 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.
8. 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.

9. 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.
10. 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.
11. 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.
12. Ethiopian Biodiversity Institute (2016). Ethiopian National Strategy and Plan of Action for Conservation.
13. 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.
14. 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.
15. 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.
16. Hu., Scheben, A., Edwards, D. (2018). Advances in integrating genomics and bioinformatics in the breeding program. *Agriculture,***8**:56-69.
17. 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.
18. 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.
19. 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.
20. Liu, K and Muse, S.V., 2005. Power Marker integrated analysis environment for genetic marker data. *Bioinformatics*, **21**: 2128–2129.
21. Lopez, T., Haucke, V. and Maritzen, T. (2020). Endocytosis in the adaptation to cellular stress. *Cell Stress,***4:**217- 230.
22. 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.
23. 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.
24. 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.
25. 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.
26. Nietlisbach, P., Keller, L. F., and Postma, E. (2016). Genetic variance components and heritability of multiallelic heterozygosity under inbreeding. *Heredity,***116**:1–11.
27. 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.
28. 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.
29. 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.
30. 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.
31. 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.
32. Rouleau, M. Patel, A., Hendzel, M.J. (2010). PARP inhibition: PARP1 and beyond. *Nature,* **10**:293–301.
33. 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.
34. 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.
35. 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.

36. 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.
37. 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 one.***12**:1–15.
38. 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.
39. 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.
40. 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.
41. Weir, B. S. and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution,***38**:1358-1370.
42. Yin, M., Wang, X., and Lu, J. (2020). Advances in IKBKE as a potential target for cancer therapy. *Cancer Med.* **9**:247-258.
43. 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.
44. 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.

Figures

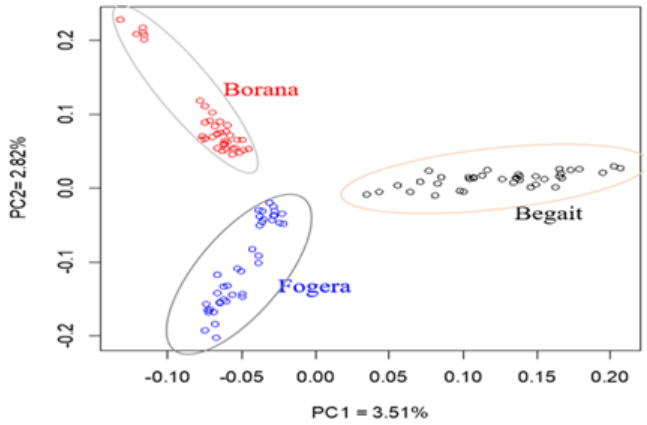

Figure 1

Clustering of Ethiopian cattle populations based on principal component analysis

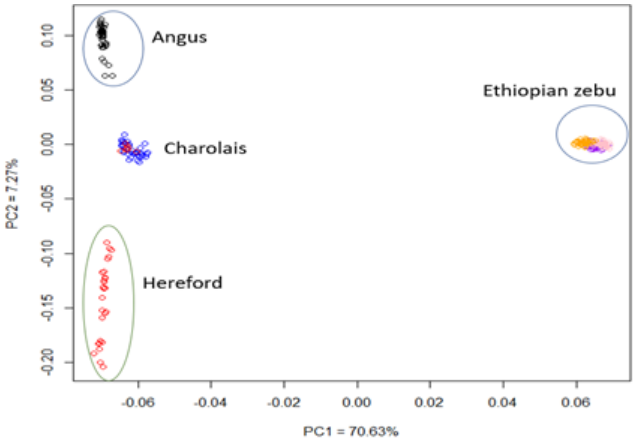

Figure 2

PCA-based Clustering of the study of cattle breeds

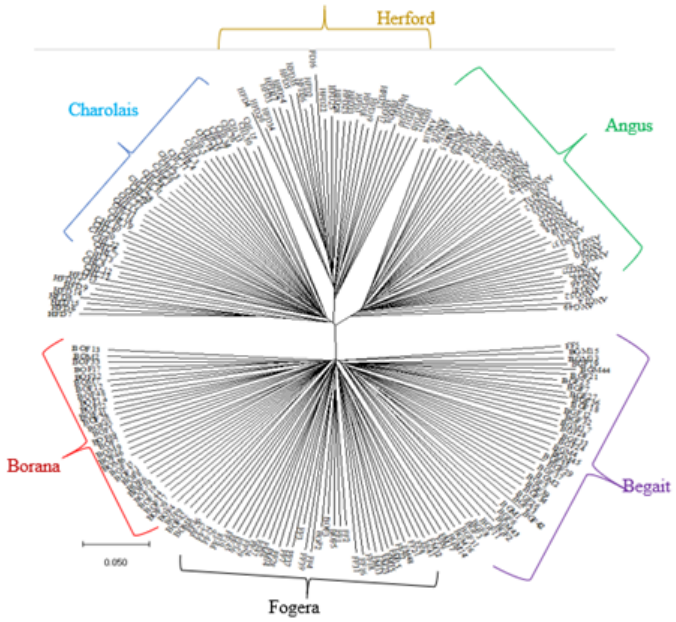

Figure 3

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

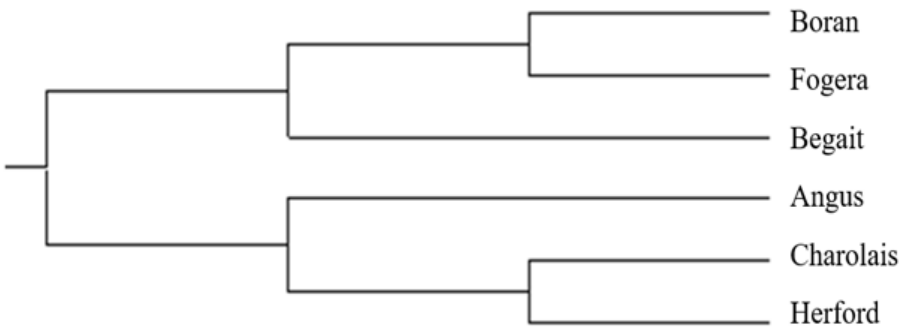

Figure 4

Phylogenetic tree revealing the genetic relationship among the six cattle breeds

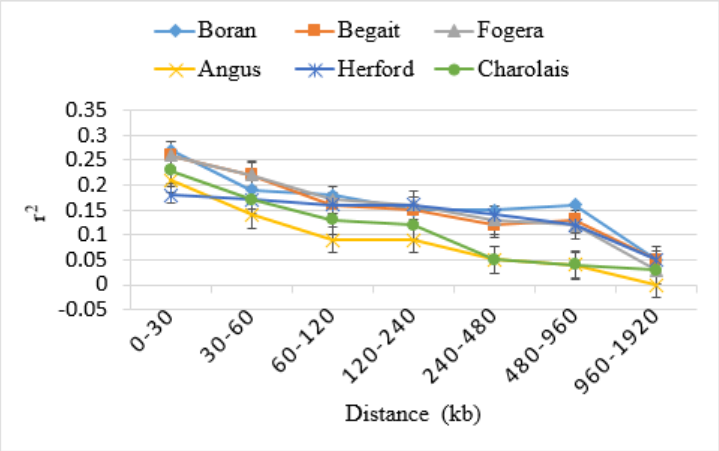

Figure 5

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

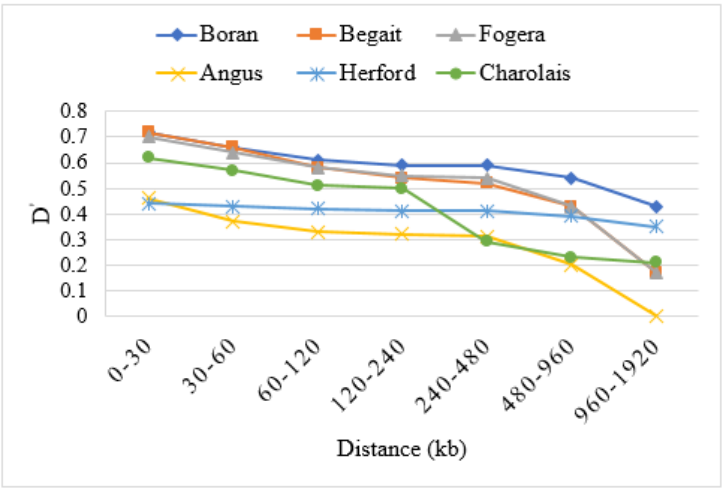

Figure 6

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 2. The mean r^2 values ranged from 0.17 in Angus and Hereford to 0.23 in Begait. Similarly, the mean D' values varied from 067 in Boran to 0.42 in Angus.

Figure 7

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.