

RESEARCH

Open Access



# Genetic diversity of *DGAT1* gene linked to milk production in cattle populations of Ethiopia

Behailu Samuel<sup>1</sup>, Dejenie Mengistie<sup>2</sup>, Ermias Assefa<sup>2</sup>, Mingue Kang<sup>3</sup>, Chankyu Park<sup>3</sup>, Hailu Dadi<sup>2</sup> and Hunduma Dinka<sup>1\*</sup>

## Abstract

**Background:** Diacylglycerol acyl-CoA acyltransferase 1 (*DGAT1*) has become a promising candidate gene for milk production traits because of its important role as a key enzyme in catalyzing the final step of triglyceride synthesis. Thus use of bovine *DGAT1* gene as milk production markers in cattle is well established. However, there is no report on polymorphism of the *DGAT1* gene in Ethiopian cattle breeds. The present study is the first comprehensive report on diversity, evolution, neutrality evaluation and genetic differentiation of *DGAT1* gene in Ethiopian cattle population. The aim of this study was to characterize the genetic variability of exon 8 region of *DGAT1* gene in Ethiopian cattle breeds.

**Results:** Analysis of the level of genetic variability at the population and sequence levels with genetic distance in the breeds considered revealed that studied breeds had 11, 0.615 and 0.010 haplotypes, haplotype diversity and nucleotide diversity respectively. Boran-Holstein showed low minor allele frequency and heterozygosity, while Horro showed low nucleotide and haplotype diversities. The studied cattle *DGAT1* genes were under purifying selection. The neutrality test statistics in most populations were negative and statistically non-significant ( $p > 0.10$ ) and consistent with a populations in genetic equilibrium or in expansion. Analysis for heterozygosity, polymorphic information content and inbreeding coefficient revealed sufficient genetic variation in *DGAT1* gene. The pairwise  $F_{ST}$  values indicated significant differentiation among all the breeds ( $F_{ST} = 0.13$ ;  $p \leq 0.05$ ), besides the rooting from the evolutionary or domestication history of the cattle inferred from the phylogenetic tree based on the neighbourhood joining method. There was four separated cluster among the studied cattle breeds, and they shared a common node from the constructed tree.

**Conclusion:** The cattle populations studied were polymorphic for *DGAT1* locus. The *DGAT1* gene locus is extremely crucial and may provide baseline information for in-depth understanding, exploitation of milk gene variation and could be used as a marker in selection programmes to enhance the production potential and to accelerate the rate of genetic gain in Ethiopian cattle populations exposed to different agro ecology condition.

**Keywords:** *DGAT1*, Ethiopian Cattle breeds, Genetic diversity, Evolution, Sequencing

## Background

The candidate gene, Diacylglycerol acyl-CoA acyltransferase (*DGAT*) activity was first described by Weiss and Kennedy [1] and *DGAT1* enzyme was found to play fundamental role in the metabolism of cellular triacylglycerol during physiological processes, such as intestinal fat absorption, lipoprotein assembly, fat tissue formation

\*Correspondence: hunduma.dinkaa@astu.edu.et

<sup>1</sup> Department of Applied Biology, Adama Science and Technology University, P. O. Box 1888, Adama, Ethiopia

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

and lactation [2]. Functionally, the *DGAT1* gene was identified as one of at least two genes that encodes *DGAT* enzyme which catalyzes the final step of triglyceride synthesis in eukaryotic cells [3]. It became a functional candidate gene for lactation traits after studies indicated that mice lacking both copies of *DGAT1* are completely devoid of milk secretion, presumably due to deficient triglyceride synthesis in the mammary gland [4]. The fluorescence in situ hybridization and radiation hybridization method identified *DGAT1* having profound effect on milk production [5] and primarily responsible for milk fat variation in dairy animals [6]. Bovine *DGAT1* was the first identified gene of about 14,117 bp and 17 exons that encodes a protein with *DGAT* activity [2]. In the centromeric region of the bovine chromosome 14 [7] a missense mutation K232A (Lys232 → Ala) was shown to be significantly associated with variation in milk fat on exon 8 region [5].

According to the studies undertaken so far the reported haplotype number, haplotype diversity and nucleotide diversity for *Bos indicus* cattle were 2, 0.536 and 0.003 respectively [8] and gene diversity for *Bos indicus* cattle range from 0.02–0.50 for *DGAT1* gene [9]. Negative estimates of  $F_{IS}$  (inbreeding coefficient) was observed in Creole and Borgou cattle of Uruguay and Benin cattle respectively [10, 11]. Pairwise  $F_{ST}$  values for pooled subpopulations showed least divergence for *Bos indicus* breeds with high milk fat percentage for *DGAT1* gene [9]. Among the seventeen exons of *DGAT1* gene, exon 8 has previously been reported to be the most polymorphic and potentially affect milk composition and yield traits [5, 7, 12]. Moreover, studies reported that the *DGAT1* gene has been associated with regulation of the synthesis of Vitamin A and somatic cell count in lactating cattle [13, 14]. These all mentioned reports indicate *DGAT1* can be used as practical genetic markers for selective breeding of dairy cattle.

The effect of any identified polymorphism may differ across different populations or breed because of specific genetic backgrounds. Ethiopia has the largest cattle population in Africa and the fifth largest in the world and estimated to be about 70 million and indigenous cattle hold great promise and potential for milk production and constitute about 97.4% of the total cattle population [15]. Some census has witnessed an increase of 1.97% in population of milking cattle from 19.7 million to 22.5 million. For instance, Boran and Begait are *Bos indicus* zebu cattle breeds with a well-developed udder, long legs, and large humps, long teats and known for their milk, resistance to heat stress and tick infestation [16–18]. Moreover Fogera and Horro are *Bos indicus* Zenga (Zebu x Sanga) cattle breeds

mainly characterized by their calm disposition and variable milk production [17–19]. To increase milk volume Boran cattle are crossed with Holstein–Friesian dairy breeds and Boran–Holstein crossbreds have increased lactation lengths, shorter calving intervals and calve at a younger age than the indigenous stock [20, 21]. Before the aforementioned marker is used for the genetic improvement of Ethiopian cattle productivity, its polymorphism should be clearly studied. Unfortunately, no study has been undertaken to understand the genetic diversity within a population, analysis of evolution, neutrality test and *DGAT1* gene population differentiation among Ethiopian cattle breeds that are critically important for future breeding programs. Therefore, the aim of this study was to characterize the genetic variability of exon 8 region of *DGAT1* gene in Ethiopian cattle breeds through sequencing.

## Results

### *DGAT1* gene genetic diversity in Ethiopian cattle breeds

The nucleotide sequences of studied breeds were compared with the reference sequence from the GenBank Accession No.AJ318490 and variability of the region was visualized (Fig. 1).

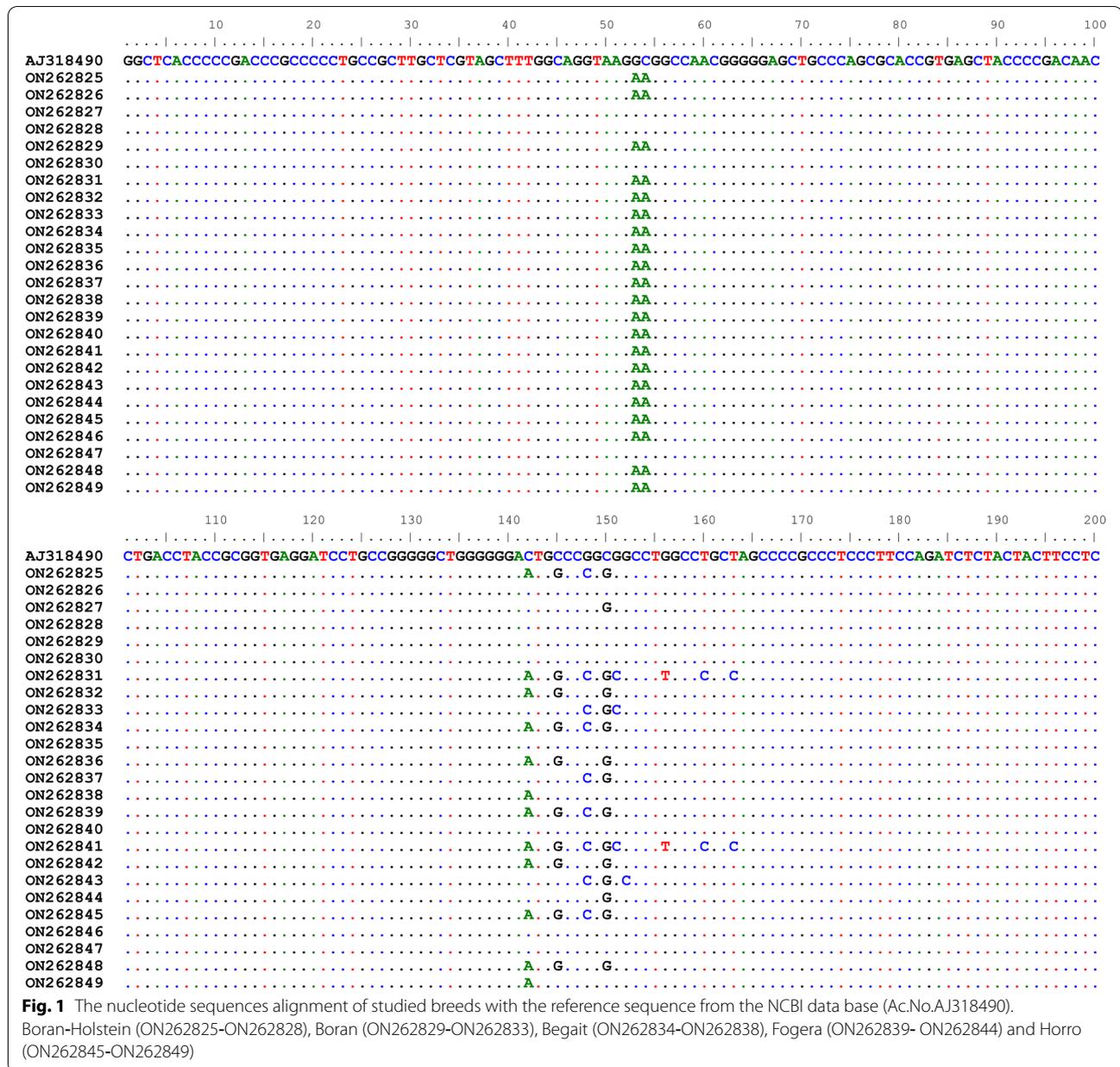
The number of polymorphic sites ( $S$ ) which is the measure of usable loci that show more than one allele per locus was analyzed. Accordingly, the value of  $S$  was higher in Boran (10) and lower in Begait (4). Horro and Boran–Holstein cattle had the same number of polymorphic sites (6). The nucleotide diversities ( $\pi$ ) were relatively high in Begait breeds (0.010) whereas in the Horro breed a relatively low number of  $\pi$  (0.005) was observed. For all breeds, non-synonymous substitutions were less as compared to the synonymous substitutions (Table 1).

Haplotypes were constructed for each breed and a total of 11 haplotypes were obtained. The haplotype values ranged from 6 (Fogera) to 4 (Boran–Holstein). Boran, Begait and Horro cattle had the same number of haplotype number (5).

The haplotype diversity ( $H_d$ ) of the Begait and Boran–Holstein cattle were relatively higher as compared to the Horro. The haplotype diversity for Horro was 0.450 which was the lowest as compared to the rest of the breed (Table 1).

### Evolution of *DGAT1* gene among Ethiopian cattle breeds

The ratio between non-synonymous ( $K_a$ ) and synonymous mutations ( $K_s$ ) was calculated to have an insight on the evolution of *DGAT1* (Table 1). The ratio of non-synonymous substitution ( $K_a$ ) to synonymous substitution ( $K_s$ ) was highest in Fogera with an estimated value of 0.758 and lowest in Boran, Begait and Horro with an estimated value of 0.00.



**DGAT1 gene genetic parameters and neutrality tests in Ethiopian cattle breeds**

For analysis of genetic diversity within a population, minor allele frequency (MAF), polymorphic information content (PIC), average expected ( $H_e$ ) and observed heterozygosity ( $H_o$ ) values, HWE in terms of  $F_{IS}$  coefficient and neutrality tests were estimated. The observed heterozygosity ( $H_o$ ) ranged between 0.157 in Boran-Holstein to 0.413 in Begait cattle (Table 2).

The expected heterozygosity ( $H_e$ ) ranged from 0.120 (Boran-Holstein) to 0.256 (Boran). Boran-Holstein cattle had the lowest number of PIC (0.101), whereas in

Begait the highest numbers of PIC (0.2005) was detected. MAF values ranged from 0.078 in Boran-Holstein to 0.206 in Begait cattle. The different sequences of *DGAT1* gene were observed in heterozygote excess in studied breeds that ultimately lead to low and negative  $F_{IS}$  values (Table 2). Fu's  $F_s$  was negative and statistically non-significant ( $p > 0.10$ ) in Fogera and Horro cattle breeds, while positive in Boran and Boran-Holstein. The Tajima's  $D$  value obtained was negative and statistically non-significant ( $p > 0.10$ ) in Boran, Fogera, and Horro, while positive in Boran-Holstein. Both tests are positive and statistically significant in Begait cattle (Table 2).

**Table 1** DNA polymorphism and evolution of *DGAT1* gene for considered breeds

Breeds	$\pi$	S	H	Hd	Ka	Ks	Ka/Ks
BR	0.009	10	5	0.507	0	0.014	0
BG	0.010	4	5	0.700	0	0.013	0
FG	0.009	9	6	0.514	0.0047	0.0062	0.758
HR	0.005	6	5	0.450	0	0.005	0
BHC	0.008	6	4	0.636	0.0037	0.01	0.37

Nucleotide diversity ( $\pi$ ), number of polymorphic sites (S), haplotype number (H), haplotype diversity (Hd), synonymous (Ks), non-synonymous (Ka), Boran-Holstein (BHC), Boran (BR), Begait (BG), Fogera (FG), Horro (HR)

**Table 2** Genetic parameters estimates and neutrality tests for the considered breeds

Breeds	N	MAF	$H_o$	$H_e$	PIC	$F_{IS}$	Fu's $F_s$	Tajima's D
BR	17	0.129	0.259	0.176	0.148	-0.446	0.941 <sup>ns</sup>	-1.204 <sup>ns</sup>
BG	16	0.206	0.412	0.256	0.200	-0.591	0.461 <sup>*</sup>	2.364 <sup>*</sup>
FG	17	0.153	0.306	0.196	0.163	-0.538	-0.631 <sup>ns</sup>	-0.956 <sup>ns</sup>
HR	16	0.119	0.238	0.146	0.118	-0.606	-1.037 <sup>ns</sup>	-1.214 <sup>ns</sup>
BHC	23	0.078	0.157	0.120	0.101	-0.283	2.463 <sup>ns</sup>	0.258 <sup>ns</sup>

Number of sequences (N), significant (\*) at  $p < 0.10$ , non-significant (ns)

**DGAT1 gene population differentiation among Ethiopian cattle breeds**

Two main parameters ( $F_{ST}$  index and exact G test) for the inspection of population genetic structure and differentiation levels among cattle breed were used. The  $G_{ST}$  (coefficient of interpopulational genetic differentiation) test which is used to measure among-population differentiation, relative to the total diversity value, was compared between cattle breeds and highest value was observed between Boran and Begait (0.096). The lowest  $G_{ST}$  value (-0.016) was found between Boran and Fogera (Table 3). The pairwise  $F_{ST}$  showed significant differences across all cattle breeds ( $F_{ST} = 0.13$ ;  $p$ -value  $\leq 0.0001$ ).

The estimated  $F_{ST}$  values between *Bos taurus* breeds and individual cattle breed considered in this study ranged from 0.005 (*Bos taurus* vs. Boran-Holstein) to 0.389 (*Bos taurus* vs. Begait). Among the Ethiopian breeds, the  $F_{ST}$  value was relatively high ( $F_{ST} = 0.346$ ) between Begait and Boran-Holstein, and a low  $F_{ST}$  value of -0.029 was noted between Boran and Fogera. Moreover,  $F_{ST}$  values were computed across species leading to the detection of higher genetic differences and *Bubalus bubalis* and *Camelus dromaderies* showed least differentiation and maximum differentiation was observed between cattle and other species (Fig. 2).

**Phylogenetic relationships and Median Joining network**

The evolutionary history was inferred by using neighbor joining method based on the Tamura-Nei model with bootstrap replications of 1000 and the evolutionary tree inferred from cattle.

**Table 3** Genetic distances between pairs of populations based on Wright's F-statistics  $F_{ST}$  below the diagonal and Nei's genetic distance  $G_{ST}$  above the diagonal estimated

	BHC	BR	BG	FG	HR
BHC	0	0.03659	0.07786	0.06447	0.05644
BR	0.06655	0	0.09618	-0.01613	-0.02029
BG	0.34631	0.21648	0	0.07531	0.08628
FG	0.15855	-0.02984	0.13728	0	-0.02013
HR	0.12778	-0.03543	0.25748	-0.01595	0

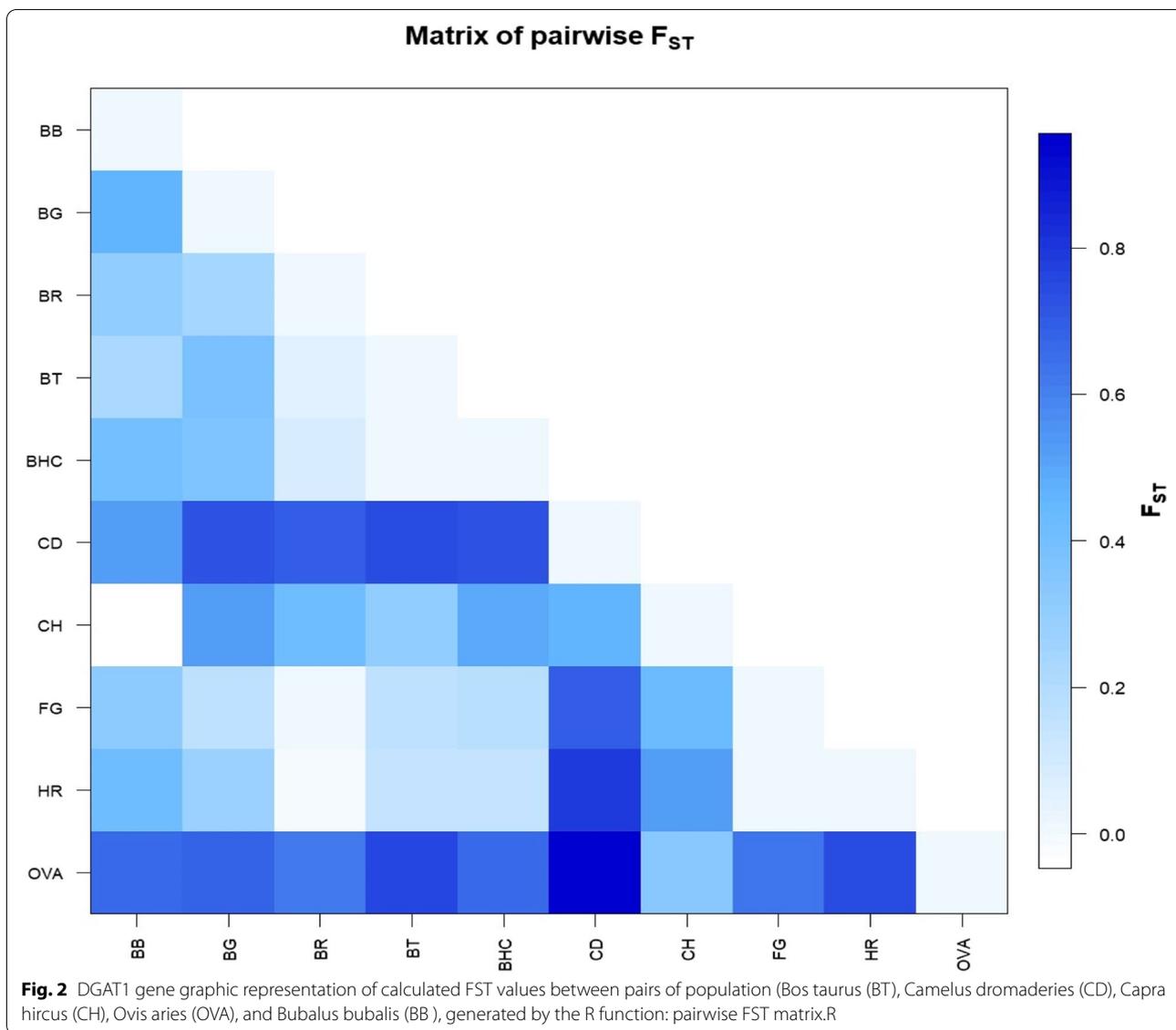
*DGAT1* locus sequences are presented in Figs. 3 and 4.

The analysis revealed four separate clusters among the studied cattle breeds sharing a common node from the constructed tree (Fig. 3). It looks that *DGAT1* alleles might have evolved through multiple lineages.

Furthermore, a median-joining network tree was constructed based on the haplotypes (Fig. 4). The number of individual sequences in H-1, H-2, H-3, H-4, H-5, H-6, H-7, H-8, H-9, H-10 and H-11 was 11,53,1,12,2,4,1,1,2,1, and 1, respectively. Here, H11 being with most probable ancestral haplotype. Haplotypes H5, H7 and H10 separated from the rest of the clusters.

**Discussion**

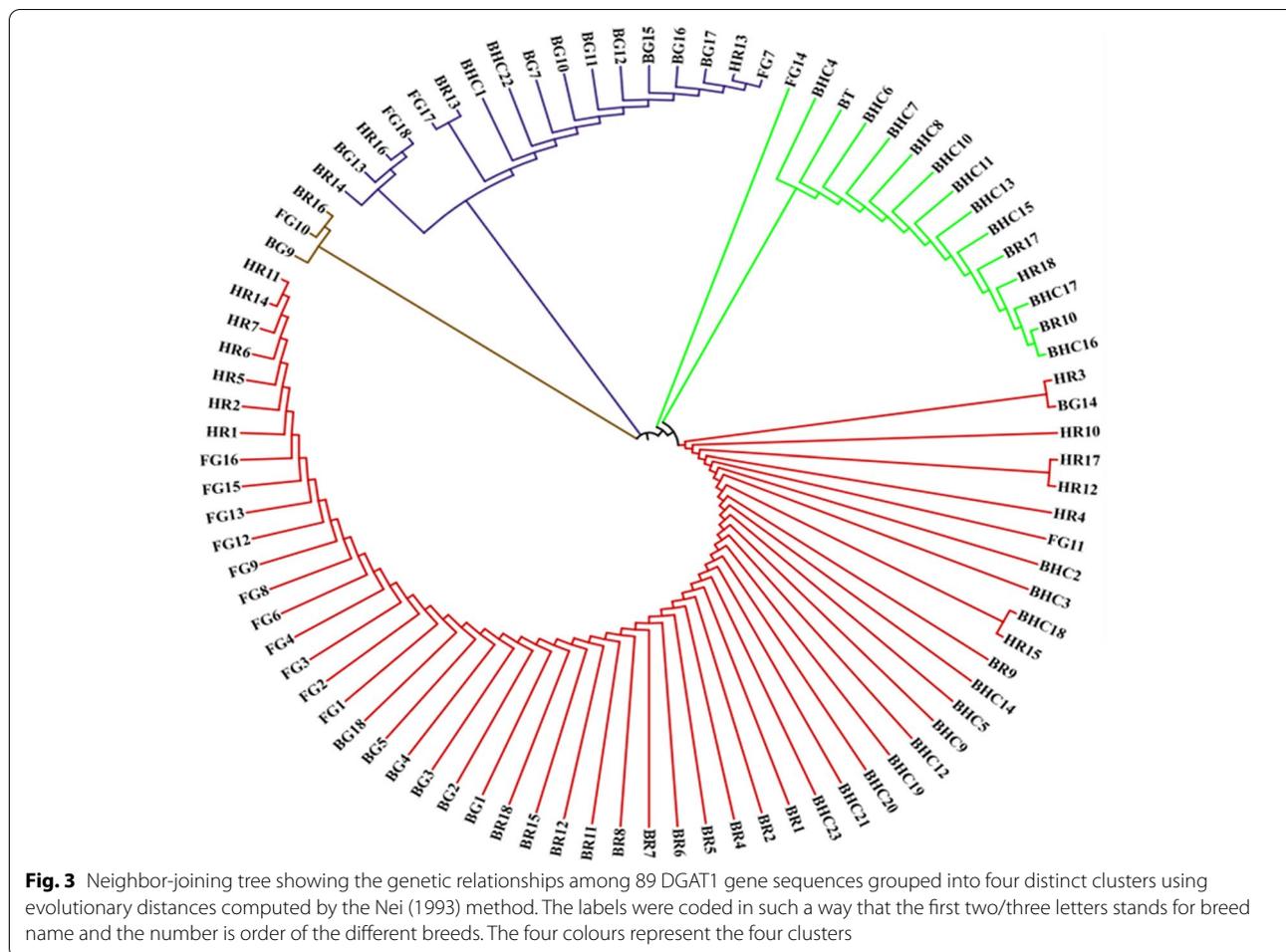
The reported haplotypes number, haplotype diversity and nucleotide diversity were 11, 0.615 and 0.0100 respectively. These values are significantly higher than the values reported by Faraj and his colleagues [8] where



number of haplotypes, haplotype diversity and nucleotide diversity of *bos indicus* cattle were 2, 0.536 and 0.0031, respectively. The combination of high haplotype diversity and low nucleotide diversity, suggested small differences between haplotypes and can be a signature of a rapid population expansion from a small effective population size [22]. The unique breeding histories of some population may account for the high variability of genetic diversity. The ratio of (Ka/Ks) was evaluated and the values recorded were less than one for all the breeds. We found that cattle *DGAT1* gene have been subjected to purifying selection (Ka/Ks=0.00–0.708). The average Ka/Ks ratio for *DGAT1* gene was 0.564, indicating that the evolution of bovine *DGAT1* is largely shaped by strong purifying selection through the removal of alleles that are

deleterious, resulting in stabilizing selection in the phenotypic outcomes. It has been shown that a Ka/Ks ratio of or close to 1 indicates no strong selection pressure, a ratio larger than 1 indicates that the protein is subjected to positive selection, whereas less than 1 indicates the presence of purifying selection [23]. Positive selection is observed less often than purifying selection and most mammalian genes are under strong to moderate purifying selection [23]. *DGAT1* genes were under purifying selection in Ethiopian cattle. Purifying selection against newly arising deleterious mutations is essential to preserve biological function of *DGAT1* gene among Ethiopian cattle.

Minor allele frequency(MAF) refers to the frequency of the second most common allele in a population, and



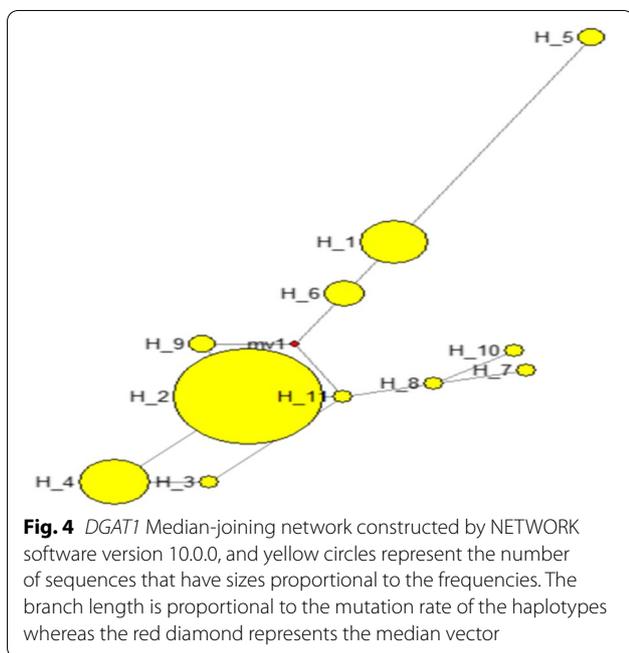
it affects heritability and predictive ability and multiple studies have shown that MAF affects predictive ability [24]. Higher MAF recorded for indigenous cattle than Boran-Holstein. The higher values for indigenous breeds can be explained by the fact that loci used in this study were detected in indicine breeds, and their average MAF was much lower in taurine-indicine breeds. The present MAF values were lower than the report of Edea and his colleagues [25] for Ethiopian cattle. This could be attributed to the differences in genotyping platforms used and causal variants have lower MAF than SNPs in a panel [26].

The heterozygosity situation in the present study for  $H_e$  ranged from 0.120- 0.256 were in the previously reported range of 0.02 to 0.50 for *Bos indicus* cattle [9]. Observed heterozygosity at *DGAT1* locus in this study was higher than the reports of Borgou (0.388) and White Fulani (0.155) cattle breeds of Benin [11]. The values of heterozygosity were lower when compared with previous study on Rathi, Sahiwal and Kankrej cattle breeds of India [27]. Similarly higher values of heterozygosity in Holstein cattle breeds

was reported by previous studies ( $H_o$  ranged from 0.313–0.938 and  $H_e$  ranged from 0.264–0.498) [28]. The highest observed heterozygosity and haplotype diversity in Begait indicates that Begait cattle are more genetically variable at the *DGAT1* locus compared to the other breeds.

A marker with  $PIC > 0.5$  can be considered as highly informative, whereas,  $0.5 > PIC > 0.25$  recognized as reasonably informative and below 0.25 is measured as slightly informative [29]. In the present study breeds PIC values range between 0.101 and 0.200. Thus, the marker is slightly informative for the studied breeds.

Genetic divergence among populations of the same or different breeds is usually quantified by fixation indices or F statistics [30]. The  $F_{IS}$  coefficients are the classical Wright's F-statistic, which estimates the variation within populations. Specifically, it measures the reduction in heterozygosity in an individual caused by nonrandom mating within its subpopulation [30]. The  $F_{IS}$  coefficient value was negative for the studied breeds ranged from -0.284 to -0.606. The present study is in agreement with the findings of Rincon et al. [10] who



also observed negative estimates of  $F_{IS}$  for *DGAT1* loci in Uruguayan Creole cattle population and for Borgou cattle of Benin [11]. Negative estimates of  $F_{IS}$  coefficient value for Ethiopian cattle were also observed [25]. This suggests there is no heterozygosity deficiency in all studied cattle population as a result of uncontrolled mating leading to higher diversity.

The negative  $F_{ST}$  values recorded by Fogera and Horro with Boran as well as between Horro and Fogera indicated that genetic subdivision could not be established among these populations. Moreover, the genetic differentiation of all the breeds based on  $G_{ST}$  was small and supported by the previous studies [25, 27]. The little genetic distance inferred from the breeds could be the result of the evolutionary or domestication history of cattle breeds [31] and which could be due to their common ancestral origin.

Neutrality tests of Tajima's  $D$  [32] and Fu's  $F_s$  statistics [33] were carried out to assess signatures of recent historical demographic events. Tajima's  $D$  test is based on comparison of the allelic frequency of segregating nucleotide sites, while Fu's  $F_s$  test is based on the alleles or haplotypes distribution [32, 33]. They estimate the deviation from neutrality, which is based on the expectation of a constant population size at mutation-drift equilibrium and negative values of both tests signifies an evidence for an excess number of alleles and could be expected from a recent population expansion or genetic hitchhiking and positive value signifies deficiency of alleles, as would be expected from a recent population bottleneck and/or balancing selection [32–34]. Overall Tajima's  $D$  (-0.147)

and Fu's  $F_s$  (-1.534) tests statistics in all populations were negative and statistically non-significant ( $p > 0.10$ ) and consistent with a populations in genetic equilibrium or in expansion [35]. However, the two tests of neutrality were not statistically significant for Boran-Holstein, Boran, Horro and Fogera cattle populations (Table 2). The results suggested that these populations are in genetic equilibrium or in expansion and Tajima's and Fu's neutrality tests were both significant for Begait population, indicate deficiency of alleles, as would be expected from a recent population bottleneck (Table 2). The Begait cattle have been reduced in population size; however, they maintained genetic diversity which is comparable to other studied breeds. The influence of factors which affect genetic diversity can complicate an attempt to interpret the genetic diversity of any population in terms of population size as observed in Begait cattle populations [36]. We observed a close genetic relationship between the Ethiopian cattle breeds from the inferred phylogenetic tree and median joining-network tree. The current result is consistent with the report of other researchers [25, 37]. This is generally interpreted as indicative of a population that has recently expanded in size from a small number of founders following a population bottleneck [31].

## Conclusions

The overall diversity indices showed the cattle populations studied were polymorphic for *DGAT1* locus. *DGAT1* genes were under purifying selection and the presence of high gene diversity, heterozygosity and polymorphic information content revealed sufficient genetic variation in the studied cattle breeds. Fixation indices indicated significant differentiation among all the breeds. This study confirmed that the *DGAT1* gene locus is extremely crucial and may provide baseline information for in-depth understanding, exploitation of milk linked gene variation and could be used as a marker in selection programmes to enhance the production potential and to accelerate the rate of genetic gain in Ethiopian cattle populations exposed to different agro ecology condition.

## Methods

### Sampled populations and genomic DNA extraction

A total of eighty nine animals comprising of five Ethiopian cattle (Boran,  $n = 17$ ; Begait,  $n = 16$ ; Fogera,  $n = 17$ ; Horro,  $n = 16$  and Boran-Holstein,  $n = 23$ ) were considered in this study (Table 1). Animals were selected randomly from each breed with intentional exclusion of closely related ones. Blood samples were collected from the tail head with a volume of 4 ml under aseptic conditions and gently mixed with ethylene diamine tetraacetic acid (EDTA) anticoagulant placed into an ice box containing ice. Extraction of genomic DNA was carried out

using modified salting out extraction procedure [38]. Briefly, blood samples were first thawed at room temperature and about 500  $\mu$ l blood was poured into 2 ml centrifuge tube. Exactly 800  $\mu$ l of lysis buffer (0.3 M sucrose, 0.01 M Tris HCl, pH7.5, 5mMmgcl and1%tritonX100) was added into the centrifuging tube and blood samples were mixed gently by inversion. Then it was centrifuged for 5 min at 4725 rpm. The supernatant was removed carefully and the step was repeated until white pellet obtained. The pellet was vigorously vortexed and re-suspended with the 60  $\mu$ l of 10 mM trisHCl pH 8 and centrifuged again at 2500 rpm for 2 min. After discarding the supernatant, the pellet was re-suspended in 66  $\mu$ l 10 mM tris HCl, 66  $\mu$ l laundry powder solution (30 mg/ml laundry powder solution) and glass beads and vortexed for 2 min. Exactly 50  $\mu$ l of 6 M NaCl was added and vortexed again for 20 s, then centrifuged for 5 min at 11573 rpm and the supernatant was transferred into 2 ml eppendorf tube and 150  $\mu$ l of 96% ethanol was added and the solution was mixed by inversion and centrifuged for 3 min at 13,000 rpm until the genomic DNA was precipitated. The precipitate was then rinsed twice with 100  $\mu$ l of 70% ethanol in 1.5 ml eppendorf tube by inverting the solution for 5 min and centrifuged for 2 min at 12,000 rpm. Then ethanol was carefully poured off using tissue paper. The DNA was allowed to air dry for 30 min. Finally, the extracted DNA was allowed to dissolve in 60  $\mu$ l TE buffer (10 mM tris–HCl pH 8; 0.1 mM EDTA, pH 7.4) overnight and stored at 4°C. The quality of the DNA and its concentration were quantified via NanoDrop1000 and electrophoresis in 1.7% agarose gels. Those DNA samples with good quality and quantity were considered for amplification and sequencing.

#### PCR Amplification and Sequencing of *DGAT1* Gene region

To amplify 278 bp product size of *DGAT1* exon 8 region, primers were designed with accession number AJ318490 as reference sequence using primer 3 plus software [39]. The region was amplified by PCR using the two primers: Forward 5'-AAGGCCAAGGCTGGTGAG -3' and Reverse: 5'-GGCGAAGAGGAAGTAGTAG -3'.

Polymerase chain reaction (PCR) was carried out in a total volume of 25  $\mu$ L containing, 5X PCR buffer (5  $\mu$ l), 1.5 mM MgCl<sub>2</sub> (3  $\mu$ l), 10 Mm dNTP's mix(1  $\mu$ l), forward primer 70 pmol/ $\mu$ l (0.5  $\mu$ l), reverse primer 70 pmol/ $\mu$ l (0.5  $\mu$ l), genomic DNA 25 ng/ $\mu$ l (2  $\mu$ l), Taq DNA polymerase 5U/ $\mu$ l (0.3  $\mu$ l) and DNAase free water (12.7  $\mu$ l). The optimized thermal profile include an initial denaturation at 94 °C for 3 min, 30cycles of denaturation at 94 °C for 1 min, annealing at 57 °C for 45 s, elongation at 72 °C for 1 min and a final extension at 72 °C for 7 min. Finally, the PCR products were visualized post electrophoresis on 1.7% agarose gel with acetate EDTA (TAE)

buffer followed by GelRed staining. For sequencing, the PCR products were sent to Konkuk University, Seoul, South Korea.

Before sequencing, sequencing reaction was performed for the PCR products by using one of a pair of PCR primers used for amplification of the *DGAT1* (Forward) gene as the size was short (278 bp). After completion of the reaction, reaction products were purified using a sodium acetate–ethanol purification method. The purified products of sequencing reactions were analyzed on an ABI3730 capillary genetic analyzer (Sanger Sequencing Machine). Finally, the sequences were analyzed and deposited to GenBank with the following accession numbers (ON262825- ON262849).

#### Data management and statistical analysis

Genetic diversity at sequence level was performed encompassing partial intron 7, exon 8 and intron 8 region of *DGAT1* gene. Prior to analysis, all the chromatograms were visualized and sequence fragments were edited using Bio-edit version 7.0.5.3 and aligned by clustalX2 software package [40]. DNA polymorphism, observed and unbiased expected heterozygosity was computed using ARLEQUIN software version 3.5.2.2 [41]. Minor allele frequency (MAF), Polymorphic information content (PIC) and coefficient of inbreeding ( $F_{IS}$ ) was calculated using Power Marker (version 3.25) [42]. Evolutionary analysis of *DGAT1* exon 8 regions was carried out through analysis of rates of synonymous (Ks) and non-synonymous (Ka) substitutions. Ka/Ks ratio, the average rates of non-synonymous (Ka) over synonymous substitutions (Ks) per site were computed using DnaSP v5 [43].

Population differentiation due to the population genetic structure was also assessed from sequence data, population pairwise Wright's  $F_{ST}$  [44] values were calculated by using ARLEQUIN software version 3.5.2.2 applying 1000 replication values [41]. The pairwise  $F_{ST}$  graph was displayed by Rcmd (console version of the R statistical package) installed on the computer integrated with ARLEQUIN. Moreover, to strengthen the analysis fifteen haplotypes/sequences of milk producing farm animals from Genbank including *Bos taurus* (AJ318490, EU077528, MF069174 and MF445056), *Bubalus bubalis* (MZ230553, MZ230553, MF069172 and KX965992), *Camelus dromedaries* (MF069170 and MF069171), *Capra hircus* (LT221856 and FJ415876) and *Ovies aries* (KJ918741, FJ415875 and EU178818) from Germany, Turkey, India, Iran and Benin were included in the analysis. Hence, population genetic differentiations based on the *DGAT1* genes of the breeds were evaluated by Ne's genetic distance ( $G_{ST}$ ) by DnaSP software [43].

To test for past population expansion, we used two statistical tests Tajima's  $D$  [32] and Fu's  $F_s$  [33]. The analyses

were implemented in the program ARLEQUIN software version 3.5.2.2 [41] and p-values were generated using 1,000 simulations under a model of infinite site neutrality.

Phylogenetic analysis of coding region was carried out for *DGAT1* gene in accordance to neighbor-joining algorithm [45] based on the Tamura-Nei model [46] using MEGA 11 [47] via implementing 1000 bootstrap values [48]. Positions from DNA sequences containing gaps or missing data including identical sequences were excluded from the analysis. A median-joining network (MJN) tree was constructed using the NETWORK software (version 10.0.0) [49]. To evaluate the median network, the nucleotide sequences were first converted into binary data, while identical sites were omitted from the analysis. Each split was programmed as a binary character, satisfying the values of 0 and 1. The haplotypes were denoted as a binary vector in this method [49].

#### Abbreviations

*DGAT1*: Diacylglycerol O-Acyltransferase 1; MAF: Minor allele frequency; NCBI: National Center for Biotechnology information; PIC: Polymorphic information content; rpm: Revolution per minute; TAGs: Triacylglycerols.

#### Acknowledgements

Not applicable

#### Authors' contributions

BS, HD and HDA conceived and designed the study. HD and HDA supervised and reviewed the study. BS, EA and DM analyzed the data. BS collected the blood samples, isolated the DNA, and performed PCR work, analyzed, interpreted and wrote the manuscript. MK and CP sequenced the gene. All authors read and approved for publication.

#### Funding

This research work was supported by the Ethiopian Ministry of Education for data collection and Konkuk University of south Korea for sequencing.

#### Availability of data and materials

The datasets generated and/or analysed during the current study are available in the NCBI- GenBank® repository with accession numbers (ON262825-ON262849).

#### Declarations

##### Ethics approval and consent to participate

All authors declare that animal samples were obtained in compliance with local/national laws in force at the time of sampling. The animals used in this study were owned by governmental institutions (ranches) established for the purpose of research. Data exchange was in accordance with national and international regulations, and approved by the owners. The procedure involving sample collection followed the recommendation of directive 2010/63/EU. In addition, Scientific Ethical Review Board of Adama Science and Technology University approved the experimental design and animal data collection for the present study (certificate reference number RECSOANS/BIO/07/2021). All methods were carried out in accordance with relevant guidelines and regulations.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Department of Applied Biology, Adama Science and Technology University, P. O. Box 1888, Adama, Ethiopia. <sup>2</sup>Bio and Emerging Technology Institute, P. O. Box 5954, Addis Ababa, Ethiopia. <sup>3</sup>Department of Stem Cell and Regenerative Biotechnology, Konkuk University, Seoul, South Korea.

Received: 16 June 2022 Accepted: 3 August 2022

Published online: 10 August 2022

#### References

- Weiss SB, Kennedy EP. The Enzymatic Synthesis of Triglycerides. *J Am Chem Soc* 1956;78. <https://doi.org/10.1021/ja01595a088>.
- Cases S, Smith SJ, Zheng YW, Myers HM, Lear SR, Sande E, et al. Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. *Proc Natl Acad Sci U S A* 1998;95. <https://doi.org/10.1073/pnas.95.22.13018>.
- Yen C-LE, Stone SJ, Koliwad S, Harris C, Farese RV. Thematic Review Series: Glycerolipids. DGAT enzymes and triacylglycerol biosynthesis. *J Lipid Res* 2008;49. <https://doi.org/10.1194/jlr.R800018-jlr200>.
- Smith SJ, Cases S, Jensen DR, Chen HC, Sande E, Tow B, et al. Obesity resistance and multiple mechanisms of triglyceride synthesis in mice lacking Dgat. *Nat Genet* 2000;25. <https://doi.org/10.1038/75651>.
- Winter A, Krämer W, Werner FAO, Kollers S, Kata S, Durstewitz G, et al. Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:Diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. *Proc Natl Acad Sci U S A* 2002;99. <https://doi.org/10.1073/pnas.142293799>.
- Dokso A, Ivanković A, Zečević E, Brka M. Effect of DGAT1 gene variants on milk quantity and quality in Holstein, Simmental and Brown Swiss cattle breeds in Croatia. *Mljekarstvo* 2015;65. <https://doi.org/10.15567/mljekarstvo.2015.0403>.
- Grisart B, Farnir F, Karim L, Cambisano N, Kim JJ, Kvasz A, et al. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proc Natl Acad Sci U S A*. 2004;101:2398–403. <https://doi.org/10.1073/pnas.0308518100>.
- Faraj SH, Ayied AY, Seger DK. DGAT1 gene polymorphism and its relationships with cattle milk yield and chemical composition. *Period Tche Chim* 2020;17. [https://doi.org/10.52571/ptq.v17.n35.2020.16\\_faraj\\_pgs\\_174\\_180.pdf](https://doi.org/10.52571/ptq.v17.n35.2020.16_faraj_pgs_174_180.pdf).
- Kaube B, Winter A, Fries R, Erhardt G. DGAT1 polymorphism in *Bos indicus* and *Bos taurus* cattle breeds. *J Dairy Res*. 2004;71:182–7. <https://doi.org/10.1017/S0022029904000032>.
- Rincón G, Armstrong E, Postiglioni A. Analysis of the population structure of Uruguayan Creole cattle as inferred from milk major gene polymorphisms. *Genet Mol Biol* 2006;29. <https://doi.org/10.1590/S1415-4752006000300016>.
- Houaga I, Muigai AWT, Kyallo M, Githae D, Youssao IAK, Stomeo F. Effect of breed and Diacylglycerol acyltransferase 1 gene polymorphism on milk production traits in Beninese White Fulani and Borgou cows. *Glob J Anim Breed Genet*. 2017;5:403–12.
- Spelman RJ, Ford CA, McElhinney P, Gregory GC, Snell RG. Characterization of the DGAT1 gene in the New Zealand dairy population. *J Dairy Sci*. 2002;85:3514–7. [https://doi.org/10.3168/jds.S0022-0302\(02\)74440-8](https://doi.org/10.3168/jds.S0022-0302(02)74440-8).
- Liu L, Zhang Y, Chen N, Shi X, Tsang B, Yu YH. Upregulation of myocellular DGAT1 augments triglyceride synthesis in skeletal muscle and protects against fat-induced insulin resistance. *J Clin Invest* 2007;117. <https://doi.org/10.1172/JCI30565>.
- Manga I, Riha H. The DGAT1 gene K232A mutation is associated with milk fat content, milk yield and milk somatic cell count in cattle (Short Communication). *Arch Anim Breed* 2011;54. <https://doi.org/10.5194/aab-54-257-2011>.
- CSA. Federal democratic republic of Ethiopia. Central statistical agency. Agricultural sample survey, Volume II, Report on livestock and livestock. Cent Stat Agency (CSA), Addis Ababa, Ethiop. 2013;2:34–5.
- Haile A, Joshi BK, Ayalew W, Tegegne A, Singh A. Genetic evaluation of Ethiopian Boran cattle and their crosses with Holstein Friesian for growth performance in central Ethiopia. *J Anim Breed Genet* 2011;128. <https://doi.org/10.1111/j.1439-0388.2010.00882.x>.

17. Rege JEO, Tawah CL. The state of African cattle genetic resources II. Geographical distribution, characteristics and uses of present-day breeds and strains. *Anim Genet Resour Inf* 1999;26. <https://doi.org/10.1017/s101423390001152>.
18. Zerabruk M, Vangen O, Haile M. The status of cattle genetic resources in North Ethiopia: On-farm characterization of six major cattle breeds. *Anim Genet Resour Inf* 2007;40. <https://doi.org/10.1017/s1014233900002169>.
19. Mekonnen A, Haile A, Dessie T, Mekasha Y. On farm characterization of Horro cattle breed production systems in western Oromia, Ethiopia. *Livest Res Rural Dev*. 2012;24:6–17.
20. Kiwuwa GH, Trail JCM, Kurtu MY, Worku G, Anderson FM, Durkin J. Cross-bred dairy cattle productivity in Arsi Region, Ethiopia. *ILCA Res Rep* No 11. 1983.
21. Effa K. Analysis of longevity traits and lifetime productivity of crossbred dairy cows in the Tropical Highlands of Ethiopia. *J Cell Anim Biol* 2013;7. <https://doi.org/10.5897/jcab2013.0375>.
22. Brown JH. Phylogeography: The History and Formation of Species. John C. Avise. *Q Rev Biol* 2000;75. <https://doi.org/10.1086/393655>.
23. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, et al. Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002;420. <https://doi.org/10.1038/nature01262>.
24. ZHU B, ZHANG J jing, NIU H, GUAN L, GUO P, XU L yang, et al. Effects of marker density and minor allele frequency on genomic prediction for growth traits in Chinese Simmental beef cattle. *J Integr Agric* 2017;16. [https://doi.org/10.1016/S2095-3119\(16\)61474-0](https://doi.org/10.1016/S2095-3119(16)61474-0).
25. Edea Z, Dadi H, Kim SW, Dessie T, Kim K-S. Comparison of SNP Variation and Distribution in Indigenous Ethiopian and Korean Cattle (Hanwoo) Populations. *Genomics Inform* 2012;10. <https://doi.org/10.5808/gi.2012.10.3.200>.
26. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010;42. <https://doi.org/10.1038/ng.608>.
27. Agrawal V, Gahlot G, Ashraf M, Kumar A, Dhakad G. Genetic Analysis of DGAT1 Loci related to Milk production traits in native Sahiwal Cattle. *Int J Livest Res*. 2018;8:136. <https://doi.org/10.5455/ijlr.20170928050917>.
28. Asmarasari SA. The relationship of Diacylglycerol acyltransferase (DGAT1) Gene Diversity to Friesian Holstein Dairy Cattle's Milk Production and Fatty Acid Profile. 2013.
29. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 1980;32:314–31.
30. WRIGHT S. The genetical structure of populations. *Ann Eugen* 1951;15. <https://doi.org/10.2307/2407273>.
31. Bradley DG, Loftus RT, Cunningham P, Machugh DE. Genetics and domestic cattle origins. *Evol Anthropol* 1998;6. [https://doi.org/10.1002/\(SICI\)1520-6505\(1998\)6:3<79::AID-EVAN2>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1520-6505(1998)6:3<79::AID-EVAN2>3.0.CO;2-R).
32. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 1989;123. <https://doi.org/10.1093/genetics/123.3.585>.
33. Fu YX. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 1997;147. <https://doi.org/10.1093/genetics/147.2.915>.
34. Pichler FB. Genetic assessment of population boundaries and gene exchange in Hector's dolphin. vol. 44. Department of Conservation Wellington (New Zealand); 2002.
35. Ramos-Onsins SE, Rozas J. Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* 2002;19. <https://doi.org/10.1093/oxfordjournals.molbev.a004034>.
36. Amos W. Factors affecting levels of genetic diversity in natural populations. *Philos Trans R Soc B Biol Sci* 1998;353. <https://doi.org/10.1098/rstb.1998.0200>.
37. Dadi H, Tibbo M, Takahashi Y, Nomura K, Hanada H, Amano T. Microsatellite analysis reveals high genetic diversity but low genetic structure in Ethiopian indigenous cattle populations. *Anim Genet* 2008;39. <https://doi.org/10.1111/j.1365-2052.2008.01748.x>.
38. Nasiri H, Forouzandeh M, Rasaei MJ, Rahbarizadeh F. Modified salting-out method: High-yield, high-quality genomic DNA extraction from whole blood using laundry detergent. *J Clin Lab Anal* 2005;19. <https://doi.org/10.1002/jcla.20083>.
39. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 2000;132. <https://doi.org/10.1385/1-59259-192-2365>.
40. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics* 2007;23. <https://doi.org/10.1093/bioinformatics/btm404>.
41. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 2010;10. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>.
42. Liu K, Muse SV. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*. 2005;21:2128–9.
43. Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 2009;25. <https://doi.org/10.1093/bioinformatics/btp187>.
44. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution* (N Y) 1984;38. <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x>.
45. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987;4. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>.
46. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 1993;10. <https://doi.org/10.1093/oxfordjournals.molbev.a040023>.
47. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol* 2021;38. <https://doi.org/10.1093/molbev/msab120>.
48. Felsenstein J. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evolution* (N Y) 1985;39. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>.
49. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 1999;16. <https://doi.org/10.1093/oxfordjournals.molbev.a026036>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

