



Genetic diversity and population structure among indigenous and imported goat breeds in Kenya

Ruth W Waaineina ^{*,a,b}, Kiplangat Ngeno ^a, Tobias O Okeno ^a and Evans D Ilatsia ^b

^a Department of Animal Sciences, Animal Breeding and Genomics Group, Egerton University, P.O. Box 536-20115, Egerton, Kenya

^b Dairy Research Institute, Agricultural and Livestock Organization, P.O. Box 25-20117, Naivasha, Kenya

Abstract: Population structure and relationship information among goats is critical for genetic improvement, utilization, and conservation. This study explored population structure and level of introgression among four goat breeds in Kenya: the indigenous Galla (n = 12) and three imported breeds, the Alpine (n = 29), Toggenburg (n = 31), and Saanen (n = 24). Genetic diversity was analyzed using four indices (polymorphic SNPs, mean allele frequency, observed and expected heterozygosity and inbreeding coefficient) within each breed. Population structure assessed using model-based clustering (ADMIXTURE) revealed four breeds according to their geographic regions in Kenya. Kenyan Alpine goats were the most admixed breed with about 10 % of its genome derived from Galla, 10 % and 6 % from Saanen and Toggenburg respectively. The association of Galla with other breeds was anticipated since the Galla breed was used as the founder population for crossbreeding with Saanen, Alpine and Toggenburg breeds. The relationship information evaluated by computing Reynolds genetic distance revealed five distinctive clusters: Alpine, Galla, Saanen, Toggenburg and some mixture of Alpine and Toggenburg. Saanen and Galla breeds seem to be the most genetically distinct among the sampled populations. The genetic variation among the goat populations observed will provide a good opportunity for sustainable utilization, conservation, and future genetic resource improvement programmes in goat breeds in Kenya.

Keywords: Admixed breeds, Breed relationship, Crossbred goats, Gene introgression

Citation: Waaineina, R. W., Ngeno, K., Okeno, T. O., Ilatsia, E. D. (2021). Genetic diversity and population structure among indigenous and imported goat breeds in Kenya. *Genetic Resources* 2 (3), 25–35. doi: [10.46265/genresj.EQFQ1540](https://doi.org/10.46265/genresj.EQFQ1540).

© Copyright 2021 the Authors.

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

Goats are known to be the most adaptable and widespread species of domestic animals, thriving across various geographical conditions, ranging from mountains to deserts and the tropics, Africa included. The importance of goats in supporting rural household economies in developing countries is well documented (Deshingkar *et al*, 2008; Herrero *et al*, 2013). They are an important source of food and nutritional security through the supply of milk and meat, income generation through sale of surplus stock, and insurance against unforeseen risks in addition to having important, non-tangible cultural values (Herrero *et al*, 2013; Mbuku

et al, 2015; Ogola *et al*, 2010). Recent studies have shown that goat farming is one of alternative climate-smart agricultural practices that could build farmers resilience to climate change-related challenges (Ojango *et al*, 2016). The diminishing land sizes in the medium to high potential areas for agriculture due to human population pressure, expansion of urban areas and climate change-related challenges, call for alternative farming practices such as intensive dairy goat production, which offers more multi-functionality, flexibility, and adaptability to varied production conditions (Scarpa *et al*, 2003).

In Kenya, dairy goat production has mainly been supported by imported breeds such as Toggenburg, Anglo-Nubian, German Alpines, Saanen and Boer, and crossbreeds between imported and selected local breeds such as Galla and small East African goat (Ahuya *et al*, 2006; Bett *et al*, 2007; Krause, 2006). Galla

*Corresponding author: Ruth W Waaineina
(waineinaruth@gmail.com)

in Kenya is also referred to as Boran/Somali goat. They are indigenous in arid and semi-arid northern Kenya, and pure Galla are maintained in various Government breeding and conservation stations in the Country. Their characteristics, such as resistance to dehydration, preference for browsing and a wide range of feeding habits (Chenyambuga *et al*, 2004) have allowed them to adapt to the massive arid and semi-arid regions in the country, and could potentially be advantageous traits for goat breeding programmes. The imported breeds were introduced to various parts of the country by the Government of Kenya and Non-Governmental Organizations, with the aim of increasing goat productivity through appropriate husbandry and disease interventions (Peacock, 2005) and targeted breeding strategies such as crossbreeding (Bett, 2009; Peacock *et al*, 2011). The crossbreeds were kept in different geographical locations as isolated populations and subjected to separate breeding objectives for several decades.

Crossbreeding has been the strategy of choice to improve the productivity of goats under various production systems (Ahuya *et al*, 2009; Peacock *et al*, 2011). This has resulted in an increase in population sizes of crossbred goats especially in the areas the breeds were introduced (Mburu *et al*, 2014; Peacock *et al*, 2011). However, increase in population sizes did not necessarily correspond to enhanced productivity of the goats but rather reflected large numbers of households striving to support their livelihoods through goat farming (Aziz, 2010; Bett *et al*, 2011; Mburu *et al*, 2014).

In Kenya, there has been limited technical capacity on the farmers' side on how to manage the rather complex crossbreeding programmes, a fact that may have had a bearing on the sustainability of such initiatives in the long term (Aziz, 2010; Bett *et al*, 2011; Mburu *et al*, 2014). The net result of this has been the unsystematic crossing of the existing population, poor flock management, lack of records to support decision making and general lack of simplified breeding programmes to guide in genetic improvement of goats in the country (Kosgey and Okeyo, 2007). Currently, crosses of imported and local goats are reared as dairy goats in different parts of the country under different production systems. There is a huge source of genetic diversity in the current goat populations in Kenya. This is a result of unsystematic crossbreeding and lack of record keeping by most of the smallholder farmers. This calls for the need to characterize, conserve and sustainably utilize goats under various production systems in Kenya. It is important to determine genetic diversity in populations because it provides the basis for natural and artificial selection (Qanbari and Simianer, 2014).

To measure and describe genetic diversity of animal genetic resources, phenotypic and molecular characterization tools are used as a starting point to understand the animal resources and make use of them sustainably

(FAO, 2011). Characterization starts with the gathering of all information on breed origin, development, structure, population, quantitative and qualitative characteristics in defined management and climatic conditions (Gizaw *et al*, 2011; Rege and Okeyo, 2011).

Molecular characterization, using genetic markers to detect polymorphisms in nuclear DNA, is a powerful tool which can be applied in breeding programmes. For instance, it can be used to characterize the genetic variability within, and genetic distance between, populations, as well as for genomic selection, parentage verification and genetic diversity preservation (Groeneveld *et al*, 2010).

Microsatellite markers and single-nucleotide polymorphisms (SNP) are the most commonly used markers in animal breeding related fields (FAO, 2011). Microsatellite markers have several limitations, for example in the detection of null alleles (Hoshino *et al*, 2012) and homoplasy (Jarne and Lagoda, 1996; Anmarkrud *et al*, 2008), while SNPs have several advantages over microsatellites, including being highly reproducible and informative (Vignal *et al*, 2002) and the fact that SNPs can represent either neutral or functional genetic diversity (Kohn *et al*, 2006). A SNP microarray with more than 50,000 SNPs (GoatSNP50 Bead Chip, Illumina, Inc. San Diego, CA 92122 USA), which was developed from SNP loci detected by whole-genome sequencing of six different goat breeds according to Tosser-Klopp *et al* (2014) is available. This has made SNP markers the most popular and advanced technology in molecular breed characterization in goats. Additionally, its robustness, low genotyping costs, automatic allele calling and capability to interrogate the goat genome at high resolution (Ajmone-Marsan *et al*, 2014) demonstrate practicality in implementing genomic characterization in goats.

There has been no deliberate effort to understand the genetic diversity and population admixture among the goat breed populations in Kenya by use of SNP makers. This study, therefore, investigated genetic diversity, population structure and admixture among goat breeds in Kenya. The results from this study will facilitate management efforts in conserving and utilizing the various goat genetic resources sustainably.

Materials and Methods

Study area

Blood samples were collected from goats in three counties in Kenya: Nyeri (Mukurweini Sub-County), Meru (Central Imenti Sub-County) and Homa Bay (Homa Bay Town Sub-County) located in the Central, Eastern and Western regions of Kenya, respectively (Figure 1). These areas were selected because they represent the entry points of different imported dairy goat breeds in Kenya. Mukurweini Sub-County lies in the Upper midlands, also known as the main coffee zone, at an altitude of 1460-1710 metres above sea level (masl) and receives 950-1500 mm of mean annual rainfall.

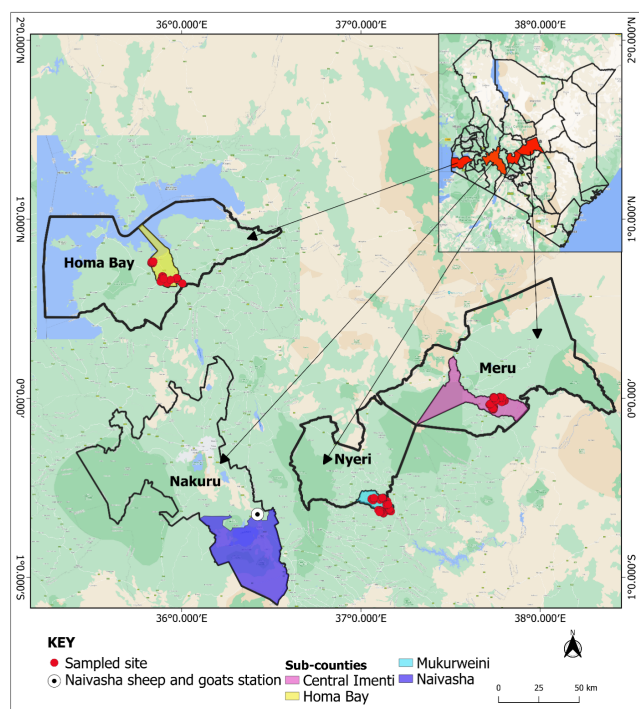


Figure 1. Map of Kenya showing area sampled within the selected sub-counties

Central Imenti is in the upper highlands, at an altitude ranging between 1830-2210 masl and has an average annual precipitation of 800-2600 mm. The Homa Bay Town Sub-County lies in the lower midlands, at 1166 masl and receives an annual rainfall of 1226 mm.

Animal resources and sampling

Goat keeping households were purposively selected based on the following criteria: 1) being a member of dairy goat farmer group and 2) having more than two mature does which matched the breed kept by the farmer group in the said county of study. The herd structures between the breeds and within the county/breed varied among the selected households. Therefore, when a farm had only two mature does, only one doe was sampled. Where more than two does were available, the relationship of the does was confirmed by the farmer to avoid sampling closely related does. Sampling of full and half siblings was avoided. To ensure the representativeness of sampling for each breed, unrelated animals were selected from various farms across the designated counties. The Galla goat breed, however, did not follow the criteria because they were from the government breeding station where breeding records were used to identify the animal to be sampled. Therefore, to minimize sampling from closely related animals within the Galla population, pedigree data were used to select against full and half sibling animals.

A total of 96 animals including three imported breeds (31 Toggenburg, 29 Alpine, 24 Saanen) and one indigenous breed (12 Galla) were incorporated in this study. The Toggenburg and Alpine were found in

Eastern and Central Kenya under Meru Goat Breeders Association (MGBA) and Dairy Goat Association of Kenya (DGAK), respectively. Saanen goats were found in Homa Bay under Nyanza Goat Breeders Association (NGBA). All samples were collected from a total of 53 farms in the three counties, Nyeri (18), Meru (19) and Homa Bay (16). Galla goats were sampled from the sheep and goats government station in Naivasha.

Whole blood (10 ml) was collected from the jugular vein into Vacutainer tubes with Ethylenediaminetetraacetic acid (EDTA) as an anticoagulant. The blood was stored at -20°C for two months before genomic DNA extraction. The procedure of blood sampling followed FAO guidelines (FAO, 2012). From each animal, duplicate samples were collected and kept separately during transportation and storage. For each sample, the following information was collected: sex of the animal, basic pedigree information, size of the flock, breed, any relevant phenotypic feature, and a photograph of the goat.

The study was conducted in strict accordance with the recommendations of the Institute of Primate Research (IPR) Ethical Guidelines on Animal Care and Use of Laboratory Animals (<https://grants.nih.gov/grants/olaw/guide-for-the-care-and-use-of-laboratory-animals.pdf>).

The protocol was approved by the committee on the ethics of Animal Experiment of Egerton University of Egerton in Kenya (ISERC/03/2020). A qualified veterinary officer collected the whole blood following FAO guidelines (FAO, 2012) to reduce pain and discomfort to a minimum.

DNA extraction and genotyping

Genomic DNA was extracted from the whole blood using the DNeasy Blood and Tissue Kit (Qiagen[®], Hilden, Germany) according to the manufacturer's instructions. Ten μl of each ten randomly selected samples were subjected to a preliminary estimate of the DNA quality and quantity on a 1 % agarose gel electrophoresis. Secondary quantification and purity analysis of the DNA were confirmed using one μl for each sample on both Nanodrop Spectrophotometer (Nanodrop[®] ND-2000) and Qubit[®] dsDNA BR (Broad-Range) Assay Kit on the Qubit 3.0 fluorimeter (Invitrogen). The extraction and quality control check of genomic DNA was done at Kenya Agricultural and Livestock Research Organization biotechnology laboratory in Kabete, Kenya.

The DNA samples were genotyped using the Goat-SNP50 Bead Chip (Illumina, Inc. San Diego, CA 92122 USA), developed by the International Goat Genome Consortium (IGGC), which features 53347 SNPs across the whole genome with inter-SNP spacing of approximately 40 kb (Tosser-Klopp *et al*, 2014). The genotyping was outsourced to Neogen Europe Limited in Scotland (<https://genomics.neogen.com/en/>).

SNP quality control and data analysis

The SNP genotype quality control process was applied to raw reads for both merged (all breeds) and then per breed using PLINK v1.07 (Purcell *et al*, 2007).

Table 1. Goat breed, number of goats and SNPs excluded and remaining after quality control processes on genotyping data. N, number of animals; MIND, genotype missing (< 0.1), GENO, SNP missing (< 0.15), MAF, minor allele frequency (< 0.05); HWE, Hardy-Weinberg equilibrium (P-value < 0.001).

Breed	N	Excluded SNPs					Remaining SNPs	Remaining samples
		MIND	GEN	HWE	MAF	TOTAL		
Saanen	24	0	2453	38	3658	6149	47198	24
Alpine	29	1	2496	47	2586	5129	48218	28
Galla	12	0	2413	26	8249	10688	42659	12
Toggenburg	31	1	2345	50	4690	7085	46262	30
Merged	96	2	2235	663	644	3542	49805	94

First, individuals with a missing genotype call rate of more than or equal to 10 % were removed from further analysis using the mind function in PLINK with default settings. The remaining individuals were then exposed to further filtering. SNPs with less than 95 % call rate, Minor Allele Frequency (MAF) of less than 0.05 and $P < 0.001$ Hardy Weinberg Equilibrium (HWE) were excluded from downstream analyses. The SNP data set used for downstream analysis is accessible from the Mendely Digital Repository (<https://doi.org/10.17632/hhb9rhdztt.1>).

Basic genetic diversity indices, which include the proportion of polymorphic markers, inbreeding coefficient, observed (H_o) and expected (H_e) heterozygosity were calculated within breeds using PLINK (Purcell et al, 2007). The proportion of polymorphic SNPs (P_N) offers the fraction of the total SNPs that showed both alleles within each population. The P_N was calculated as the proportion of SNPs with more than 1 % MAF within each breed. The MAF is the approximate frequency of the second most common allele per breed. The output from PLINK for observed and expected heterozygosity per animal within breed was subjected to further calculation to get an average estimate of H_o and H_e per breed. The heterozygosity values were calculated by getting the average of all SNPs (that is the sum of all heterozygosity values averaged over the total number of SNPs passed the quality control).

The population structure and relatedness were estimated by principal components analysis (PCA) using the R package SNPRELATE (Zheng et al, 2012) and admixture proportion inference using model-based clustering ADMIXTURE 1.3.0 software (Alexander et al, 2009). The PCA analysis allowed for visual investigation and solid quantitative summaries. The admixture analysis inferred the proportions of ancestry within the populations by use of prior defined K-values matching the assumed number of ancestral populations. The admixture procedure employs a maximum-likelihood based method by converging the ancestry proportions and allele frequencies that maximize the likelihood function. The most optimal population structure was determined by cross-validation error procedure (McVean, 2009) with assumed ADMIXTURE runs from $K = 2$ to $K = 4$. The K-value with the lowest CV error was selected as the optimal value. A phylogenetic tree based on Reynolds

genetic distances representing relationships among goat breeds was visualized using iTOL software (Letunic and Bork, 2019).

Results

Quality control procedure on the 53,347 SNPs included on the SNP chip excluded a total of 3,542 SNPs retaining 49,805 SNPs for downstream analyses as shown in Table 1. Of the excluded SNPs, 2,235 had less than 0.1 missing per SNP, 663 SNPs significantly deviated from HWE ($P < 0.001$) and 644 SNPs had MAFs lower than 0.05. The Galla breed had the highest number of SNPs excluded in total (10,688), whereas Alpine revealed the lowest number of SNPs excluded (5,129). It is worth noting that some SNPs were left out due to more than a single criterion.

Genetic diversity

The four indices of genetic diversity (polymorphic SNPs, mean allele frequency, observed and expected heterozygosity and inbreeding coefficient) were calculated within each breed (Table 2). The assessment of the proportion of SNPs that exhibited both alleles within each breed indicated high levels of diversity. The percentage of within-breed polymorphic SNPs ranged from 94.6% to 80.7%. The highest values of polymorphic loci were found in Alpine (94.6%) and Saanen (92.2%) while the lowest proportion was found in the Galla breed (80.7%). Across all loci, the lowest MAF was found in Galla (0.291) and the highest in Alpine (0.323).

Results revealed differences in genetic diversity between breeds. The expected heterozygosity was, in all cases, higher than the observed heterozygosity ($H_e > H_o$). The Alpine had the lowest observed heterozygosity ($H_o = 0.558 \pm 0.026$) while Toggenburg had the highest ($H_o = 0.580 \pm 0.032$). Inbreeding coefficients for all the breeds were negative and ranged between -0.013 (Toggenburg) and -0.042 (Galla).

Population structure analysis

Principal components analysis was used to cluster goats and explore the association among individuals and breed groups. In Figure 2, the principal component 1, which accounts for 15.2% of the total variance, separated Galla breed from the other three breeds. The

Table 2. Proportion of polymorphic SNPs(P_N), mean allele frequency (MAF), observed (H_o) and expected (H_e) heterozygosity and inbreeding coefficient (F) for the goat breeds. N, number of animals.

Breed	N	P_N [%]	MAF	$H_o \pm SD$	$H_e \pm SD$	F
Toggenburg	30	89.9	0.297	0.580±0.032	0.580±0.001	-0.013
Alpine	28	94.6	0.323	0.558±0.026	0.564±0.001	-0.015
Galla	12	80.7	0.291	0.563±0.025	0.580±0.000	-0.042
Saanen	24	92.2	0.311	0.559±0.019	0.573±0.001	-0.034

principal component 2 accounts for 14.1% of the total variance, split the goat breeds into four clusters (Alpine, Saanen, Galla and Toggenburg clusters). One outlier, was, however, observed for the Saanen population mixing with Alpine population.

To examine admixture between the breeds, model-based clustering was performed and the most likely number of genetic population (cluster or K) between the four goat breeds were deduced using ADMIXTURE cross-validation procedure (McVean, 2009). The K-value with the lowest CV error was $K = 4$ and was selected as an optimal number of ancestral populations (Figure 3).

A population structure plot (Figure 4) showed proportions of ancestral populations for all breeds (Alpine, Galla, Saanen and Toggenburg) for $K = 2$ to $K = 4$. At $K = 2$, Galla goats were separated from the other three goat breeds (Toggenburg, Saanen and Alpine). Moreover, Galla goats largely do not carry ancestral components present in Saanen, Alpine and Toggenburg goats (shown in light blue, Figure 4). At $K = 3$, Alpine and Saanen goats carry ancestral components largely absent from either the Galla or Toggenburg goats. At $K = 4$, Galla had the lowest level of admixture, whereas Toggenburg, Alpine and Saanen demonstrated some signs of admixture with Galla.

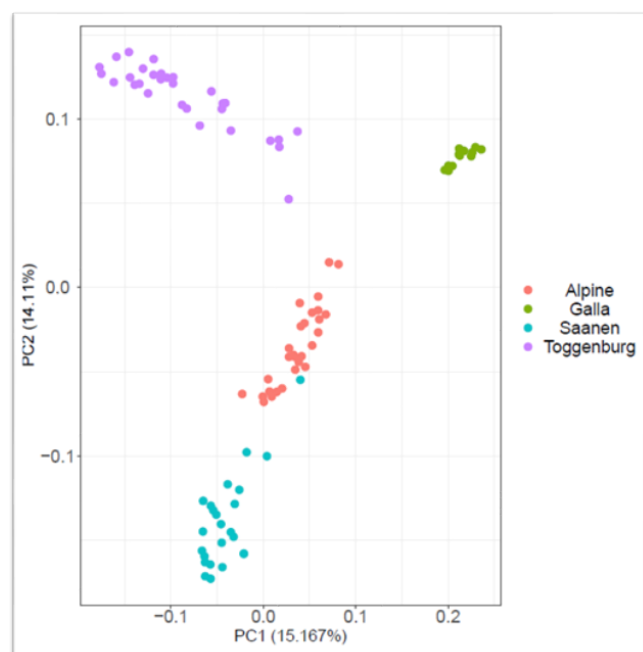


Figure 2. Principal components analysis plot based on SNP array data of goat breeds

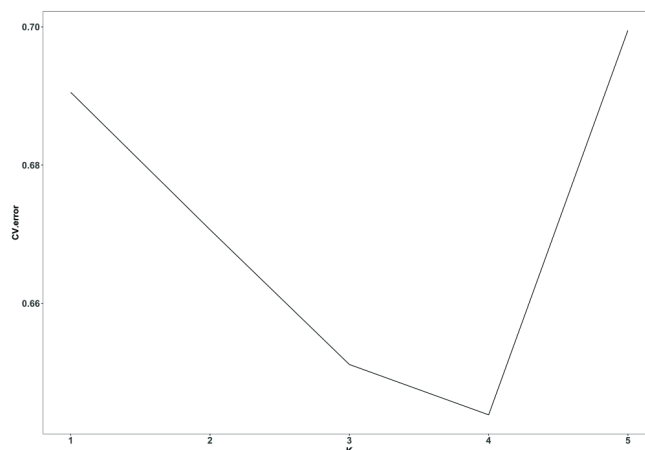


Figure 3. A cross-validation plot indicating the choice of optimal K-value

The proportions of individuals in each breed in the four most likely clusters estimated by ADMIXTURE is shown in Table 3. 71 % of Alpine breed were allocated to cluster one, 97 % of Galla were assigned to cluster two with one percent (1 %) of its genome assigned to cluster one, three and four, 84 % percent of Saanen were in cluster three with seven percent (7 %) of its genome assigned to cluster one. On the other hand, 78 % of Toggenburg were assigned to cluster four with 17 %, and three percent (3 %) of its genome allocated to cluster two and one respectively.

Breed relationships were evaluated by computing the genetic distance between all pairwise combinations of individuals (D) from the average proportion of shared alleles. Based on the calculated Reynolds genetic distances, a phylogenetic tree was constructed to represent breed clustering (Figure 5). The results revealed five clusters for the four populations (Alpine, Saanen, Toggenburg and Galla). Some Toggenburg goats were found to be grouped together with Alpine, forming the fifth cluster.

Discussion

Genetic diversity

Livestock has been exposed to various forces that contributed to the genetic diversity underlying phenotypic dissimilarities ever since domestication. These forces include natural selection, artificial selection for specific traits, migration, genetic drift and inbreeding (Andersson and Georges, 2004; Groeneveld *et al*, 2010). Genetic drift plays an important role during short-term evolu-

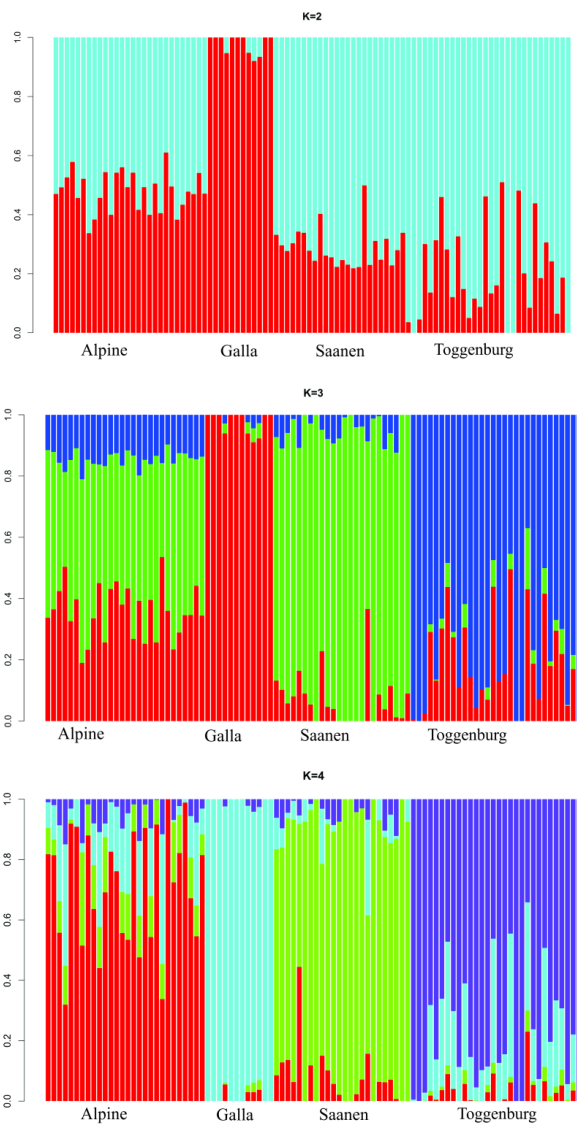


Figure 4. Population structure plots showing proportions of ancestral populations for individuals of sampled goat breeds (Alpine, Galla, Saanen and Toggenburg) for K= 2 to K = 4

tion in situations where populations are reproductively isolated (Laval et al, 2002).

Genotyping with the GoatSNP50 Bead Chip revealed some levels of diversity within the goat breeds in this study. In each breed, fewer than 80% of SNPs exhibited polymorphisms, and heterozygosity ranged from 0.558 to 0.580 (Table 2). A large number of polymorphic SNPs were detected for Alpine (94.6) and Saanen (92.2)

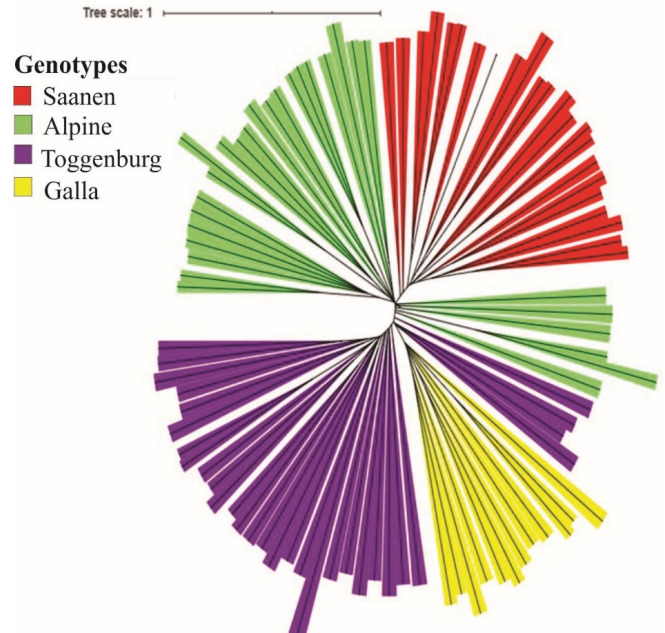


Figure 5. Phylogenetic tree based on Reynolds genetic distances representing breed relationships among goat breeds.

breeds; this was expected because sequenced data from Alpine and Saanen were included in the 50K SNPs panel discovery (Tosser-Klopp et al, 2014). Other results from various authors using different numbers of samples and goat breeds showed >93 % of polymorphism (Visser et al, 2016; Onzima et al, 2018). However, it is difficult to compare and conclude the estimates of SNPs stated as polymorphic by other authors, because the number of samples genotyped per breed and proportion of genotyped samples used for SNP discovery varied.

The diversity amongst the four breeds showed Galla had the lowest diversity among sampled individuals in comparison with the other breeds. The polymorphic variance is dependent on the history of each breed. As opposed to Galla goats, the three introduced breeds showed greater diversity, likely as a result of crossbreeding with local goats (Galla and small East African). Therefore, each breed may contain genetic contributions from various breeds, thus revealing higher polymorphisms than Galla goats. In contrast, Galla goats sampled from the government breeding station with detailed pedigree records still maintained levels of diversity ($P_N = 81\%$). The Galla population had been introduced in the Naivasha Sheep and Goat Station in the early 1970s during a sheep and goat

Table 3. Proportion of individuals of goat breeds in each of the four clusters estimated by ADMIXTURE. The diagonal indicates the inferred cluster. N, number of animals sampled

Predefined populations	Inferred clusters N	1	2	3	4
Alpine	28	0.708±0.197	0.136±0.128	0.101±0.071	0.055±0.048
Galla	12	0.013±0.020	0.970±0.044	0.008±0.013	0.009±0.015
Saanen	24	0.073±0.095	0.049±0.073	0.838±0.149	0.040±0.039
Toggenburg	30	0.030±0.047	0.173±0.142	0.015±0.019	0.782±0.193

project funded by FAO (Palian and Racokzi, 1976). The population is registered with the Galla Goat Breeders Society of Kenya (GGBSK) and the Kenya Stud Book. The population is inspected every year by inspectors from GGBSK using the Galla goat breed standards. Therefore, the station has maintained pure breed Galla goats which produce meat, milk and reproduce under harsh conditions while maintaining or conforming to the set standards of excellence defined by the GGBSK, where the objective is genetic improvement of target traits while controlling the level of inbreeding. The negative inbreeding coefficient in this study can effectively be taken as zero values, which means that there is no inbreeding observed in the reference populations. It could also mean that many heterozygotes were observed although the sample size for the four breeds was small. The increased heterozygosity could be due to random mating within the herd rather than random differences between herds.

The observed heterozygosity was lower than the expected heterozygosity ($H_o < H_e$) in the three breeds apart from Toggenburg, which recorded the same value for both observed and expected heterozygosity. The difference between the observed and expected heterozygosity was small, which may not be due to inbreeding but a Wahlund effect (Garnier-Géré and Chikhi, 2013). The observed heterozygosity in the current study was from a sample of individuals from a structured population even though all sub-divisions were in Hardy-Weinberg equilibrium. Over and above the semi and intensive production systems practised by smallholders, there is the presence of artificial selection, gene flow and non-random mating, hence not holding the law of HWE in these populations. In this study the observed and expected heterozygosity for Alpine ($H_o = 0.558$; $H_e = 0.564$), Saanen ($H_o = 0.559$; $H_e = 0.573$) and Toggenburg ($H_o = 0.580$; $H_e = 0.580$) were higher than those stated in Canada for Alpine ($H_o = 0.385$; $H_e = 0.388$), Saanen ($H_o = 0.379$; $H_e = 0.382$) and Toggenburg ($H_o = 0.353$; $H_e = 0.336$) (Brito *et al*, 2017). Moreover, Saanen in Italy recorded the same trend as in Canada ($H_o = 0.41$; $H_e = 0.41$) (Nicoloso *et al*, 2015). Differences in effective population sizes, length of isolation, selection and breeding management practices in the various production system may be the cause of these variances.

Toggenburg and Galla breeds had the highest expected heterozygosity. This could be explained by the types of crossbreeding programmes practised by farmers keeping Toggenburg breeds resulting in an admixed population. Organized breeding strategies using artificial selection are practised for Galla goats under the government breeding station resulting in genetic variability and lack of inbreeding for the populations.

Population structure and relationship

Principal component and population structure analyses confirmed distinctiveness among the goat breeds (Saa-

nen, Galla, Toggenburg and Alpine) according to their geographic regions in Kenya. This can be explained by the demographic history of these breeds that have been reared for a long time in separate geographic locations (Ahuya *et al*, 2009; Peacock *et al*, 2011). Although goats from each breed clustered separately, model-based clustering revealed some signs of admixture and genetic links between Alpine, Toggenburg, Saanen and Galla.

The results (Figure 4 and Table 3) of this study indicate that Kenyan Alpine goats were the most admixed breed with about 14 % of its genome derived from Galla, while ten and six percent of its genome is resulting from Saanen and Toggenburg respectively. It is worth noting that Saanen were introduced in the sampling region (Nyeri County) already before the Alpine were imported in the late 1970s. Therefore, the 10 % of Saanen genes in the Alpine genome may be a result of Saanen being one of the Kenyan Alpine ancestors. According to Waaineina *et al* (2019), lack of breeding stock was one of the challenges Alpine farmers were encountering, thus driving them to source breeding animals from local markets, friends, neighbours and commercial farms notwithstanding their undefined genetic composition. Furthermore, the increase in demand for dairy goats in the country has resulted in several farms setting up nucleus flocks with a significant proportion of the crossbred flocks as a source of breeding material for distribution to lower cadre farmers (Ahuya *et al*, 2006; Bett, 2009; Ogola *et al*, 2010). Through such arrangements, most of the breed-types have migrated to other areas apart from their original entry in the country (Mburu *et al*, 2014; Peacock, 2007). Toggenburg and Alpine goats shared some linkage with Galla goats, 17 % and 14 %, respectively. This was expected because Galla goat was used as the founder population for crossbreeding with Alpine and Toggenburg breeds (Ahuya *et al*, 2009; Bett *et al*, 2011; Mburu *et al*, 2014; Peacock *et al*, 2011; Shivairo *et al*, 2013).

As expected, Galla was the least admixed breed, in agreement with the history of this breed as the first indigenous goat for which a breed society was formed in Kenya. Moreover, the particular population in this study has been managed in seclusion within the government farm, and only animals registered within the society are allowed into the population. The Galla breed displayed isolation by distance and seemed to be at equilibrium under dispersal and genetic drift. In comparison with the other breeds in this study, Galla arrived in their current locations long before these breeds were introduced in Kenya and that is why there has been sufficient time for isolation by distance to take effect and, that long distance gene dispersal is sufficiently common to prevent genetic divergence.

The phylogenetic analysis categorized the breeds into five clusters (Figure 5). The outcomes show a clear differentiation of Galla, Saanen, some Toggenburg and Alpine. A group of some Alpine and Toggenburg, however, remained clustered together, which may

be attributed to the adjacent regions of the breeds (Figure 1). Lack of differentiation in some of Alpine and Toggenburg breeds signified a high level of genetic resemblance and low divergence, which may be a result of gene flow among Alpine and Toggenburg breeds. Common ancestry, short domestication history, lack of selection pressure and movement of the goats may play a role in lack of differentiation in varied geographically separated populations. Furthermore, in Kenya, as well as other parts of Africa, goats are also used for religious and other cultural ceremonies such as payment of dowry and gifts (Herrero et al, 2013; Mbuku et al, 2015; Ogola et al, 2010). Therefore, some of the Alpine and Toggenburg breeds clustering together may be a result of movement of breed animals between the communities in those two regions due to the forementioned cultural ceremonies. As mentioned earlier, one of the criteria for selecting the goat keeping households in this study was them being members of a dairy goat farmer group association (DGAK, MGBA, NGBA). The associations are responsible for buck rotation among the group members, maintaining the purity of the breed and providing technical backstopping. However, the results indicate a need to technically strengthen the Dairy Goat Association of Kenya for Alpine and the Meru Goat Breeders Association for Toggenburg, because urgent management efforts are essential to improve on breeding aspects, utilization and conservation of the various goat genetic resources.

All Saanen goats formed one cluster in the phylogenetic analysis. Indeed, the long distance (over 450 km) between the regions where Saanen and the rest of the breeds are kept may be the barrier to gene flow from other breeds. Through adaptive hitchhiking, natural selection can play an essential role in shaping this variability (Andolfatto, 2001). Therefore, the observed genetic divergence of Saanen from Alpine, Toggenburg and Galla breeds could have been contributed by random genetic drift and natural selection for adaptation to their environment/region.

Genetic uniqueness can be determined from the magnitude of genetic distances and phylogenetic relationships between populations if supporting indications such as genetic history, records of production, reproduction and on adaptation are lacking (Eding and Laval, 1999; Tosser-Klopp et al, 2014; Zheng et al, 2012). Embracing this principle with respect to the results of this study, Saanen and Galla breeds seem to be the most genetically distinct among the populations sampled, and can be categorised as important genetic resources. It will be interesting to enlarge this breed level investigation in later studies through addition of all Kenyan goat breeds to better appreciate the genetic relationship among them.

Conclusion

The study revealed clear divergence between some goat breeds, which provides a wide prospect on the current genetic diversity of goats in Kenya. This will be vital in planning breeding strategies for genetic

resources that should be sustainably utilized and conserved. Of the breeds studied, Galla breed displayed isolation by distance and seemed to be at equilibrium under dispersal and genetic drift. This shows that stronger efforts of genetic conservation and sustainable management of its gene pool have been undertaken. However, further studies are required for the on-farm Galla population. The most admixed breeds were Alpine and Toggenburg. Therefore, there is need to technically strengthen the Dairy Goat Association of Kenya for urgent management efforts that are essential for genetic improvement, utilization and conservation of the various goat genetic resources. Additional studies on phenotypic similarities and performance evaluation of the breeds in this study could add value to the information generated from this study to form the basis for future genetic resource conservation and improvement of goat breeds in Kenya.

Acknowledgements

The authors wish to acknowledge the African Development Fund (AfDB), USAID and the Government of Kenya under the Kenya Climate Smart Agriculture Project (KCSAP CGS/CRGs-AD-2019) for their financial support to the first author. We sincerely thank Egerton University (Njoro, Kenya) for providing technical support to undertake the study, Kenya Agricultural and Livestock Research Organizations for granting the study leave for the first author, dairy goat farmers and Naivasha Sheep and Goat Breeding Station for making their animals available for sample collection.

Author contributions

RW, KN, TO and EI conceived the study, RW analyzed the data and drafted the manuscript. All authors read and approved the manuscript.

Conflict of interest statement

The authors declare that there is no conflict of interest with any organization concerning the material discussed in the manuscript.

References

- Ahuya, C. O., Ojango, J. M. K., Mosi, R. O., Peacock, C. P., and Okeyo, A. M. (2009). Performance of Toggenburg dairy goats in smallholder production systems of the eastern highlands of Kenya. *Small Ruminant Research* 83(1-3), 7–13. doi: <https://doi.org/10.1016/j.smallrumres.2008.11.012>
- Ahuya, C. O., Okeyo, A. M., and Murithi, F. M. (2006). Productivity of cross-bred goats under smallholder production systems in the Eastern highlands of Kenya. In *Small stock in development: Proceedings of a workshop on enhancing the contribution of small livestock to the livelihoods of resource-poor communities*, 7p.

- Ajmone-Marsan, P., Colli, L., Han, J. L., Achilli, A., Lancioni, H., Joost, S., Crepaldi, P., Pilla, F., Stella, A., Taberlet, P., Boettcher, P., Negrini, R., and Lenstra, J. A. (2014). The characterization of goat genetic diversity: Towards a genomic approach. *Small Ruminant Research* 121(1), 58–72. doi: <https://doi.org/10.1016/j.smallrumres.2014.06.010>
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19(9), 1655–1664. doi: <https://doi.org/10.1101/gr.094052.109>
- Andersson, L. and Georges, M. (2004). Domestic-animal genomics: deciphering the genetics of complex traits. *Nature Reviews Genetics* 5(3), 202–212. doi: <https://doi.org/10.1038/nrg1294>
- Andolfatto, P. (2001). Adaptive hitchhiking effects on genome variability. *Current Opinion in Genetics & Development* 11(6), 635–641. doi: [https://doi.org/10.1016/S0959-437X\(00\)00246-X](https://doi.org/10.1016/S0959-437X(00)00246-X)
- Anmarkrud, J. A., Kleven, O., Bachmann, L., and Lifjeld, J. T. (2008). Microsatellite evolution: Mutations, sequence variation, and homoplasmy in the hypervariable avian microsatellite locus HrU10. *BMC Evolutionary Biology* 8(1), 138–138. doi: <https://doi.org/10.1186/1471-2148-8-138>
- Aziz, M. A. (2010). Present status of the world goat populations and their productivity. *Lohmann Information* 45(2), 42–52. url: https://www.lohmann-information.com/content/li_45_artikel17.pdf.
- Bett, R. C. (2009). Design and evaluation of breeding strategies for low input dairy goat production systems in Kenya.
- Bett, R. C., Kosgey, I. S., Bebe, B. O., and Kahi, A. K. (2007). Genetic improvement of the Kenya Dual Purpose Goat: Influence of economic values and prospects for a practical breeding programme. *Tropical Science* 47(3), 105–119. doi: <https://doi.org/10.1002/ts.204>
- Bett, R. C., Kosgey, I. S., Kahi, A. K., and Peters, K. J. (2011). Definition of breeding objectives and optimum crossbreeding levels for goats in the smallholder production systems. *Small Ruminant Research* 96(1), 16–24. doi: <https://doi.org/10.1016/j.smallrumres.2010.11.008>
- Brito, L. F., Kijas, J. W., Ventura, R. V., Sargolzaei, M., Porto-Neto, L. R., Cánovas, A., Feng, Z., Jafarikia, M., and Schenkel, F. S. (2017). Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers. *BMC Genomics* 18(1), 229–229. doi: <https://doi.org/10.1186/s12864-017-3610-0>
- Chenyambuga, S. W., Hanotte, O., Hirbo, J., Watts, P. C., Kemp, S. J., Kifaro, G. C., Gwakisa, P. S., Petersen, P. H., and Rege, J. E. O. (2004). Genetic Characterization of Indigenous Goats of Sub-saharan Africa Using Microsatellite DNA Markers. *Asian-Australasian Journal of Animal Sciences* 17(4), 445–452. doi: <https://doi.org/10.5713/ajas.2004.445>
- Deshingkar, P., Farrington, J., Rao, L., Akter, S., Sharma, P., Freeman, H. A., and Reddy, J. (2008). Livestock and poverty reduction in India: findings from the ODI Livelihood Options Project. url: <https://hdl.handle.net/10568/281>.
- Eding, J. H. and Laval, G. (1999). Measuring genetic uniqueness in livestock. In *Genebanks and the conservation of farm animal genetic resources*, ed. Oldenbroek, J. K., (Lelystad, the Netherlands: DLO-Institute for Animal Science and Health (ID-DLO), Instituut voor Veehouderij en Diergezondheid, Research Branch Zeist), 33–58.
- FAO (2011). Molecular genetic characterization of animal genetic resources volume 9 of *FAO Animal Production and Health Guidelines*. (Rome) . url: <http://www.fao.org/3/i2413e/i2413e00.htm>.
- FAO (2012). Phenotypic characterization of animal genetic resources volume 11 of *FAO Animal Production and Health Guidelines*. (Rome) . url: <http://www.fao.org/3/i2686e/i2686e00.htm>.
- Garnier-Géré, P. and Chikhi, L. (2013). Population Subdivision, Hardy–Weinberg Equilibrium and the Wahlund Effect. In *eLS*, ed. John Wiley & Sons, L.
- Gizaw, S., Komen, H., Hanote, O., van Arendonk, J., Kemp, S., Haile, A., Okeyo, A. M., and Dessie, T. (2011). Characterization and conservation of indigenous sheep genetic resources: A practical framework for developing countries. In *ILRI Research Report*, volume 27, ILRI, Nairobi, Kenya.
- Groeneveld, L. F., Lenstra, J. A., Eding, H., Toro, M. A., Scherf, B., Pilling, D., Negrini, R., Finlay, E. K., Jianlin, H., Groeneveld, E., and and, S. W. (2010). Genetic diversity in farm animals - a review. *Animal Genetics* 41, 6–31. doi: <https://doi.org/10.1111/j.1365-2052.2010.02038.x>
- Herrero, M., Grace, D., Njuki, J., Johnson, N., Enahoro, D., Silvestri, S., and Rufino, M. C. (2013). The roles of livestock in developing countries. *Animal* 7, 3–18. doi: <https://doi.org/10.1017/s1751731112001954>
- Hoshino, A. A., Bravo, J. P., Nobile, P. M., and Morelli, K. A. (2012). Microsatellites as tools for genetic diversity analysis. In *Genetic Diversity in Microorganisms*, ed. Caliskan, M.
- Jarne, P. and Lagoda, P. J. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution* 11(10), 424–429. doi: [https://doi.org/10.1016/0169-5347\(96\)10049-5](https://doi.org/10.1016/0169-5347(96)10049-5)
- Kohn, M. H., Murphy, W. J., Ostrander, E. A., and Wayne, R. K. (2006). Genomics and conservation genetics. *Trends in Ecology & Evolution* 21(11), 629–637. doi: <https://doi.org/10.1016/j.tree.2006.08.001>
- Kosgey, I. S. and Okeyo, A. M. (2007). Genetic improvement of small ruminants in low-input, smallholder production systems: Technical and infrastructural issues. *Small Ruminant Research* 70(1), 76–88. doi: <https://doi.org/10.1016/j.smallrumres.2007.01.007>

- Krause, A. K. (2006). Breeding programmes for small ruminants in the tropics with special reference to the crossbreeding programme of the Dairy Goat Association of Kenya (DGAK). Ph.D. thesis, Humboldt University, Berlin, Germany.
- Laval, G., SanCristobal, M., and Chevalet, C. (2002). Measuring genetic distances between breeds: use of some distances in various short term evolution models. *Genetics Selection Evolution* 34(4), 1–27. doi: <https://doi.org/10.1186/1297-9686-34-4-481>
- Letunic, I. and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research* 47(W1), W256–W259. doi: <https://doi.org/10.1093/nar/gkz239>
- Mbuku, S. M., Okeyo, A. M., Kosgey, I. S., and Kahi, A. K. (2015). Optimum crossbreeding systems for goats in low-input livestock production system in Kenya. *Small Ruminant Research* 123(1), 55–61. doi: <https://doi.org/10.1016/j.smallrumres.2014.10.001>
- Mburu, M., Mugendi, B., Makhoka, A., and Muhoho, S. (2014). Factors Affecting Kenya Alpine Dairy Goat Milk Production in Nyeri Region. *Journal of Food Research* 3(6), 160–160. doi: <https://doi.org/10.5539/jfr.v3n6p160>
- McVean, G. (2009). A Genealogical Interpretation of Principal Components Analysis. *PLoS Genetics* 5(10), e1000686–e1000686. doi: <https://doi.org/10.1371/journal.pgen.1000686>
- Nicoloso, L., Bomba, L., Colli, L., Negrini, R., Milanese, M., Mazza, R., et al. (2015). Genetic diversity of Italian goat breeds assessed with a medium-density SNP chip. *Genetics Selection Evolution* 47(1). doi: <https://doi.org/10.1186/s12711-015-0140-6>
- Ogola, T. D. O., Nguyo, W. K., and Kosgey, I. S. (2010). Economic contribution and viability of dairy goats: implications for a breeding programme. *Tropical Animal Health and Production* 42(5), 875–885. doi: <https://doi.org/10.1007/s11250-009-9501-x>
- Ojango, J. M. K., Audho, J., Oyieng, E., Recha, J., Okeyo, A. M., Kinyangi, J., and Muigai, A. W. T. (2016). System characteristics and management practices for small ruminant production in “Climate Smart Villages” of Kenya. *Animal Genetic Resources/Ressources génétiques animales/Recursos genéticos animales* 58, 101–110. doi: <https://doi.org/10.1017/s2078633615000417>
- Onzima, R. B., Upadhyay, M. R., Mukiibi, R., Kanis, E., Groenen, M. A. M., and Crooijmans, R. P. M. A. (2018). Genome-wide population structure and admixture analysis reveals weak differentiation among Ugandan goat breeds. *Animal Genetics* 49(1), 59–70. doi: <https://doi.org/10.1111/age.12631>
- Palian, B. and Racokzi, G. (1976). Sheep and goat development project, Kenya. Breeding and research - Technical report 3.
- Peacock, C. (2005). Goats: Unlocking their potential for Africa’s farmers. In Proceedings of the Seventh Conference of Ministers Responsible for Animal Resources, 23p.
- Peacock, C. (2007). The goats model. A proven approach to reducing poverty among smallholder farmers in Africa by developing profitable goat enterprises and sustainable support services. url: <https://www.farmafrica.org/downloads/resources/WP9%20The%20Goat%20Model.pdf>.
- Peacock, C., Ahuya, C. O., Ojango, J. M. K., and Okeyo, A. M. (2011). Practical crossbreeding for improved livelihoods in developing countries: The FARM Africa goat project. *Livestock Science* 136(1), 38–44. doi: <https://doi.org/10.1016/j.livsci.2010.09.005>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., and Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics* 81(3), 559–575. doi: <https://doi.org/10.1086/519795>
- Qanbari, S. and Simianer, H. (2014). Mapping signatures of positive selection in the genome of livestock. *Livestock Science* 166, 133–143. doi: <https://doi.org/10.1016/j.livsci.2014.05.003>
- Rege, J. and Okeyo, A. (2011). Improving our knowledge of tropical indigenous animal genetic resources. In Animal Genetics Training Resource Version 3 Training Module 2, ILRI, Nairobi, Kenya; Uppsala, Sweden.
- Scarpa, R., Ruto, E. S., Kristjanson, P., Radeny, M., Drucker, A. G., and Rege, J. E. (2003). Valuing indigenous cattle breeds in Kenya: an empirical comparison of stated and revealed preference value estimates. *Ecological Economics* 45(3), 409–426. doi: [https://doi.org/10.1016/s0921-8009\(03\)00094-6](https://doi.org/10.1016/s0921-8009(03)00094-6)
- Shivairo, R. S., Matofari, J., Muleke, C. I., Migwi, P. K., and Lugairi, E. (2013). Production Challenges and Socio-Economic Impact of Dairy Goat Farming amongst Smallholder Farmers in Kenya. *Food Science and Quality Management* 17, 54–61.
- Tosser-Klopp, G., Bardou, P., Bouchez, O., Cabau, C., Crooijmans, R., Dong, Y., Donnadiou-Tonon, C., Eggen, A., Heuven, H. C. M., Jamli, S., Jiken, A. J., Klopp, C., Lawley, C. T., McEwan, J., Martin, P., Moreno, C. R., Mulsant, P., Nabihoudine, I., Pailhoux, E., Palhière, I., Rupp, R., Sarry, J., Sayre, B. L., Tircazes, A., Wang, J., Wang, W., and Zhang, W. (2014). Design and Characterization of a 52K SNP Chip for Goats. *PLoS ONE* 9(1), e86227–e86227. doi: <https://doi.org/10.1371/journal.pone.0086227>
- Vignal, A., Milan, D., SanCristobal, M., and Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution* 34(3), 275–305. doi: <https://doi.org/10.1186/1297-9686-34-3-275>
- Visser, C., Lashmar, S. F., Marle-Köster, E. V., Poli, M. A., and Allain, D. (2016). Genetic Diversity and Population Structure in South African, French and Argentinian Angora Goats from Genome-Wide SNP Data. *PLOS ONE* 11(5), e0154353–e0154353. doi: <https://doi.org/10.1371/journal.pone.0154353>

Waineina, R. W., Okeno, T. O., Ngeno, K., and Ilatsia, E. D. (2019). Performance of Dairy Goat Breeds in different Production Systems in Kenya. In Proceedings of the Tanzania Society of Animal Production (TSAP), 42nd Scientific Conference, 23-25.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., and Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28(24), 3326–3328. doi: <https://doi.org/10.1093/bioinformatics/bts606>