Genetic Diversity and Population Structure of Ethiopian Sheep Populations Revealed by High-Density SNP Markers

Zewdu Edea[1], Tadelle Dessie[2], Hailu Dadi[3], Kyoung-Tag Do[4] and Kwan-Suk Kim[1]*

[1] Department of Animal Science, Chungbuk National University, Cheongju, South Korea, [2] International Livestock Research Institute, Addis Ababa, Ethiopia, [3] Ethiopian Biotechnology Institute, Addis Ababa, Ethiopia, [4] Department of Animal Biotechnology, Faculty of Biotechnology, Jeju National University, Jeju, South Korea

Sheep in Ethiopia are adapted to a wide range of environments, including extreme habitats. Elucidating their genetic diversity is critical for improving breeding strategies and mapping quantitative trait loci associated with productivity. To this end, the present study investigated the genetic diversity and population structure of five Ethiopian sheep populations exhibiting distinct phenotypes and sampled from distinct production environments, including arid lowlands and highlands. To investigate the genetic relationships in greater detail and infer population structure of Ethiopian sheep breeds at the continental and global levels, we analyzed genotypic data of selected sheep breeds from the Ovine SNP50K HapMap dataset. All Ethiopian sheep samples were genotyped with Ovine Infinium HD SNP BeadChip (600K). Mean genetic diversity ranged from 0.29 in Arsi-Bale to 0.32 in Menz sheep, while estimates of genetic differentiation among populations ranged from 0.02 to 0.07, indicating low to moderate differentiation. An analysis of molecular variance revealed that 94.62 and 5.38% of the genetic variation was attributable to differences within and among populations, respectively. Our population structure analysis revealed clustering of five Ethiopian sheep populations according to tail phenotype and geographic origin—i.e., short fat-tailed (very cool high-altitude), long fat-tailed (mid to high-altitude), and fat-rumped (arid low-altitude), with clear evidence of admixture between long fat-tailed populations. North African sheep breeds showed higher levels of within-breed diversity, but were less differentiated than breeds from Eastern and Southern Africa. When African breeds were grouped according to geographic origin (North, South, and East), statistically significant differences were detected among groups (regions). A comparison of population structure between Ethiopian and global sheep breeds showed that fat-tailed breeds from Eastern and Southern Africa clustered together, suggesting that these breeds were introduced to the African continent via the Horn and migrated further south.

Keywords: Ethiopian sheep, high-density chip, genetic diversity, population structure, SNP

INTRODUCTION

In the Horn of Africa and especially in Ethiopia where the economy is predominantly agriculture-based, sheep and their products play a critical role in the livelihood of millions of farmers and pastoralists (Wilson, 2011). Sheep serve as a source of income, mutton, and manure; provides an economic buffer in the event of crop failures; and fulfill many other socio-cultural functions. In some areas such as the cool alpine and arid lowlands where crop production is not a viable economic option, sheep production is the sole option for livelihood (Tibbo, 2006). Sheep are also important for the national economy; indigenous populations have evolved in diverse and harsh environments in Ethiopia where they face disease and parasite burdens, feed shortage, and extreme temperatures. Consequently, these animals likely harbor gene variants uniquely adapted to specific environmental conditions that may not be present in commercial breeds. The economic and agricultural value of sheep is expected to increase as a result of climate change (Seo, 2008; Seo et al., 2010); genetic characterization of local breeds adapted to extreme environments using modern genomic tools can ensure the breeding of hardy sheep populations (Boettcher et al., 2015; Yang et al., 2016).

Given its proximity to the Arabian Peninsula, Ethiopia is considered as a genetic corridor for the introduction of livestock species including sheep to the African continent (Hanotte et al., 2002; Muigai and Hanotte, 2013). Extensive hybridization has occurred between sheep breeds introduced at various times via different routes, making the Horn Africa in general and Ethiopia in particular an excellent resource for the study of genetic diversity in domestic livestock breeds. The ecological, climatological, ethnic, and cultural diversity of Ethiopia is reflected in its large sheep populations (25.5 million heads) (Leta and Mesele, 2014), which can be phenotypically classified into 14 native populations (Gizaw et al., 2007) in addition to populations distributed along the northern, southwestern, and western borders of the country that have yet to be described. These local populations are mainly named after the geographic location or ethnic group/community rearing them, or based on phenotypic characteristics; for instance, the 14 Ethiopian sheep populations are broadly categorized according to their tail phenotypes as thin-tailed (one breed), fat-tailed (11 populations), and fat-rumped (two populations) (Gizaw et al., 2007). The short fat-tailed population mainly inhabits the sub-alpine regions; long fat-tailed sheep are predominant in mid- to high-altitude environments; and fat-rumped sheep are distributed in dry lowland areas (Gizaw et al., 2007).

Characterizing genetic diversity is a key aspect of developing sustainable breed improvement strategies (Groeneveld et al., 2010) and understanding adaptation to extreme environments (Boettcher et al., 2015). Although several studies have investigated the origin of African sheep breeds, many breeds and populations have yet to be fully characterized. The genetic diversity and population structure of Ethiopian sheep populations have been examined using non-recombinant (mitochondrial DNA) and selection-neutral markers (Gizaw et al., 2007; Helen, 2015). However, microsatellite-based studies have provided limited global picture as it included only local sheep breeds of Ethiopia. In general, at the African continent level, there have been far fewer studies on sheep diversity and population structure using genome-wide nuclear markers as compared to non-recombinant markers (Muigai, 2003a; Bruford and Townsend, 2006; Aswani, 2007; Horsburgh and Rhines, 2010; Helen, 2015). It was recently reported that genetic diversity estimated using microsatellites was not correlated with genome-wide single nucleotide polymorphism (SNP) diversity estimates, with larger genetic differentiation values obtained by the former approach (Ciani et al., 2013; Fischer et al., 2017). On the other hand, the large number of genome-wide SNP markers makes it superior to microsatellites for inferring population structure (Glover et al., 2010; Gärke et al., 2012). The recently developed genome-wide high-density ovine SNP array has provided a tool for investigating genetic diversity at a high resolution, inferring population history, and mapping genomic regions subject to selection and adaptation (Kijas et al., 2009; Yang et al., 2016; Zhao et al., 2017). Despite the richness of Ethiopia's sheep genetic resources, only one population was represented in previous genome-wide global sheep analyses (Kijas et al., 2009). Therefore, the extent of genetic variation and patterns of admixture are not known for most Ethiopian sheep populations. Additionally, polymorphisms in the Ovine HD chip in non-reference African/Ethiopia sheep populations have not been identified or validated.

The present study provides the first analysis of high density (~600K) ovine SNPs in Ethiopian sheep breeds. We sampled and genotyped five Ethiopian sheep populations adapted to diverse agro-ecologies using the Infinium HD SNP BeadChip (600K). A detailed understanding of the genetic landscape of national populations requires sampling of representative breeds from wider geographic regions, particularly from a center of domestication and along migration routes (Zhao et al., 2017). To establish historical patterns of admixture and the genetic relatedness of Ethiopian sheep breeds on a broader geographic scale, we compared these breeds with 12 others extracted from Ovine SNP50K HapMap datasets as well as one from Morocco, the data for which was generated by the NextGen Consortium. Two North African sheep breeds (Egyptian Barki and Moroccan) were not previously analyzed (Kijas et al., 2012) but were included here to examine their genetic influence on Ethiopian/East African sheep genetic composition.

MATERIALS AND METHODS

Breeds/Populations and Samples

Nasal samples were collected using Performagene LIVESTOCK's nasal swab DNA collection kit (DNA Genotek, Kanata, ON, Canada) from a total of 72 animals representing five Ethiopian sheep populations: Arsi-Bale, Horro, Menz, Adilo, and Blackhead Somali. Three of these (Horro, Arsi-Bale, and Adilo) are long fat-tailed hairy sheep; Menz is a short fat-tailed coarse-wool sheep; and Blackhead Somali belong to the fat-rumped group (Gizaw et al., 2008). Both female and male animals were randomly sampled from multiple flocks.

Blackhead Somali sheep (also known as Blackhead Ogaden or Berbera Blackhead) exist at low altitudes (500–1000 m above sea level [a.s.l.]) and are well adapted to arid and semi-arid environments characterized by high ambient temperature, low precipitation (200–400 mm), and recurrent drought (Wilson, 1991). The breed is distinguished by the absence of horns in both sexes, black head and neck and white body and limbs, and a fat rump (Wilson, 2011), and is reared across the Horn of African (Ethiopia, Djibouti, Somali, Kenya, and Sudan) under a mobile pastoral management system that includes heavy heat stress, long walks in search of pasture and water, long watering intervals, and few health management practices. In contrast, three of the populations (Horro, Arsi-Bale, and Menz) are reared under sedentary farming systems. Horro sheep are mainly distributed throughout western and southwestern parts of the country inhabiting mid- to high-altitude (1400–2000 m a.s.l.) areas with a mean precipitation of 1000–2000 mm. Horro sheep are characterized by a larger body size and higher twinning rate than other indigenous breeds (Gizaw et al., 2013). Arsi-Bale is the predominant breed in the eastern and south-central parts of Ethiopia, spreading from the Central Great Rift Valley to the Bale mountains (>3000 m a.s.l.). Menz sheep have a relatively small body size with an average live weight of 20.1 ± 0.3 kg, are raised for meat and coarse wool production, and are well adapted to cool highland areas (2500–3000 m a.s.l.) (Haile et al., 2002; Gizaw et al., 2008; Getachew et al., 2015). The Adilo (Wolaita) sheep breed is distributed in southern Ethiopia and characterized by long fat-tail and large body size (Melesse et al., 2013). Phenotypic descriptions and environmental variables of the study sheep populations are summarized in **Table 1**.

To compare the genetic relationship between sheep breeds in Ethiopia and those on other continents and investigate historical patterns of admixture, we also used genotype data of 228 animals representing 12 breeds from North Africa, Middle East, South Africa, Europe, and Asia from the Ovine HapMap project (International Sheep Genomics Consortium[1]). We also included Moroccan sheep data generated by the NextGen Consortium[2]. Details regarding sample sizes, breeds, and geographic origins are summarized in **Table 2**.

Genotyping, Quality Control, and Markers Screening

Ethiopian sheep samples were genotyped with the Ovine Infinium HD BeadChip (Illumina, San Diego, CA, United States) by GeneSeek/Neogen (Lincoln, NE, United States). Among the 606,006 SNPs, 577,401 were autosomal, 1291 were unmapped to any ovine chromosome (OAR), and 27,314 were located on the X chromosome.

Autosomal SNPs with call rates <90% and minor allele frequency (MAF) <0.01 were filtered out, leaving 497,294 SNPs with average and median gaps of 4.92 and 3.58 kb, respectively. Additionally, 11 samples with call rates ≤ 85% were excluded from further analysis. To test for potential effects of

[1] http://www.sheephapmap.org/download.php
[2] http://projects.ensembl.org/nextgen/

TABLE 1 | Summary of phenotypic characteristics and environmental variables of the study sheep populations.

Phenotypes and environmental variables	Adilo (Wolaita)	Arsi-Bale	Horro	Blackhead Somali	Menz
			Population		
Tail-type	Long fat-tail	Long fat-tail	Long fat-tail	Short fat-rump	Short fat-tail
Body weight (kg)	27.40–28.10	26.96–28.60	30.71–35.40	24.92–27.90	20.1
Coat color	Brown, white, gray, and combinations of different colors	Brown, red, black, gray, white, and combinations of different colors	Brown, creamy white/fawn	White body, black head and neck	Black, brown, white and combinations of different colors
Horn	Male short horned, most females polled	Males and most females are horned	Polled	Polled	Horned
Altitude (m)	1618–2043	2492–2810	1400–2000	500–1500	2500–3000
Rainfall (mm)	1175–1449	1020–1076	1000–2000	200–400	900–1102
Temperature (°C)	13.60–25.60	7.87–22.20	10.62–22.53	14.40–35.0	7.6–22.1
Agro-ecology	Wet highland	Wet highland	Wet highland	Arid-lowland	Sub-alpine
Use	Meat	Meat	Meat	Meat	Meat and wool
Management	Mixed crop-livestock	Mixed crop-livestock	Mixed crop-livestock	Pastoral/agro-pastoral	Sheep-barley
Community	Wolaita/Hadiya	Oromo	Oromo	Oromo/Somali	Amhara

Source: Gizaw et al. (2008), Ethiopia Sheep and Goat Productivity Improvement Program (ESGPIP) (2009), Melesse et al. (2013).

TABLE 2 | Diversity indices in 18 sheep breeds estimated from different single nucleotide polymorphism (SNP) datasets.

| Breed/population | Country/origin | n | SNP dataset | | | | | |
| | | | 497,294 | | | | 40,770 | |
			PI_HAT	HO	HE	F	HO	HE
Adilo	Ethiopia/Africa	11	0.09	0.30	0.30	−0.01	0.30	0.30
Arsi-Bale	Ethiopia/Africa	8	0.09	0.30	0.29	−0.03	0.30	0.29
Blackhead Somali	Ethiopia/Africa	15	0.094	0.32	0.30	−0.04	0.31	0.30
Horro	Ethiopia/Africa	15	0.03	0.3	0.31	0.02	0.30	0.31
Menz	Ethiopia/Africa	12	0.08	0.33	0.32	−0.04	0.33	0.31
Overall		61	0.01	0.31	0.33	0.06	0.31	0.33
Red Maasai	Africa/Kenya	20					0.33	0.32
Namaqua Afrikaner	South Africa	12					0.29	0.24
African Dorper	South Africa	21					0.34	0.33
Egyptian Barki	Egypt/Africa	12					0.35	0.35
Moroccan	Morocco/Africa	21					0.36	0.37
Local Awassi	Israel/Middle East	21					0.36	0.36
Afshari	Iran/Middle East	20					0.35	0.34
Indian Garole	Indian/Asia	20					0.29	0.29
Barbados Black Belly	Americas/Caribbean	20					0.32	0.33
Brazilian Creole	Brazil/Americas	20					0.33	0.36
Churra	SW Europe	20					0.36	0.36
Australian Merino	SW Europe	21					0.37	0.37
Dorset Horn	British/North Europe	21					0.32	0.29

HE, expected heterozygosity; HO, observed heterozygosity; PI_HAT, average relatedness; SW, South-west.

ascertainment bias on diversity index estimates, 497,294 SNPs were subjected to linkage disequilibrium (LD) pruning using the parameter (50 5 0.20), yielding 80,602 SNPs.

Genotypic data for the 600K and 50K platforms were merged using SNP and Variation Suite v.8.5.0 (Golden Helix, Bozeman, MT, United States[3]). A total of 41,752 SNPs overlapping between the two platforms were filtered according to quality control criteria; SNPs with call rates <90% and MAF <0.01 were removed, leaving 40,770 SNPs for subsequent analyses. A total of 6163 SNPs remained for population structure analysis after 40,770 SNPs in each population were pruned based on LD using the parameter (50 5 0.80).

Statistical Analysis
Genetic Diversity
Minor allele frequency and deviation from Hardy–Weinberg equilibrium (HWE) were estimated by SNPs for each of the five Ethiopian sheep populations using SNP and Variation Suite v.8.5.0. Alleles were categorized into different bins based on their frequency: fixed alleles (MAF = 0.00), rare alleles (>0.00–<0.05), intermediate alleles (≥0.05–<0.10) and common alleles (≥0.10 and ≤0.5). Diversity indices were estimated from three datasets: (i) 497,294 SNPs that passed the quality control threshold of MAF ≥ 0.01 and call rate ≥ 90%; and (ii) 80,602 SNPs that remained after pruning 497,294 based on LD using the parameter (50 5 0.20) in SNP and Variation Suite v.8.5.0; and (iii) 40,770 SNPs common to 600K and 50K platforms.
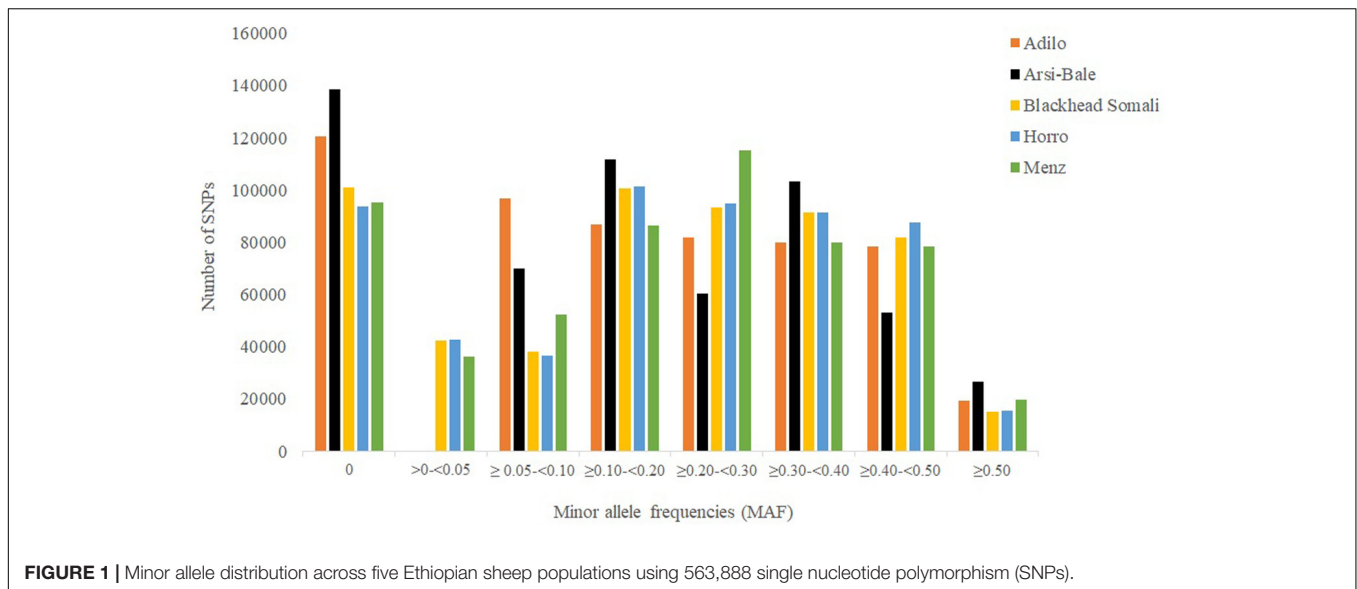
---
[3]www.goldenhelix.com

To estimate within-population genetic diversity, we calculated observed heterozygosity (HO), expected heterozygosity (HE), and inbreeding coefficients for the three datasets using PLINK (Purcell et al., 2007). Animal relatedness was estimated as the proportion of gene identity-by-descent between sample pairs within the breed/population as an average relatedness (PI_HAT) value using the same software.

Genetic Relationships and Population Structure
Pairwise genetic differentiation (fixation index, F_{ST}) (Weir and Cockerham, 1984) and Reynolds' genetic distances (Reynolds et al., 1983) between all pairs of sheep populations were calculated using Arlequin v.3.5.2 (Excoffier and Lischer, 2010). The significance of genetic differences was determined from 10,000 permutation tests. Analysis of molecular variance (AMOVA) with 10,000 permutations was carried out using the same software. Using Reynolds' genetic distance, a neighbor-net tree was constructed using SPLITTREE4 v.14.5 (Huson and Bryant, 2006).

Population genetic structure was assessed using STRUCTURE v.2.3.4 software (Pritchard et al., 2000) using a Bayesian model based on 80,602 SNPs in the five Ethiopian sheep populations and 18 breeds and on 6163 SNPs overlapping between OvineSNP50 and 600K and remaining after pruning based on LD. An admixture ancestry model with correlated allele frequencies was generated for a putative number of subpopulations (K) ranging from 2 to 18. Five runs of 20,000 Markov chain Monte Carlo iterations after a burn-in period of 10,000 iterations were carried out for each K-value. The STRUCTURE output was

FIGURE 1 | Minor allele distribution across five Ethiopian sheep populations using 563,888 single nucleotide polymorphism (SNPs).

analyzed in HARVESTER (Earl, 2012). The most likely number of clusters was identified by the ΔK method (Evanno et al., 2005). Population structure was separately inferred by principal component analysis (PCA) based on 497,294 SNPs for the five Ethiopian sheep populations and 40,770 SNPs for all breeds using SNP and Variation Suite v.8.5.0.

RESULTS AND DISCUSSION

Intra-population Genetic Variability

The mean MAFs for Arsi-Bale, Horro, Adilo, Menz, and Blackhead Somali sheep were 0.19 ± 0.16, 0.21 ± 0.16, 0.20 ± 0.16, 0.21 ± 0.16, and 0.20 ± 0.16, respectively, with an overall mean of 0.20 across populations. For all genotyped animals, the mean MAF ranged from 0.21 for OAR 11, 12, 14, 2, and 24 to 0.23 for OAR 23. These were lower than the reported average value (0.255 ± 0.136) for New Zealand sheep breeds based on an analysis of 517,902 SNPs and those reported for Corriedale and Merino sheep (0.27), but were higher than the value observed in Creole sheep (Grasso et al., 2014).

Minor allele frequency distribution for different categories is shown in **Figure 1**. The percentage of fixed SNPs (MAF = 0.00) varied from 16.60% in Horro to 24.60% in Arsi-Bale sheep, with an overall mean of 8.10% across populations, which is lower than that reported for Creole (27%) but higher than those in Merino (3%) and Corriedale (4%) breeds (Grasso et al., 2014). In total, 45,723 fixed SNPs were shared by the five Ethiopian sheep populations; the common SNPs (≥ 0.10 and ≤ 0.5) accounted for 71.03% of the total and ranged from 58.03% in Adilo to 66.56% in Horro sheep. On average, highly polymorphic SNPs (MAF ≥ 0.30) accounted for 32.69% of total SNPs and ranged from 31.54% in Adilo to 33.40% in Blackhead Somali sheep. The levels of polymorphic SNPs (80.52%, MAF $>$ 0.01) observed in Ethiopian sheep populations were lower than those observed in Merino (89.4%) and Corriedale (86%) sheep, but were higher

than the 69% reported in Creole sheep based on a 50K chip analysis (Grasso et al., 2014). The observed difference between the current and previous studies may be explained by a difference in genotyping platforms and ascertainment bias.

The number of breed-specific SNPs detected for comparison of each breed is given in **Supplementary Table S1**. The highest number of breed-specific SNPs (68,265) was detected in the Menz sheep with frequency ranging from 0.04 to 0.50 and mean of 0.15. The lowest number of breed-specific SNPs (14870) was observed in the Arsi-Bale sheep with frequency ranging from 0.062 to 0.50 and mean of 0.09. Breed-specific SNPs have been detected and used for breed assignment and product traceability in several livestock animals including pigs (Ramos et al., 2011), cattle (Negrini et al., 2009; Ripoli et al., 2013), and sheep (Grasso et al., 2014; Heaton et al., 2014). The population-specific SNPs identified in our Ethiopian sheep populations could be used in a similar manner once they have been validated.

On average, 23,649 (4.19%) loci in the five Ethiopian sheep populations deviated significantly from HWE ($P < 0.05$), with the largest number observed in Adilo (26,348) followed by Blackhead Somali (26,056) sheep. Deviation from HWE is due to inbreeding or genetic substructures within populations (i.e., the Wahlund effect) (Robertson and Hill, 1984; Hart and Clark, 1997; Choi et al., 2009).

The PI_HAT estimated based on 497,294 loci between pairs of individuals was 0.09, 0.03, 0.09, 0.08, and 0.09% for Arsi-Bale, Horro, Adilo, Menz, and Blackhead Somali sheep, respectively, and 0.01% across populations (**Table 2**). HO over all loci (497,292 SNPs) varied from 0.30 in Arsi-Bale, Horro, and Adilo sheep to 0.33 in Menz. The average gene diversity or HE across the five populations was 0.30 and ranged from 0.29 (Arsi-Bale) to 0.32 (Menz). In all populations, HO was higher than or equal to HE, except in Horro sheep. The levels of within-breed genetic variation for Ethiopian sheep populations were within the range reported for New Zealand sheep breeds (0.249–0.383) analyzed using a 600K SNP chip (Brito et al., 2017).

TABLE 3 | Analysis of molecular variance of Ethiopian sheep populations grouped based on tail phenotype from analysis of 497,294 SNPs analysis.

Source of variation	Degree of freedom	Sum of squares	Variance components	Percentage of variation
Among groups	2	452560.919	2614.34624	3.33
Among populations within groups	2	244943.771	2229.52922	2.84
Within populations	117	8605746.777	73553.39126	93.82
Total	121	9303251.467	78397.26672	100

Sample size and the population in which SNPs are detected affect population parameter estimates (Lachance and Tishkoff, 2013; McTavish and Hillis, 2015). Variability is often overestimated in individuals from which the genotyping panel is developed (Rosenblum and Novembre, 2007). We also investigated the effect of ascertainment bias on genetic diversity parameters using loci pruned based on LD. The HE of the unpruned dataset (0.33) was reduced (0.26) after pruning SNPs with high LD within each breed (**Supplementary Table S2**). Removing SNPs in high LD minimizes the effects of ascertainment bias and reduces heterozygosity (Kijas et al., 2012; Edea et al., 2015). In both datasets, estimated inbreeding coefficients (F) were negative in all populations, except in Horro sheep ($F = 0.00$–0.02). Overall inbreeding in all populations was estimated as 0.06. The most inbred individual was an Adilo sheep ($F = 0.30$), whereas the most outbred individual was a Blackhead Somali sheep ($F = -0.33$).

Population Divergence and Relationships

Analysis of molecular variance based on 497,294 autosomal SNPs revealed variations of 5.38% ($P < 0.0001$) and 94.62% among and within populations, respectively. The large within-population variation observed in Ethiopian indigenous sheep populations can be exploited through appropriate breeding strategies to improve productivity. When an analysis was performed for sheep populations grouped based on tail phenotype (long fat-tailed, short fat-tailed, and fat-rumped), among-groups variance was 3.33, with 93.82% within individuals (**Table 3**). Further analysis of populations grouped according to ecological distributions (high- vs. lowland) revealed that 1.28% of the variance was among groups, 4.70% ($P < 0.0001$) among populations within groups, 1.61% among individuals within populations, and 92.47% within populations.

When we previously grouped Ethiopian cattle populations based on their ecological distribution (high- vs. lowland), the estimated among-group variation was 0.42% (Edea et al., 2013), which is lower than the value observed here. The variability among Ethiopian sheep populations was higher than the value of 3.64% reported among five Moroccan sheep breeds based on microsatellite markers (Gaouar et al., 2016).

F_{ST} values and Reynolds' genetic distances among the five Ethiopian sheep populations were estimated using 497,294 SNPs (**Table 4**). The overall F_{ST} value among the five populations was low (0.053) but significant ($P < 0.0001$). F_{ST} for all pairs of populations also differed significantly from zero ($P < 0.001$) and ranged from 0.02 to 0.07, with the closest pairwise value (0.02) observed between Arsi-Bale and Horro sheep. Menz sheep were more distantly related to other Ethiopian sheep populations ($F_{ST} = 0.05$–0.07).

The average F_{ST} among Ethiopian sheep populations was higher than the values reported for Ethiopian cattle (0.01) and goats (0.0245) (Edea et al., 2013; Mekuriaw, 2016), but similar to the mean value of 0.046 obtained using microsatellite markers (Gizaw et al., 2007) and higher than the values in Moroccan [3.6%; (Gaouar et al., 2016)], Algerian [3.8%; (Gaouar et al., 2015)], and Tunisian [3%; (Sassi-Zaidy et al., 2014)] sheep breeds.

Population Structure

To illustrate relationships within individuals and among Ethiopian sheep populations, PCA was performed using 497,294 SNPs. PC1 and PC2 accounted for 26.71 and 25.20%, of the variation, respectively, and clustered the five sheep populations according to their tail phenotypes: long fat-tailed (Arsi-Bale, Horro, and Adilo), short fat-tailed (Menz), and fat-rumped (Blackhead Somali). These clustering patterns corresponded with their geographic distribution. PC1 segregated long-fat-tailed and fat-rumped populations from the short fat-tailed Menz sheep, whereas PC2 separated lowland fat-rumped Blackhead Somali sheep from highland fat-tailed populations (**Figure 2**). Menz sheep formed a tight cluster, whereas outliers were detected in the other populations. The unique genetic background of Menz sheep was corroborated by the STRUCTURE analysis results. At $K = 2$, the three-long fat-tailed sheep populations formed a single group while Menz sheep formed an independent cluster with some admixture from the other populations. Blackhead Somali sheep shared the genetic background of the long fat-tailed populations (**Figure 3**). At $K = 3$, Blackhead Somali sheep tended to segregate, yet shared about 35% of its genome with long fat-tailed populations. The PCA and STRUCTURE analysis revealed clear signatures of admixture among Ethiopian sheep populations—particularly among long-fat tailed breeds—as well as genetic introgression from short-fat tailed Menz into other populations.

TABLE 4 | Pairwise genetic differentiation (F_{ST}) (below diagonal) and Reynolds' genetic distance (above diagonal) among five Ethiopian sheep populations based on an analysis of 497,294 SNPs.
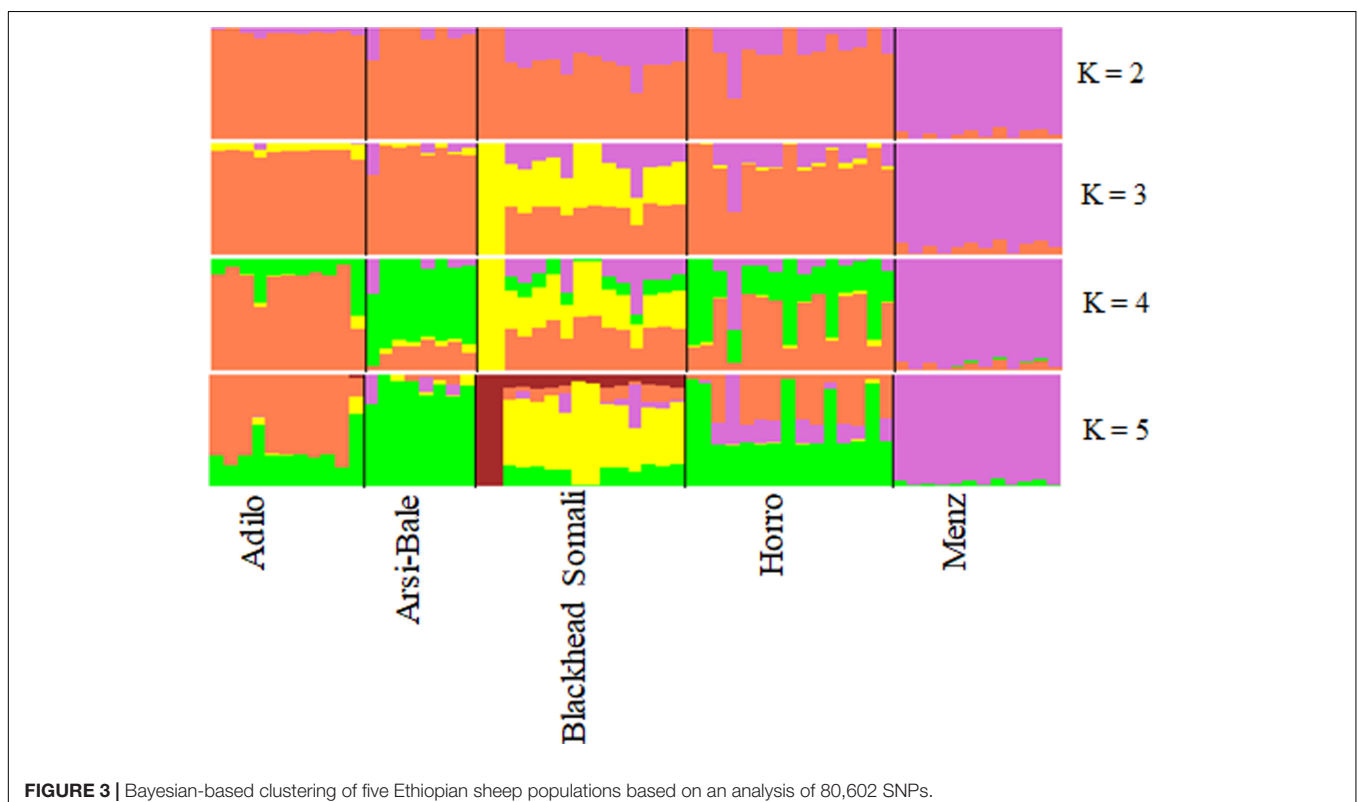
Population	Arsi-Bale	Horro	Adilo	Menz	Blackhead Somali
Adilo	0.046	0.029		0.075	0.062
Arsi-Bale		0.017	0.047	0.076	0.066
Blackhead Somali	0.064	0.052	0.060	0.070	
Horro	0.017		0.030	0.053	0.054
Menz	0.073	0.051	0.073		0.073

FIGURE 2 | Clustering of individual animals in five Ethiopian sheep populations based on an analysis of 497,294 SNPs.
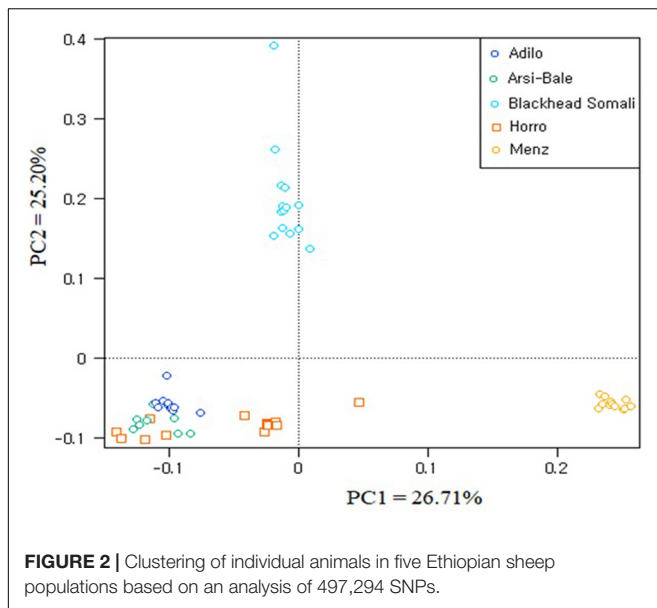
Grouping of populations according to tail-phenotype and ecology is in line with the previous microsatellite based analysis (Gizaw et al., 2008). Morphological variation analysis also grouped Ethiopian sheep populations according to their tail-phenotype (long fat-tailed, short fat-tailed and fat-rumped) and ecological distribution [sub-alpine, wet highland and arid lowlands (Gizaw et al., 2008)]. These results further support

the independent introduction of fat-tailed and fat-rumped sheep into Africa. Accordingly, it was thought that fat-tailed sheep were introduced into Africa during the third wave of migration following thin-tailed hair sheep and thin-tailed wool sheep, whereas fat-rumped sheep entered much later (Epstein, 1971; Ryder, 1984).

As indicated by our genetic distance, PCA and STRUCTURE results, the Menz sheep showed greater genetic differentiation and clearly separated from the rest of the populations. Differences in allele frequencies between Menz sheep and other populations might have been due to selection for ecological adaptation, differences in migration histories and geographical isolation. Menz sheep are evolved in the cool sub-alpine climate of highlands (2500–3000 m a.s.l.) and are kept for meat and coarse wool production (Wilson, 1991; Tibbo, 2006), and are one of the most primitive coarse-wool breeds imported from Arabia via the Bab-el-Mandeb Strait (Wilson, 1991). It is thought that fat-tailed coarse-wooled sheep were introduced to Africa after thin-tailed breeds about 3,000 years ago (Wilson, 2011) for which adequate time has elapsed for adaptive evolution to take place. Furthermore, historical data show that the Amhara ethnic group of Ethiopia have inhabited altitudes more 2500 m for at least 5 ky (Alkorta-Aranburu et al., 2012). The Menz sheep have migrated to new areas and co-exist with humans for centuries under such extreme environments. On the other hand, fat-rumped Blackhead Somali sheep are well adapted to semi-arid to extremely arid lowlands with high temperatures and sparse and erratic precipitation (Wilson, 2011). The breed is kept for meat production and selected for higher fat deposition on the rump

FIGURE 3 | Bayesian-based clustering of five Ethiopian sheep populations based on an analysis of 80,602 SNPs.

as a source of energy-dense food during prolonged dry spells (Muigai and Hanotte, 2013).

The low genetic differentiation between the two-long fat-tailed populations (Arsi-Bale and Horro) was further supported by our population STRUCTURE analysis results. Arsi-Bale and Horro sheep populations are predominantly maintained by the Ethiopian Oromo ethnic group. In addition to geographic isolation, ethnic, cultural, and religious differences may act as barriers to gene flow that shape population genetic structure (Madrigal et al., 2001). The chances of animal exchange are greater within the same ethnic group or tribe than between any two different ethnic groups or tribes (Gizaw et al., 2007). Arsi-Bale and Horro sheep both inhabit highland environments and face common selective pressures, which may have shaped their genomes in a similar manner. We previously reported that Arsi and Horro cattle had the lowest level of genetic differentiation among examined breeds (Edea et al., 2013); our current results imply that sheep dispersal accompanied that of cattle.

Genetic Diversity of Ethiopian Sheep Populations and Their Relationships to Global Sheep Breeds

Genetic Diversity and Relationships

To compare genetic diversity and trace historical patterns of Ethiopian sheep population structure on a broader geographic scale, we analyzed 41,752 SNPs that overlap between Ovine50SNP and 600K chips. Polymorphic (MAF > 0.01) and highly polymorphic (MAF > 0.30) SNPs accounted for 92 and 37% of SNPs in Ethiopian sheep populations, respectively. These values were lower than those observed for Australian Merino (96 and 45%, respectively), but higher than those for Dorset Horn (89 and 34%, respectively). Using the OvineSNP50 chip, highly polymorphic (MAF > 0.30) SNPs accounted for 50% of the total in Merino and Corriedale sheep and for 36% of the total in Creole sheep (Grasso et al., 2014). The relatively high levels of genomic variability observed in Merino sheep may be partly ascribed to ascertainment bias, as these breeds were used in SNP discovery of the OvineSNP50 chip (Kijas et al., 2012). Despite their small sample size, Ethiopian sheep populations show moderate genetic variability relative to southern African Namaqua, Indian Garole, and Dorset Horn (**Table 2**). However, Ethiopian sheep populations show slightly lower levels of genetic diversity than the presumed ancestral breeds of the Near East (Afshari; HE = 0.34) and northern Africa (HE = 0.35–0.37). Breeds from or close to domestication centers are expected to retain higher allelic diversity than those that migrated farther away (Canon et al., 2006; Peter et al., 2007). The higher diversity estimates in North African as compared to East African breeds can be further explained by the fact that these populations reflect a high degree of admixture between fat- and thin-tailed sheep, as demonstrated by our STRUCTURE analysis. Given its close proximity to the Near East and Mediterranean sea, North Africa served as a gateway for early livestock introduction to the African continent and is considered as a secondary hotspot of genetic variation (Gautier, 2002).
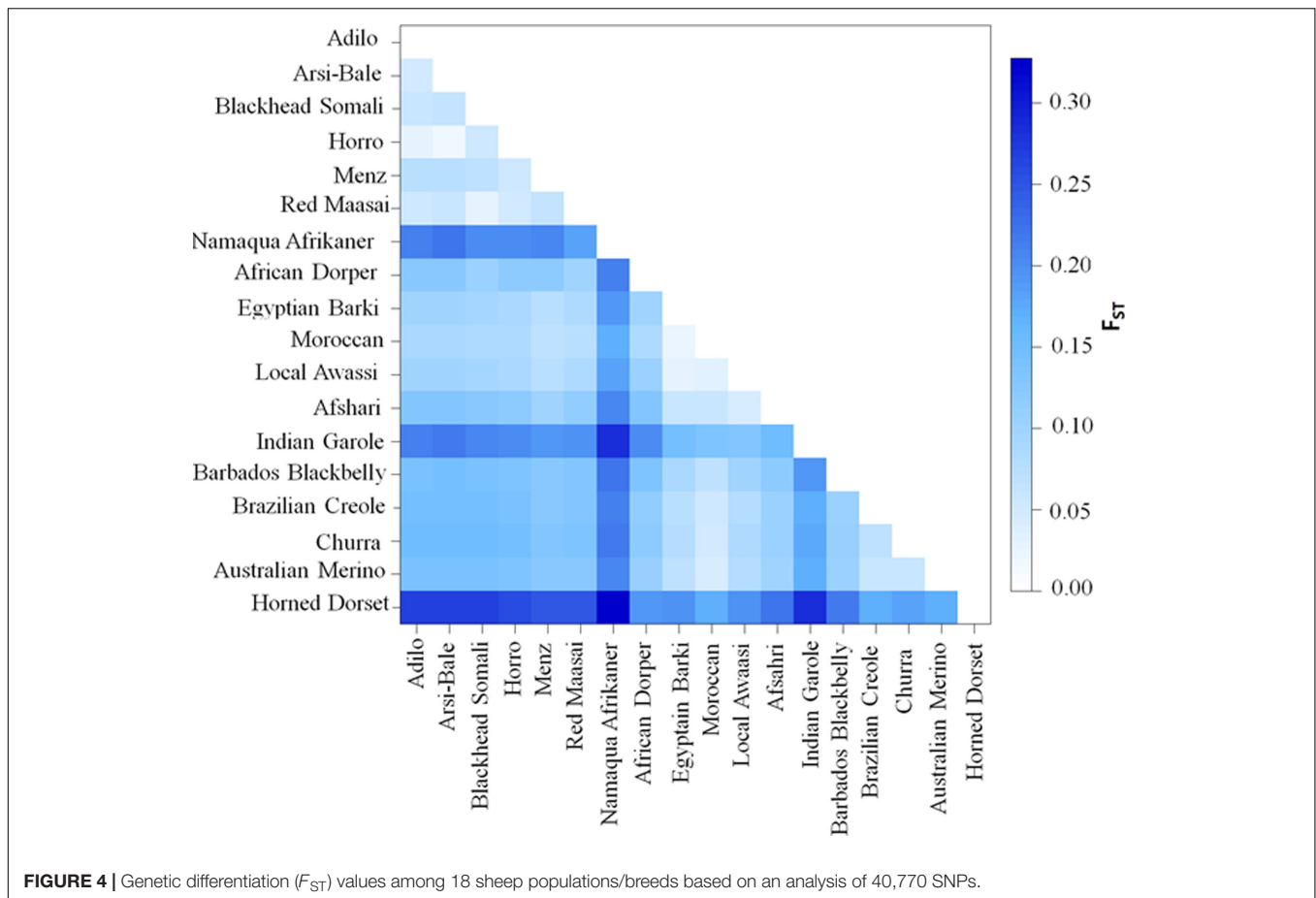
Pairwise F_{ST} (**Figure 4** and **Supplementary Table S3**) and Reynolds' genetic distances (**Supplementary Table S4**) were calculated between the 18 sheep breeds/populations. The lowest differentiation was in Ethiopian populations (Arsi-Bale and Horro; F_{ST} = 0.02) and in North African breeds (Egyptian Barki and Moroccan sheep; F_{ST} = 0.02). Pairwise genetic differentiation comparisons revealed that the highest F_{ST} value (F_{ST} = 0.33) was obtained between the Dorset Horn and Namaqua Afrikaner. Within African sheep breeds, the highest differentiation (mean of 0.21) was observed between Ethiopian and Namaqua Afrikaner breeds. The low within-breed genetic diversity in Namaqua Afrikaner and high genetic differentiation between this breed and other East African sheep populations was likely due to genetic drift, which is consistent with the significantly smaller population size of Namaqua Afrikaner (Qwabe et al., 2013). Ethiopian and North African sheep breeds showed moderate genetic differentiation (F_{ST} = 0.08–0.09), while a higher value detected between East African and Middle Eastern breeds (F_{ST} = 0.12). It is well documented that the Nile River Valley served as a genetic corridor for human and livestock gene flow between the northern and southern parts of the continent across sub-Saharan Africa (Krings et al., 1999; Horsburgh et al., 2013).

Analysis of molecular variance for the 18 global populations grouped based according to geographical regions (Africa, Asia and western) revealed that 3.68% ($P < 0.0001$) of the variance was among groups, 10.64% among populations within groupings, and 85.69% within populations. The F_{ST} value was 0.1431 ($P < 0.0001$), which showed that 14.31% of the total genetic variation was due to population differences. The variation observed among the geographic regions in this study was lower than the reported value of 5.8% (Kijas et al., 2012). To assess genetic differences among the geographic regions within the African continent, we further ran AMOVA by grouping African sheep breeds according to their geographic distribution (North, East, and South). Results indicated that 8.23% ($P = 0.01$) of the variation was among groups and 4.20% among populations within groups. The F_{ST} value was 0.1243, which revealed 12.43% of the total genetic difference was attributed to population differences, and the remaining 87.57% accounted by variation within populations.

Phylogenetic Cluster Analysis

A Neighbor-Net network constructed using 40, 770 SNPs clustered the study population according to their geographic region (**Figure 5**), with close clustering of breeds or populations within a region. Among Ethiopian sheep, the two highland and fat-tailed sheep (Arsi-Bale and Horro) were closely clustered. Despite the observed phenotypic differences, fat-rumped Blackhead Somali sheep were more closely associated with fat-tailed Red Maasai sheep than with fat-tailed Ethiopian sheep populations. These populations are reared under mobile pastoral and agro-pastoral systems, and there is a high chance of inter-population mating in Kenya (Wilson, 1991). The African Dorper—a composite breed developed from Dorset Horn and Blackhead Persia (Kovács et al., 2008)—was closer to Dorset Horn than to Blackhead Somali, which is a strain of Blackhead Persian sheep.

FIGURE 4 | Genetic differentiation (F_{ST}) values among 18 sheep populations/breeds based on an analysis of 40,770 SNPs.
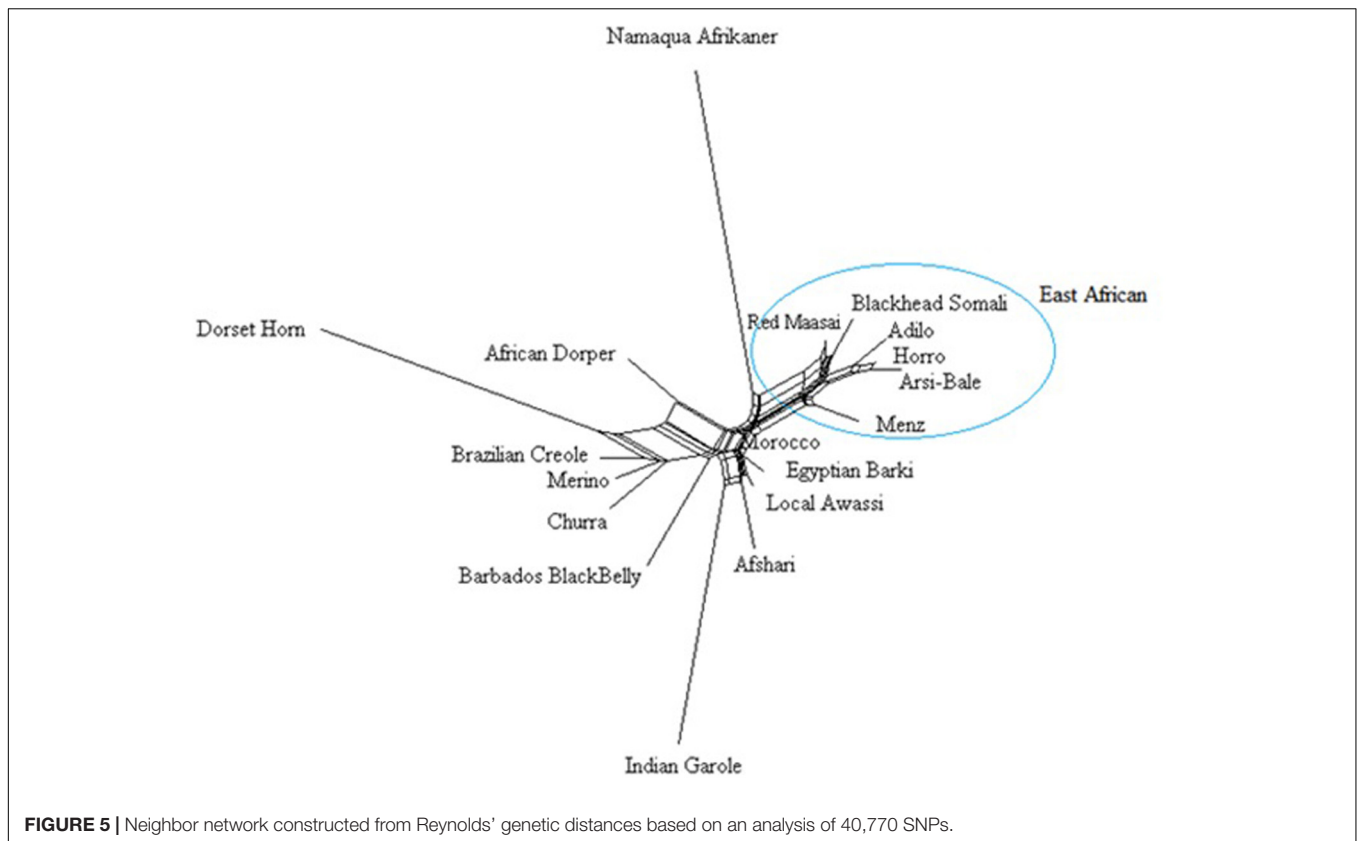
The Middle Eastern breeds (Afsahri and Awassi) formed another group with Egyptian Barki sheep. In the phylogenetic tree, the Moroccan sheep was positioned in an intermediate position. The Brazilian Creole clustered with the Iberian populations (**Figure 5**). Long branches were noted for Namaqua Afrikaner, Dorset Horn and Indian Garole, possibly due to small effective population size, which concurs with previous reports (Kijas et al., 2009; Spangler et al., 2017). These results are supported by population structure and admixture analyses. Despite the observed significant effect of ascertainment bias on genetic diversity, we did not detect any differences in the phylogenetic tree results for 40,770 and 6163 loci subjected to LD pruning (data not shown). In agreement with our results, it has been demonstrated that increasing the number of loci does not improve the reliability of the phylogenetic tree (Litt and Luty, 1989).

Population Structure Analyses

Principal component analysis was carried out using 40,770 SNPs overlapping between OvineSNP50 and Ovine HD SNPs and the 6163 SNPs left after LD pruning (**Figure 6** and **Supplementary Figure S1**). PC1 accounted 21.14% of the total variation and separated the African breeds, except Moroccan sheep from the Western breeds. Menz and Namaqua Afrikaner were closer to the rest of the East African population but remained as a separate

cluster. Eastern and Southern African breeds were separated from the Middle Eastern and North African breeds by PC2. Admixed populations should fall between their two ancestral populations, and the proportion of ancestry inherited from each can be linearly estimated (McVean, 2009). Accordingly, the African composite Dorper was positioned between Dorset Horn and East African populations, while Egyptian Barki sheep were proximal to the Middle Eastern Awassi breed. These results were consistent for 40,770 SNPs and the 6163 SNPs remaining after pruning based on LD, revealing a lack of strong ascertainment bias (**Supplementary Figure S1**).

The results of the structural analysis for hypothetical populations ranging from 2 to 10 are shown in **Figure 7**. At $K = 2$ and $K = 3$, Eastern and Southern African sheep formed one group and Dorset Horn was an independent cluster, which was supported by the PCA results. At $K = 3$, thin-tailed Indian Garole was separated from the other breeds. From $K = 4$–10, Namaqua Afrikaner sheep clearly segregated from East African populations, which was well supported by the phylogenetic results. Northern Africa is mostly populated by fat-tailed sheep (Muigai and Hanotte, 2013), but our STRUCTURE analysis revealed substantial signatures of admixture in the genomes of North Africa populations as compared to their Eastern and Southern African counterparts. This is in accordance with

FIGURE 5 | Neighbor network constructed from Reynolds' genetic distances based on an analysis of 40,770 SNPs.

the historical introduction of sheep into Africa and their dispersion across the continent through the Nile Valley; for instance, thin-tailed sheep spread into the Western Sahara via northern Africa (Muigai and Hanotte, 2013), which may have left its genomic legacy in today's North African sheep populations.

The low genetic background of Asiatic and Iberian thin-tailed sheep detected in fat-tailed East and South African breeds is consistent with the distinct histories and non-overlapping geographic distributions of these populations (Muigai, 2003b), and support the predominance of fat-tailed sheep in the eastern and southern parts of Africa (Muigai and Hanotte, 2013). Archeological evidence traces the first fat-tailed sheep to the Eastern Ethiopian highlands (Clark and Williams, 1978). Moreover, analyses of autosomal markers and the Y chromosome have revealed the distinct evolutionary histories of thin- and fat-tailed African sheep breeds (Muigai, 2003a; Aswani, 2007).

At $K = 8$, we observed a divergence of the African Dorper from the East African populations, which was also well supported by our PCA and Neighbor-Net network results. At $K = 6$–10, Menz sheep shared 20–22% its genome with Middle Eastern fat-tailed sheep, whereas this value did not exceed 1% in the remaining Ethiopian sheep populations. The influence of Middle Eastern fat-tailed sheep detected in Menz can be explained by the fact that within Menz and adjacent areas, cross-breeding between Menz and Awassi populations has been ongoing for more than three decades (Gizaw and Getachew,

2009). At the optimum K-value of 10, Red Maasai shared between 8 and 10% of its genomes with African Dorper. It is well known that the Dorper breed was introduced into Kenya in the 1960s and was indiscriminately crossed with local breeds including Red Maasai to increase meat production in local sheep populations (Verbeek et al., 2007). Similarly, Blackhead Somali—which is a strain of Blackhead Persian sheep—was used as a maternal line in the development of African Dorper (Wilson, 1991). The sizeable genetic admixture between Iberian and North African breeds, particularly with Moroccan sheep was clearly illustrated at $K = 5$–9. This finding mirrors historical human and livestock movements between Northern Africa and the Iberian Peninsula (Boone and Benco, 1999; Botigué et al., 2013); archeological and DNA evidence demonstrates the influence of North African domestic livestock species on indigenous populations of the Iberian Peninsula (Beja-Pereira et al., 2002; Anderung et al., 2005).

The close clustering of East African sheep populations and distinct separation from their northern counterparts was well demonstrated by our phylogenetic, PCA, and STRUCTURE analyses. This result coincides with the evidence that fat-tailed sheep were introduced into Africa via two independent routes: the Horn of Africa and northern Africa from the Middle East (Ryder, 1984). The lowest genetic differentiation obtained for the two Ethiopian sheep populations (Arsi-Bale and Horro; $F_{ST} = 0.02$) was also well supported by population STRUCTURE
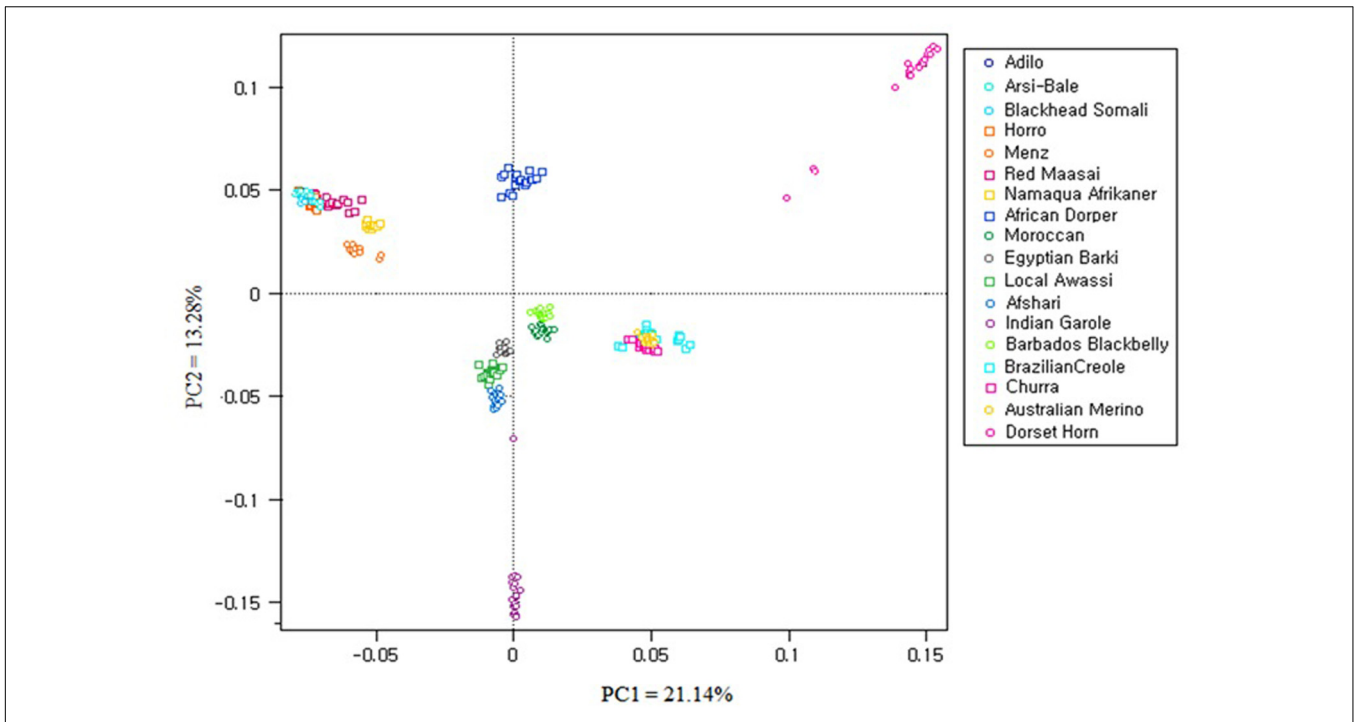
FIGURE 6 | Principal component analysis (PCA) results of 18 sheep breeds from a dataset of 40,770 SNPs.
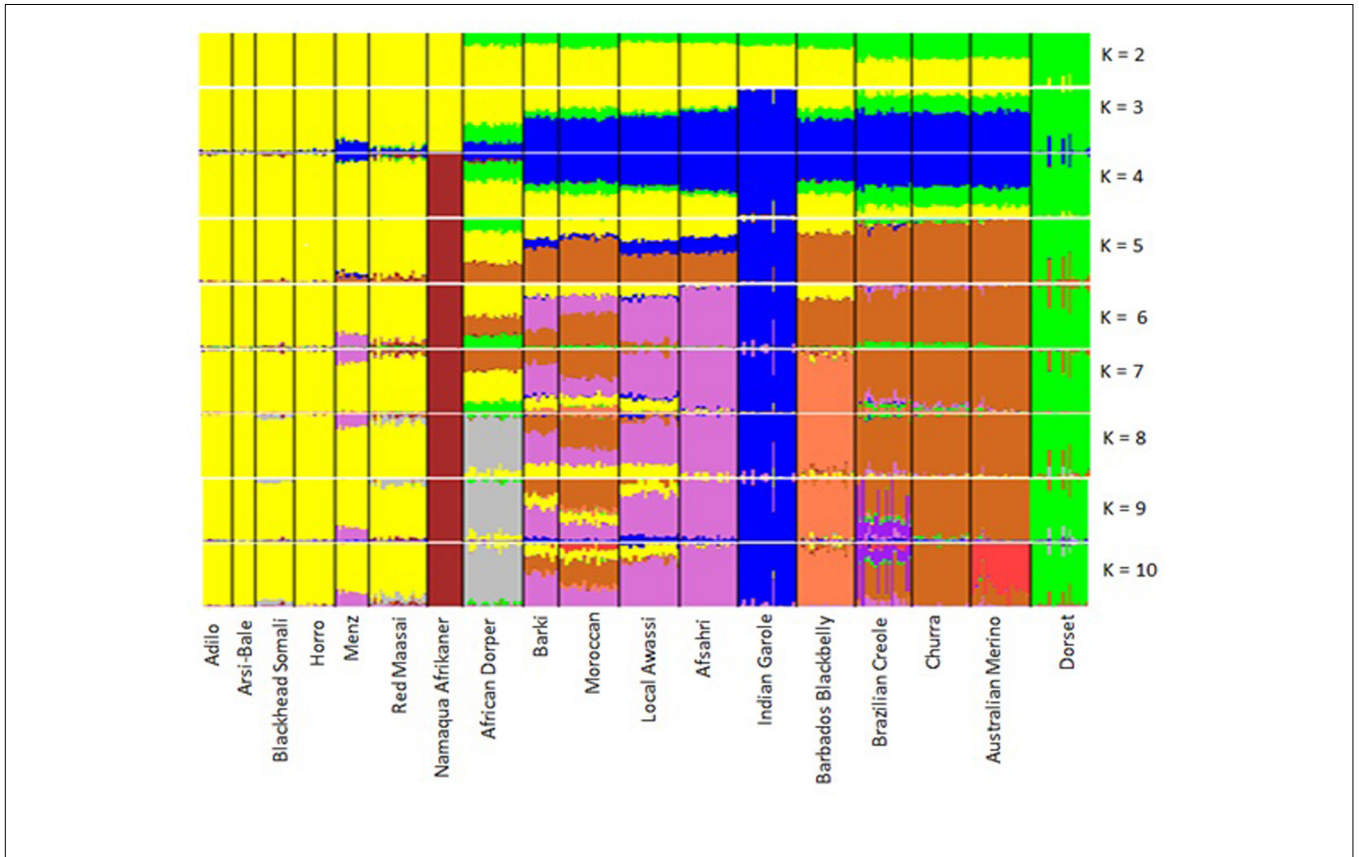
FIGURE 7 | Clustering of 18 sheep breeds based on an analysis of 6163 SNPs for *K* = 2–10.

and Neighbor network analyses. We suggest that this could be due to gene flow and similarity of production environments (Gizaw et al., 2007). On the other hand, the unique genetic composition of short fat-tailed Menz sheep is consistent with its distinct phenotypes, population histories, and ecological distribution (Gizaw et al., 2007).

CONCLUSION

Our high-density genome-wide SNP analyses revealed that Ethiopian sheep populations are roughly clustered according to their geographic distribution and tail phenotype. The genetic diversity and structure of Ethiopian sheep populations can be explained by historical events and selection for ecological adaptation. The high-density SNP data generated in this study can be used to identify genes and pathways relevant for physiological adaptation to extreme environments and variation in phenotypic traits. The close clustering of Eastern African breeds and their separation from North African breeds provide evidence that fat-tailed sheep were introduced to the continent via the Horn of Africa and migrated further southwards. Additional genome-wide analyses of thin-tailed sheep breeds from Eastern and Western Africa and fat-tailed breeds from the Arabian Peninsula can clarify the evolutionary history of sheep on the African continent and provide new insight into the genomic landscape of African sheep breeds.

DATA ACCESSIBILITY

Genotypic data of 72 animals representing five Ethiopian sheep populations are deposited and available at (https://www.animalgenome.org/repository/pub/KORE2017.1122/).

ETHICS STATEMENT

Local regulations were observed. This research used Nasal swab DNA collection kits, which does not require injure the animal nor impose pain.

AUTHOR CONTRIBUTIONS

ZE and K-SK conceived the study, analyzed the data, and wrote the manuscript. TD provided logistical support for field data collection and facilitated sample export. HD and K-TD revised the manuscript. All authors read and approved the final manuscript.

FUNDING

ACKNOWLEDGMENT

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2017.00218/full#supplementary-material

FIGURE S1 | Results of PC 1 and 2 from the dataset of 6163 SNP markers.

TABLE S1 | The number of private alleles detected in the comparison of each sheep population.

TABLE S2 | Diversity indices in 5 Ethiopian sheep populations estimated from 80,602 SNPs obtained after LD pruning.

TABLE S3 | Genetic differentiation (FST) between the study sheep breeds based on analysis of 40,770 SNPs.

TABLE S4 | Reynolds' genetic distance between the 18 sheep breeds based on analysis of 40,770 SNPs.

REFERENCES

Alkorta-Aranburu, G., Beall, C. M., Witonsky, D. B., Gebremedhin, A., Pritchard, J. K., and Di Rienzo, A. (2012). The genetic architecture of adaptations to high altitude in ethiopia. *PLOS Genet.* 8:e1003110. doi: 10.1371/journal.pgen.1003110

Anderung, C., Bouwman, A., Persson, P., Carretero, J. M., Ortega, A. I., Elburg, R., et al. (2005). Prehistoric contacts over the Straits of Gibraltar indicated by genetic analysis of Iberian Bronze Age cattle. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8431–8435. doi: 10.1073/pnas.050339 6102

Aswani, O. (2007). *Genetic Diversity and Relationships of African Sheep: AY Chromosome Perspective.* Nakuru: Egerton University.

Beja-Pereira, A., Erhardt, G., Matos, C., Gama, L., and Ferrand, N. (2002). Evidence for a geographical cline of casein haplotypes in Portuguese cattle breeds. *Anim. Genet.* 33, 295–300. doi: 10.1046/j.1365-2052.2002.00871.x

Boettcher, P. J., Hoffmann, I., Baumung, R., Drucker, A. G., McManus, C., and Berg, P. (2015). Genetic resources and genomics for adaptation of livestock to climate change. *Front. Genet.* 5:461. doi: 10.3389/fgene.2014.00461

Boone, J. L., and Benco, N. L. (1999). Islamic settlement in North Africa and the Iberian peninsula. *Annu. Rev. Anthropol.* 28, 51–71. doi: 10.1146/annurev.anthro.28.1.51

Botigué, L. R., Henn, B. M., Gravel, S., Maples, B. K., Gignoux, C. R., Corona, E., et al. (2013). Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11791–11796. doi: 10.1073/pnas.1306223110

Brito, L. F., McEwan, J. C., Miller, S. P., Pickering, N. K., Bain, W. E., and Dodds, K. G. (2017). Genetic diversity of a New Zealand multi-breed sheep population and composite breeds' history revealed by a high-density SNP chip. *BMC Genet.* 18:25. doi: 10.1186/s12863-017-0492-8

Bruford, M. W., and Townsend, S. J. (2006). "Mitochondrial DNA diversity in modern sheep," in *Documenting Domestication: New Genetic and Archaeological Paradigms*, ed. M. A. Zeder (Berkeley, CA: University of California Press), 306–316.

Canon, J., Garcia, D., Garcia-Atance, M. A., Obexer-Ruff, G., Lenstra, J. A., Ajmone-Marsan, P., et al. (2006). Geographical partitioning of goat diversity in Europe and the Middle East. *Anim. Genet.* 37, 327–334. doi: 10.1111/j.1365-2052.2006.01461.x

Choi, Y., Wijsman, E. M., and Weir, B. S. (2009). Case-control association testing in the presence of unknown relationships. *Genet. Epidemiol.* 33, 668–678. doi: 10.1002/gepi.20418

Ciani, E., Cecchi, F., Castellana, E., D'Andrea, M., Incoronato, C., D'Angelo, F., et al. (2013). Poorer resolution of low-density SNP vs. STR markers in reconstructing genetic relationships among seven Italian sheep breeds. *Large Anim. Rev.* 19, 236–241.

Clark, J. D., and Williams, M. (1978). Recent archaeological research in southeastern Ethiopia (1974-1975): some preliminary results. *Ann. Ethiopie* 11, 19–44. doi: 10.3406/ethio.1978.902

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Edea, Z., Bhuiyan, M., Dessie, T., Rothschild, M., Dadi, H., and Kim, K. (2015). Genome-wide genetic diversity, population structure and admixture analysis in African and Asian cattle breeds. *Animal* 9, 218–226. doi: 10.1017/S1751731114002560

Edea, Z., Dadi, H., Kim, S. W., Dessie, T., Lee, T., Kim, H., et al. (2013). Genetic diversity, population structure and relationships in indigenous cattle populations of Ethiopia and Korean Hanwoo breeds using SNP markers. *Front. Genet.* 4:35. doi: 10.3389/fgene.2013.00035

Epstein, H. (1971). *The Origin of Domestic Animals of Africa*, Vol. 2. New York, NY: Africana Publication Corporation, 719.

Ethiopia Sheep and Goat Productivity Improvement Program [ESGPIP] (2009). *Goat Breeds of Ethiopia: a Guide for Identification and Utilization*. Technical Bulletin 27. Addis Ababa: ESGPIP.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., et al. (2017). Estimating genomic diversity and population differentiation – an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics* 18:69. doi: 10.1186/s12864-016-3459-7

Gaouar, S. B. S., Da Silva, A., Ciani, E., Kdidi, S., Aouissat, M., Dhimi, L., et al. (2015). Admixture and local breed marginalization threaten Algerian sheep diversity. *PLOS ONE* 10:e0122667. doi: 10.1371/journal.pone.0122667

Gaouar, S. B. S., Kdidi, S., and Ouragh, L. (2016). Estimating population structure and genetic diversity of five Moroccan sheep breeds by microsatellite markers. *Small Rumin. Res.* 144, 23–27. doi: 10.1016/j.smallrumres.2016.07.021

Gärke, C., Ytournel, F., Bed'hom, B., Gut, I., Lathrop, M., Weigend, S., et al. (2012). Comparison of SNPs and microsatellites for assessing the genetic structure of chicken populations. *Anim. Genet.* 43, 419–428. doi: 10.1111/j.1365-2052.2011.02284.x

Gautier, A. (2002). "The evidence for the earliest livestock in North Africa: or adventures with large bovids, ovicaprids, dogs and pigs," in *Droughts, Food and Culture: Ecological Change and Food Security in Africa's Later*, ed. F. A. Hassan (Berlin: Springer Science & Business Media), 195–207.

Getachew, T., Alemu, B., Sölkner, J., Gizaw, S., Haile, A., Gosheme, S., et al. (2015). Relative resistance of Menz and Washera sheep breeds to artificial infection with *Haemonchus contortus* in the highlands of Ethiopia. *Trop. Anim. Health Prod.* 47, 961–968. doi: 10.1007/s11250-015-0815-6

Gizaw, S., Abegaz, S., Rischkowsky, B., Haile, A., Mwai, A. O., and Dessie, T. (2013). *Review of Sheep Research and Development Projects in Ethiopia*. Nairobi: International Livestock Research Institute (ILRI).

Gizaw, S., and Getachew, T. (2009). "The Awassi × Menz sheep crossbreeding project in Ethiopia: achievements, challenges and lessons learned," in *Proceedings of the Mid-term Conference of the Ethiopian Sheep and Goat Productivity Improvement Program (ESGPIP)*, Hawassa, 13–14.

Gizaw, S., Komen, H., Hanotte, O., and Van Arendonk, J. (2008). Indigenous sheep resources of Ethiopia: types, production systems and farmers preferences. *Anim. Genet. Resour.* 43, 25–39. doi: 10.1017/S1014233900002704

Gizaw, S., Van Arendonk, J. A., Komen, H., Windig, J., and Hanotte, O. (2007). Population structure, genetic variation and morphological diversity in indigenous sheep of Ethiopia. *Anim. Genet.* 38, 621–628. doi: 10.1111/j.1365-2052.2007.01659.x

Glover, K. A., Hansen, M. M., Lien, S., Als, T. D., Høyheim, B., and Skaala, Ø. (2010). A comparison of SNP and STR loci for delineating population structure and performing individual genetic assignment. *BMC Genet.* 11:2. doi: 10.1186/1471-2156-11-2

Grasso, A. N., Goldberg, V., Navajas, E. A., Iriarte, W., Gimeno, D., Aguilar, I., et al. (2014). Genomic variation and population structure detected by single nucleotide polymorphism arrays in Corriedale, Merino and Creole sheep. *Genet. Mol. Biol.* 37, 389–395. doi: 10.1590/S1415-47572014000300011

Groeneveld, L., Lenstra, J., Eding, H., Toro, M., Scherf, B., Pilling, D., et al. (2010). Genetic diversity in farm animals–a review. *Anim. Genet.* 41, 6–31. doi: 10.1111/j.1365-2052.2010.02038.x

Haile, A., Tembely, S., Anindo, D., Mukasa-Mugerwa, E., Rege, J., Yami, A., et al. (2002). Effects of breed and dietary protein supplementation on the responses to gastrointestinal nematode infections in Ethiopian sheep. *Small Rumin. Res.* 44, 247–261. doi: 10.1016/S0921-4488(02)00080-9

Hanotte, O., Bradley, D. G., Ochieng, J. W., Verjee, Y., Hill, E. W., and Rege, J. E. (2002). African pastoralism: genetic imprints of origins and migrations. *Science* 296, 336–339. doi: 10.1126/science.1069878

Hart, D., and Clark, A. (1997). *Principles of Population Genetics*. Sunderland, MA: Sinauer Associates.

Heaton, M. P., Leymaster, K. A., Kalbfleisch, T. S., Kijas, J. W., Clarke, S. M., McEwan, J., et al. (2014). SNPs for parentage testing and traceability in globally diverse breeds of sheep. *PLOS ONE* 9:e94851. doi: 10.1371/journal.pone.0094851

Helen, N. (2015). *Phenotypic and Genetic Characterization of Indigenous Sheep Breeds of Eastern Ethiopia A*. Haramaya: Haramaya University.

Horsburgh, K. A., Prost, S., Gosling, A., Stanton, J.-A., Rand, C., and Matisoo-Smith, E. A. (2013). The genetic diversity of the Nguni breed of African cattle (*Bos* spp.): complete mitochondrial genomes of haplogroup T1. *PLOS ONE* 8:e71956. doi: 10.1371/journal.pone.0071956

Horsburgh, K. A., and Rhines, A. (2010). Genetic characterization of an archaeological sheep assemblage from South Africa's Western Cape. *J. Archaeol. Sci.* 37, 2906–2910. doi: 10.1016/j.jas.2010.06.035

Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030

Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Porto Neto, L. R., San Cristoba, M., et al. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLOS Biol.* 10:e1001258. doi: 10.1371/journal.pbio.1001258

Kijas, J. W., Townley, D., Dalrymple, B. P., Heaton, M. P., Maddox, J. F., McGrath, A., et al. (2009). A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. *PLOS ONE* 4:e4668. doi: 10.1371/journal.pone.0004668

Kovács, A., Kukovics, S., and Jávor, A. (2008). *Dorpers, the Meat Sheep of the Future*, Vol. 7. Oradea: Analele Universitatii din Oradea, 272–275.

Krings, M., Salem, A. E., Bauer, K., Geisert, H., Malek, A. K., Chaix, L., et al. (1999). mtDNA analysis of Nile River Valley populations: a genetic corridor or a barrier to migration? *Am. J. Hum. Genet.* 64, 1166–1176. doi: 10.1086/302314

Lachance, J., and Tishkoff, S. A. (2013). SNP ascertainment bias in population genetic analyses: why it is important, and how to correct it. *Bioessays* 35, 780–786. doi: 10.1002/bies.201300014

Leta, S., and Mesele, F. (2014). Spatial analysis of cattle and shoat population in Ethiopia: growth trend, distribution and market access. *Springerplus* 3:310. doi: 10.1186/2193-1801-3-310

Litt, M., and Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* 44, 397–401.

Madrigal, L., Ware, B., Miller, R., Saenz, G., Chavez, M., and Dykes, D. (2001). Ethnicity, gene flow, and population subdivision in Limon, Costa Rica. *Am. J. Phys. Anthropol.* 114, 99–108. doi: 10.1002/1096-8644(200102)114:2<99::AID-AJPA1010>3.0.CO;2-V

McTavish, E. J., and Hillis, D. M. (2015). How do SNP ascertainment schemes and population demographics affect inferences about population history? *BMC Genomics* 16:266. doi: 10.1186/s12864-015-1469-5

McVean, G. (2009). A genealogical interpretation of principal components analysis. *PLOS Genet.* 5:e1000686. doi: 10.1371/journal.pgen.1000686

Mekuriaw, G. (2016). *Molecular Characterization of Ethiopian Indigenous Goat Populations: Genetic Diversity and Structure, Demographic Dynamics and Assessment of the Kisspeptin Gene Polymorphism*. Addis Ababa: Addis Ababa University.

Melesse, A., Banerjee, S., Lakew, A., Mersha, F., Hailemariam, F., Tsegaye, S., et al. (2013). Morphological characterization of indigenous sheep in Southern Regional State, Ethiopia. *Anim. Genet. Resour.* 2, 39–50. doi: 10.1017/S2078633612000513

Muigai, A. (2003a). *Characterization and Conservation of Indigenous Animal Genetic Resources: The Fat-tailed and Thin-tailed Sheep of Africa*. Ph.D. thesis, Jomo Kenyatta University of Agriculture and Technology, Juja.

Muigai, A. (2003b). *Characterization and Conservation of Indigenous Animal Genetic Resources: Genetic Diversity and Relationships of Fat-tailed and Thin-tailed Sheep of Africa*. Juja: Jomo Kenyatta University of Agriculture and Technology.

Muigai, A. W. T., and Hanotte, O. (2013). The origin of African sheep: archaeological and genetic perspectives. *Afr. Archaeol. Rev.* 30, 39–50. doi: 10.1007/s10437-013-9129-0

Negrini, R., Nicoloso, L., Crepaldi, P., Milanesi, E., Colli, L., Chegdani, F., et al. (2009). Assessing SNP markers for assigning individuals to cattle populations. *Anim. Genet.* 40, 18–26. doi: 10.1111/j.1365-2052.2008.01800.x

Peter, C., Bruford, M., Perez, T., Dalamitra, S., Hewitt, G., and Erhardt, G. (2007). Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds. *Anim. Genet.* 38, 37–44. doi: 10.1111/j.1365-2052.2007.01561.x

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Qwabe, S. O., van Marle-Köster, E., and Visser, C. (2013). Genetic diversity and population structure of the endangered Namaqua Afrikaner sheep. *Trop. Anim. Health Prod.* 45, 511–516. doi: 10.3389/fgene.2014.00333

Ramos, A. M., Megens, H. J., Crooijmans, R. P., Schook, L. B., and Groenen, M. A. (2011). Identification of high utility SNPs for population assignment and traceability purposes in the pig using high-throughput sequencing. *Anim. Genet.* 42, 613–620. doi: 10.1111/j.1365-2052.2011.02198.x

Reynolds, J., Weir, B. S., and Cockerham, C. C. (1983). Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* 105, 767–779.

Ripoli, M. V., Wei, S., Rogberg Muñoz, A., Guo, B., Goszczynski, D. E., Fernandez, M. E., et al. (2013). Evaluation of six single nucleotide polymorphisms for bovine traceability in the context of the argentine-chinese beef trade. *BAG J. Basic Appl. Genet.* 24, 31–45.

Robertson, A., and Hill, W. G. (1984). Deviations from Hardy-Weinberg proportions: sampling variances and use in estimation of inbreeding coefficients. *Genetics* 107, 703–718.

Rosenblum, E. B., and Novembre, J. (2007). Ascertainment bias in spatially structured populations: a case study in the eastern fence lizard. *J. Hered.* 98, 331–336. doi: 10.1093/jhered/esm031

Ryder, M. L. (1984). "Sheep," in *Evolution of Domestic Animals*, ed. I. L. Mason (London: Longman), 63–85.

Sassi-Zaidy, Y. B., Maretto, F., Charfi-Cheikrouha, F., and Cassandro, M. (2014). Genetic diversity, structure, and breed relationships in Tunisian sheep. *Small Rumin. Res.* 119, 52–56. doi: 10.1016/j.smallrumres.2014.02.005

Seo, S. N. (2008). *Differential Adaptation Strategies by Agro-ecological Zones in African Livestock Management*. Washington, DC: World Bank Publications.

Seo, S. N., McCarl, B. A., and Mendelsohn, R. (2010). From beef cattle to sheep under global warming? An analysis of adaptation by livestock species choice in South America. *Ecol. Econ.* 69, 2486–2494. doi: 10.1016/j.ecolecon.2010.07.025

Spangler, G. L., Rosen, B. D., Ilori, M. B., Hanotte, O., Kim, E.-S., Sonstegard, T. S., et al. (2017). Whole genome structural analysis of caribbean hair sheep reveals quantitative link to west african ancestry. *PLOS ONE* 12:e0179021. doi: 10.1371/journal.pone.0179021

Tibbo, M. (2006). *Productivity and Health of Indigenous Sheep Breeds and Crossbreds in the Central Ethiopian Highlands*. Doctoral thesis, Swedish University of Agricultural Sciences, Uppsala.

Verbeek, E., Kanis, E., Baker, R., and Kosgey, I. (2007). Socio-economic factors influencing small ruminant breeding in Kenya. *Livestock Res. Rural Dev.* 19, 1–11.

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

Wilson, R. (1991). *Small Ruminant Production and the Small Ruminant Genetic Resource in Tropical Africa (Animal Production and Health Paper No. 88)*. Rome: Food and Agriculture Organization.

Wilson, R. T. (2011). Populations and production of fat-tailed and fat-rumped sheep in the Horn of Africa. *Trop. Anim. Health Prod.* 43, 1419–1425. doi: 10.1007/s11250-011-9870-9

Yang, J., Li, W.-R., Lv, F.-H., He, S.-G., Tian, S.-L., Peng, W.-F., et al. (2016). Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Mol. Biol. Evol.* 33, 2576–2592. doi: 10.1093/molbev/msw129

Zhao, Y. X., Yang, J., Lv, F. H., Hu, X. J., Xie, X. L., Zhang, M., et al. (2017). Genomic reconstruction of the history of native sheep reveals the peopling patterns of nomads and the expansion of early pastoralism in East Asia. *Mol. Biol. Evol* 34, 2380–2395. doi: 10.1093/molbev/msx181