Mitochondrial DNA D-loop sequence variability reveals high haplotype diversity and multiple maternal origins in twelve indigenous goat populations from Tanzania.

Athumani Nguluma[1*]; Martina Kyalo[3]; Rose Loina[1]; Zabron Nziku[1]; Sebastian Chenyambuga[2]; Roger Pelle[3]

[1]Tanzania Livestock Research Institute

[2]Sokoine University of Agriculture, P. O Box 3000, Morogoro

[3]Biosciences Eastern and Central Africa-International Livestock Research Institute (BecA-ILRI) Hub, Nairobi, Kenya

Correspondence: Athumani Nguluma, Tanzania Livestock Research Institute, P. O. Box 147, Sanya Juu, Kilimajaro Tanzania.

Email: asnguluma@yahoo.com

Abstract

The Small East African (SEA) goat (*Capra hircus*) breeds are widely distributed in different agro-ecological zones of Tanzania. We report the genetic diversity, maternal origin, and phylogenetic relationship among the 12 Tanzanian indigenous goats populations, namely Fipa (n = 44), Songwe (n = 34), Tanga (n = 33), Pwani (n = 40), Newala (n = 49), Lindi (n = 46), Gogo (n = 73), Pare (n = 67), Maasai (n = 72), Sukuma (n = 67), and Ujiji (n = 67), based on the mitochondrial DNA (mtDNA) D-loop. High haplotype (Hd = 0.9619-0.9945) and nucleotide (π = 0.0120-0.0162) diversities were revealed from a total of 389 haplotypes. The majority of the haplotypes (h = 334) drawn from all the goat populations belonged to Haplogroup A which was consistent with the global scenario on the genetic pattern of maternal origin of all goat breeds in the world. Haplogroup G comprised of 45 haplotypes drawn from all populations except the Ujiji goat population while Haplogroup B with 10 haplotypes was dominated by Ujiji goats (41%). Tanzanian goats shared four haplotypes with the Kenyan goats and two with goats from South Africa, Namibia, and Mozambique. There was no sharing of haplotypes observed between individuals from Tanzanian goat populations with individuals from North or West Africa. The indigenous goats in Tanzania have high genetic diversity defined by 389 haplotypes and multiple maternal origins of haplogroup A, B and G. There is a lot of intermixing and high genetic variation within populations which represent an abundant resource for selective breeding in the different agro-ecological regions of the country.

Keywords: Indigenous goats; genetic variation; haplogroups; demographic history

Background

The Tanzanian goat (*Capra hircus*) population is currently estimated to be 24.1 million (NBS, 2020) with 97% comprising of indigenous goats belonging to the Small East African (SEA) goat breed (MLF, 2017). Due to their adaptability to different climatic conditions, the indigenous goats are widely distributed in almost all agro-ecological zones of Tanzania. Goats are important species for the livelihood of the rural farming communities especially those residing in arid and semi-arid areas of Tanzania where other agricultural activities are not feasible. They are raised mainly for meat and manure, and as a source of income (Chenyambuga *et al.,* 2012). Additionally, goats play different socio-cultural and traditional roles as gifts, dowry payments, and spiritual offerings. Despite their wide distribution in Tanzania, the indigenous goats have low productivity in terms of growth and milk production. Efforts to improve their productivity have mainly focused on crossbreeding with exotic germplasm which has proved to be unsustainable in the long run. Selective breeding utilizing the indigenous adapted animals would have a sustainable impact on the productivity of the animals (Syrstad and Ruane, 1998). This requires the animals to be characterized to understand the level of genetic diversity and the relationship between different animal populations (FAO, 1992; FAO/UNEP, 1998). The limited information on the characteristics of indigenous goats in Tanzania is mostly based on the phenotypic features which can be subjective and dependent on the environment which makes it difficult to distinguish between populations (Falconer & Mackay, 1996). Studying the genetic history of domestic animals can provide crucial clues about past events and main pathways used for commercial transport of the animals in historical times and therefore provide us with

information about their genetic structure and relationship within and among populations. Information from such studies is needed in designing and implementing conservation and improvement programs for indigenous goats. This study, therefore, was designed to determine the genetic diversity, maternal origin, and phylogenetic relationship of 12 populations of the indigenous goats in Tanzania using the mitochondrial DNA (mtDNA) D-loop region.

Materials and Methods

Sample collection and DNA extraction

A total of 627 blood samples were collected from unrelated indigenous female goats from all major geographical agro-ecological zones of Tanzania (Figure 1). The goats represented 12 populations namely Fipa from Rukwa (n = 44), Songwe from Songwe (n = 34), Tanga from Tanga (n = 33), Pwani from Pwani (n = 40), Newala from Mtwara (n = 49), Lindi from Lindi (n = 46), Gogo from Dodoma (n = 73), Pare from Kilimanjaro (n = 67), Maasai from Arusha and Kilimanjaro (n = 72), Sukuma from Mwanza (n = 67), and Ujiji from Kigoma (n = 67). Total genomic DNA was isolated from blood using the QIAGEN DNeasy Blood & Tissue Kit (Hilden, Germany) and TANBead OptiPure Blood DNA Extraction Kit (Taiwan Advanced Nanotech Inc., Taiwan) according to the manufacturer's protocol. The concentration and purity of extracted DNA was assessed using the Nanodrop1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA).

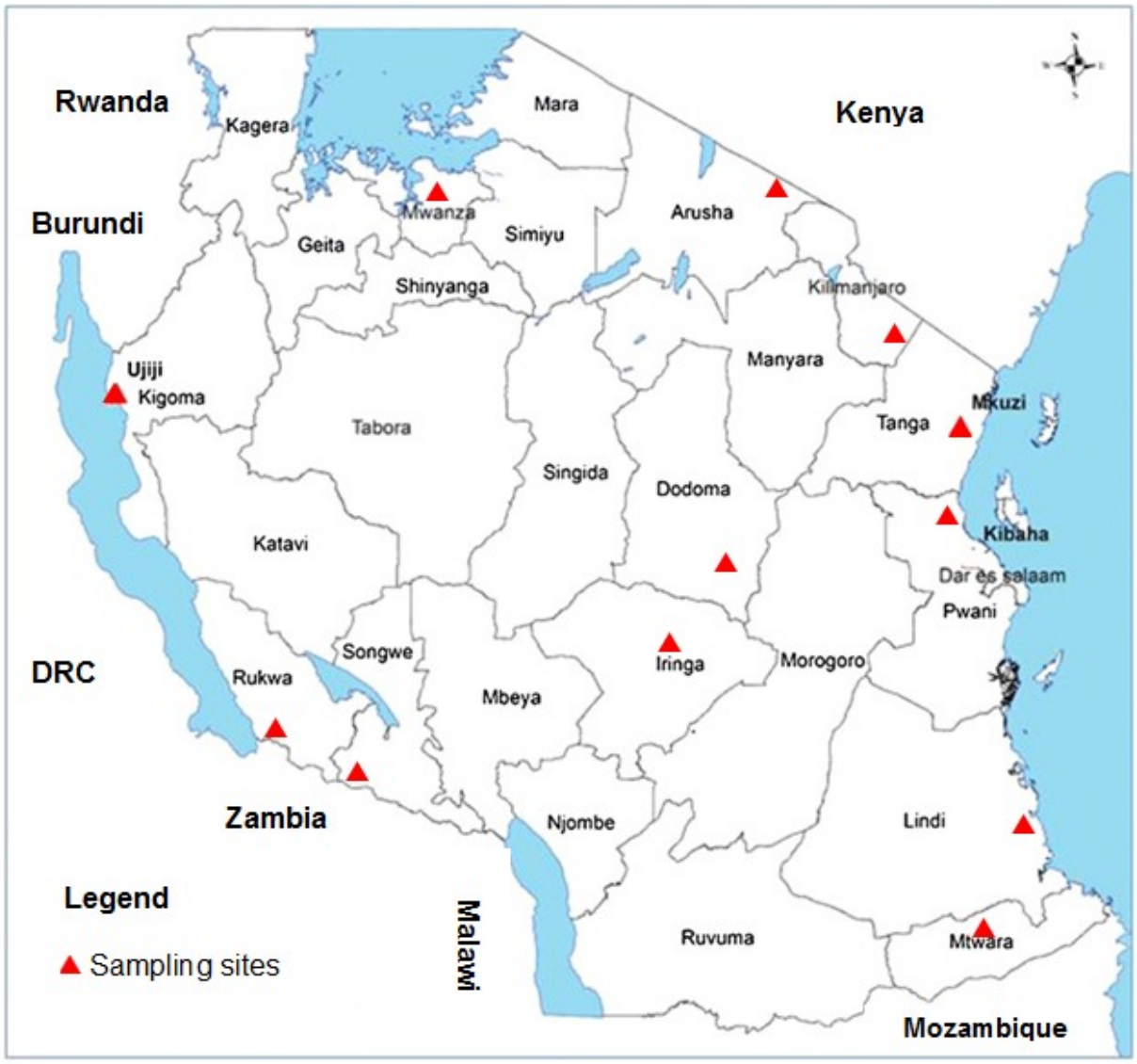


Figure 1. Map of Tanzania showing sampling sites

PCR amplification and sequencing

The primer pair 6807-F 5'-ACCAGAAAAGGAGAATAGCC-3' and 8173-R 5'-GGTACACTCATCTAGGCATT-3' flanking the mtDNA D-loop region were designed in this study to amplify the complete D-loop region. PCR amplification was performed in a total volume of 15 µL containing Phusion High-Fidelity PCR Master Mix (Thermo Fisher Scientific Inc., USA), 0.1 pM of each primer, 2% DMSO (Applied Biosystems, USA), and

50 ng of template DNA. Amplification was carried out in a GeneAmp PCR System 9700 thermal cycler using the following cycling conditions: initial denaturation at 98ºC for 30 sec, followed by 35 cycles at 98ºC for 10 sec, 62ºC for 30 sec, and 72°C for 1 min, with a final extension of 72°C for 10 min. The PCR products were fragment separated on a 1.5% agarose gel pre-stained with 0.25X GelRed (Biotium, USA) and visualized under UV light. The PCR fragment sizes were estimated using O'Gene 100 bp DNA ladder (Thermo Fisher Scientific Inc., USA). The ExoSAP-IT™ PCR Product Cleanup Reagent (Thermo Fisher Scientific Inc., USA) was used to purify PCR products before Sanger sequencing.

Data Analysis

Sequences were manually edited and aligned using Clustal W in CLC Workbench 8.0.3 (CLC Bio-Qiagen). The *Capra hircus* D-loop mtDNA sequence (Genbank accession number GU223571) was used for reference mapping during sequence assembly. Basic diversity parameters were computed for each population using Arlequin 3.5 (Excoffier & Lischer, 2010). Twenty-two goat mtDNA reference sequences belonging to six known haplogroups/lineages (Naderi et al., 2007) were downloaded from GenBank and used for haplogroup identification. Sequences of 189 individual goats belonging to 16 goat populations from nine other African countries (Table 1) were also downloaded from the GenBank and were included in the analysis. The neighbor-joining tree constructed with MEGA 6 software using all the generated sequences and the 22 reference sequences classified the populations into maternal lineages/haplogroups. The bootstrap percentage computed after 1000 replications, was used to assess the reliability of the

phylogenetic tree. To confirm the maternal origin and relationship between the Tanzania goat populations and populations from other African regions median-joining network was drawn using Network V4.6.1.3. To determine population genetic structure within and among populations, analysis of molecular variance (AMOVA) was performed using Arlequin Version 3.0 (Excoffier et al., 2005). In addition, structure analysis was performed using structure software to determine the number of gene pools in the indigenous goat populations in Tanzania.

History and demographic dynamics were investigated through mismatch distribution patterns (Rogers & Harpending, 1992) complemented by Fu's FS (Fu, 1997) and Tajima's D (Tajima, 1989) statistics calculated using the infinite sites model in Arlequin v3.5.

Table 1. Geographical location and characteristics of other African goat populations included in the study

Country	GenBank Accession number	Reference
Kenya	KP120622 - KP120681	Kibwega et al., 2015
Ethiopia	KY747687 - KY747691; KY747989 - KY747993	Getinet et al., 2018
Mozambique	AJ317804-809; EF618240 -1 EF618241	Luikart et al., 2001; Naderi et al., 2007
Zimbabwe	AJ317802 - AJ317803; EF618545 - EF618546	Luikart et al., 2001; Naderi et al., 2007
Namibia	EF618242 - EF618245	Naderi et al., 2007
Nigeria	AJ317777 - AJ317779	Luikart et al., 2001
Egypt	AJ317780 - AJ317783; AJ317795 - AJ317801; EF617711 - EF617728; EF618220	Luikart et al., 2001; Naderi et al., 2007
Algeria	AJ317777 - AJ317779	Luikart et al., 2001
South Africa	AJ317812 - AJ317815; AJ317819 - AJ317820; AJ317844; AJ317821 - AJ317822; EF618351 - EF618356; KJ466263 - KJ466273	Luikart et al., 2001; Naderi et al., 2007; Awotunde et al., 2015

Results

MtDNA D-loop variation and genetic diversity

The complete goat mtDNA D-loop region analyzed in this study corresponds to nucleotide positions 15431 to 16643 of the *C. hircus* reference sequence (Genbank accession number GU295658.1). From the 627 sequences generated for the 12 Tanzanian goat populations, 276 polymorphic sites were identified of which 223 were substitutions (214 transitions, 9 transversions) and 69 were indels. The polymorphic sites defined 389 haplotypes in total and of these 308 were unique whereas 81 were shared between individuals from at least two different populations. Maternal genetic diversity parameters for Tanzanian goat populations are presented in Table 2. All the populations showed high genetic diversity indicated by the haplotype diversity ranging between 0.9485 ± 0.011 in the Newala to 0.9945 ± 0.001 in the Sukuma goat population while haplotype proportion (number of haplotypes in relation to the sample size) was in the range of 59.2% in Newala to 95% in Fipa populations. Nucleotide diversity was the largest for Songwe (0.0162 ± 0.037) and lowest for the Lindi (0.0120 ± 0.031) goat populations. There were 69 indels in total for all the populations and, with 65 indels each, Newala and Lindi goats had the highest number of indels. Twenty-six individuals had sequences with ambiguous nucleotides at various positions and Ujiji goats had the highest number of individuals with ambiguous nucleotides (n = 14).

Table 2. Genetic diversity parameters for 627 goats from 12 Tanzanian goat populations

Population	N	H	Hd ± SD	π ± SD	I	Number of individuals (%)		
						A	B	G
Fipa	44	35	0.9749 ± 0.006	0.0145 ± 0.051	2	36 (81.8)	0	8 (18.2)
Songwe	34	29	0.9763 ± 0.005	0.0162 ± 0.037	3	25 (73.5)	1 (2.9)	8 (23.5)
Tanga	33	28	0.9734 ± 0.006	0.0144 ± 0.033	2	25 (75.8)	2 (6.0)	6 (18.2)
Pwani	40	38	0.9848 ± 0.003	0.0151 ± 0.035	3	32 (80)	2 (5)	6 (15)
Iringa	35	31	0.9789 ± 0.005	0.0169 ± 0.061	4	26 (74.3)	3 (8.6)	6 (17.1)
Maasai	71	67	0.9909 ± 0.002	0.0130 ± 0.044	3	60 (84.5)	0	11 (15.5)
Newala	49	29	0.9485 ± 0.011	0.0128 ± 0.033	65	43(87.8)	0	6 (12.2)
Lindi	46	29	0.9565 ± 0.009	0.0120 ± 0.031	65	41(89.1)	2 (4.3)	3 (6.5)
Gogo	73	65	0.9874 ± 0.002	0.0139 ±0.047	4	60 (82.2)	1 (1.4)	12 (16.4)
Sukuma	67	51	0.9945 ± 0.001	0.0134 ± 0.046	2	58 (86.6)	2 (3.0)	7 (10.4)
Pare	67	54	0.9847 ± 0.003	0.0139 ± 0.126	19	57 (85.1)	0	10 (14.9)
Ujiji	68	53	0.9619 ± 0.007	0.0135 ± 0.044	1	59 (86.8)	9 (13.2)	0
Overall	**627**	**389**	**0.9945 ± 0.001**	**0.0139 ± 0.046**	**69**	**522 (83.3)**	**22 (3.5)**	**83 (13.2)**

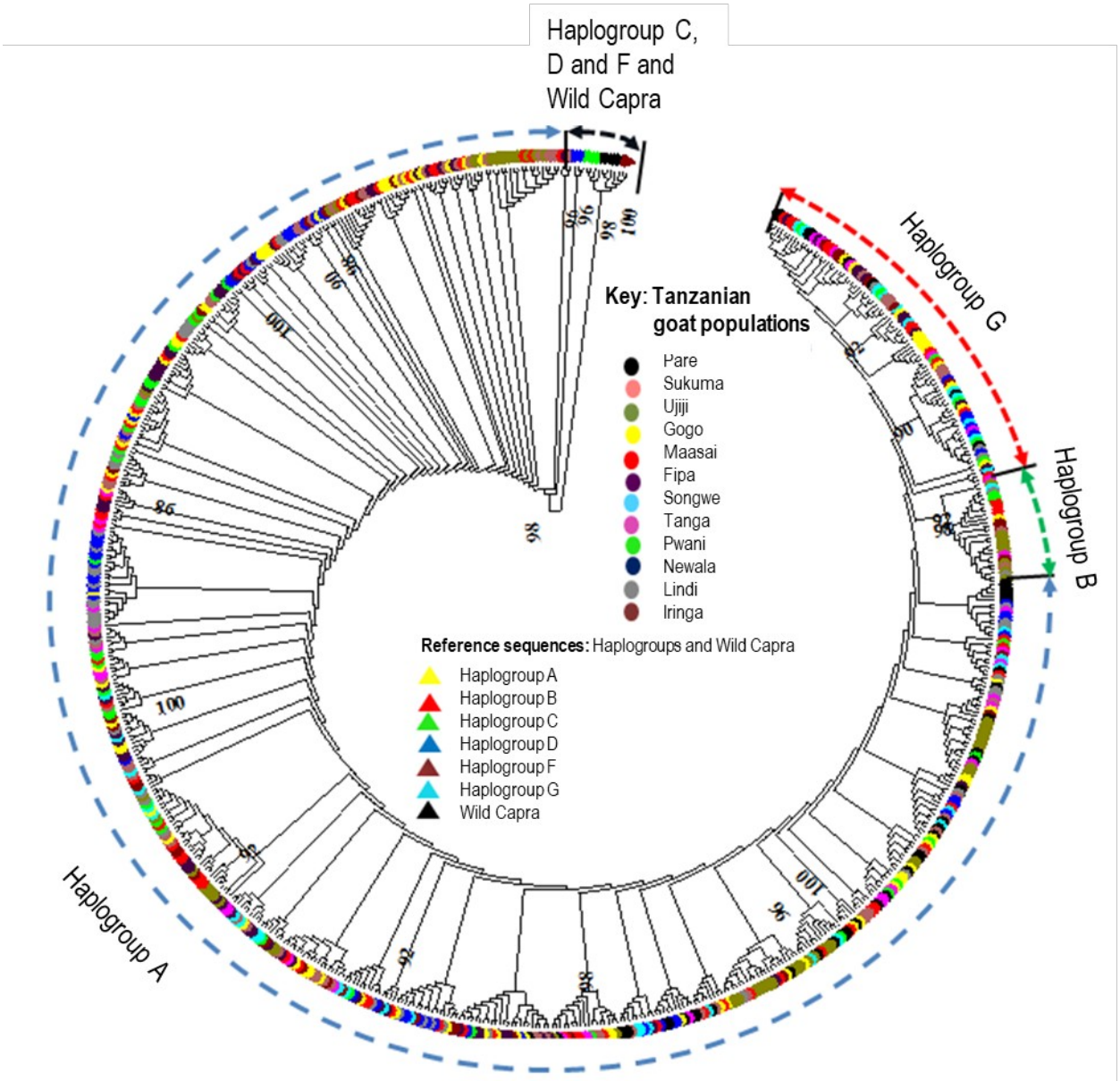
N, sample size; H, number of haplotypes; Hd, haplotype diversity; π, nucleotide diversity; I, Number of indels; SD, Standard deviation

159

160 **Phylogenetic relationship**

161 Phylogenetic analysis was conducted to assess the relationship among the 12
162 Tanzanian goat populations and assign individual goats to their respective maternal
163 origins. The hypervariable region I (HV1) of the mtDNA D-loop which is 481 bp long
164 corresponding to the positions 15737 to 16189 on the *C. hircus* complete mitochondrial
165 reference sequence (Accession number GU295658.1) were used to construct the
166 neighbor-joining (NJ) tree and median-joining (MJ) network. Both, the NJ tree (Figure 2)
167 and MJ network (Figure 3) showed that the Tanzanian goat populations were classified

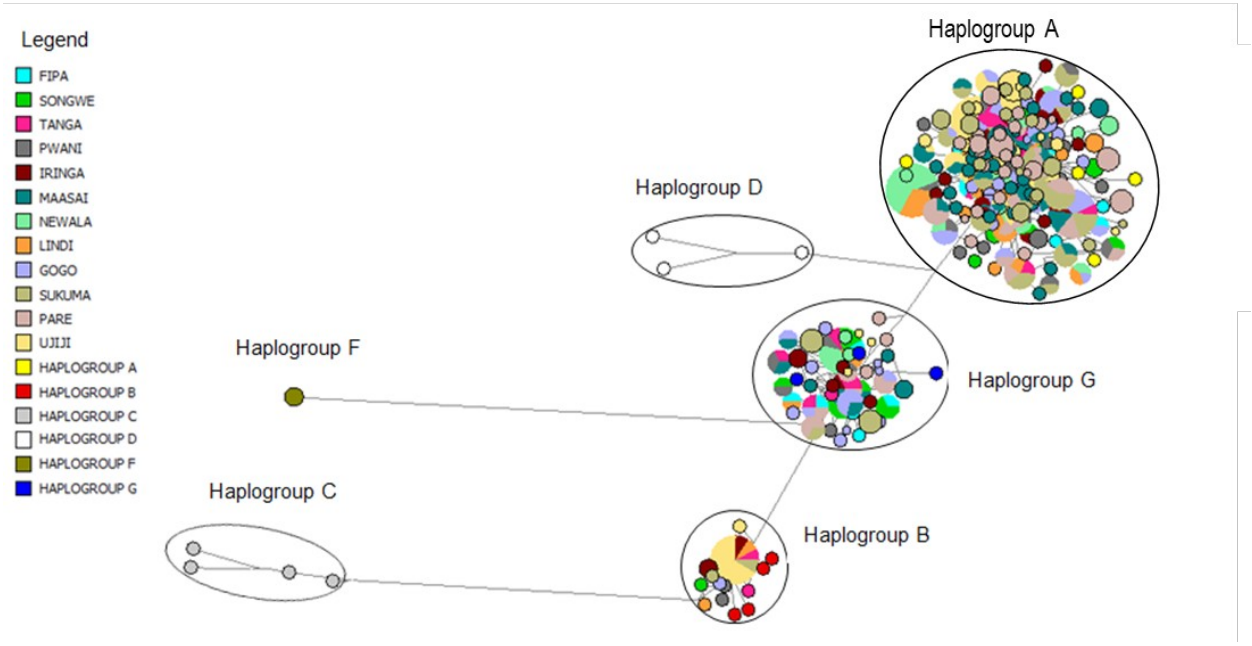
into three distinct groups which represented Haplogroup A, B, and G. Haplogroup A was the most predominant and contained 334 haplotypes representing 522 individuals drawn from all goat populations. Haplogroup G contained 83 individuals from all populations except Ujiji which was comprised of 45 haplotypes. Haplogroup B had only 22 individuals and 10 haplotypes representing 3.5% of all goats mostly from the Ujiji population. An MJ network used to assess the relationship among the Tanzanian goat populations revealed that Maasai and Gogo populations had the highest number of shared haplotypes (n = 10) while only one haplotype was shared by Ujiji-Newala, Ujiji-Pwani, Sukuma–Lindi, and Sukuma–Tanga pairs of populations. The most commonly shared haplotype (H119) was shared among six populations namely Newala, Gogo, Iringa, Ujiji, Maasai, and Sukuma while the most frequent haplotype (H85) occurred in 14 individuals from Newala, Pwani, and Lindi populations.



180

181 **Figure 2**. Neighbor-joining tree constructed using the HV1 region of the mtDNA D-loop
182 of 12 Tanzanian goat populations, reference sequences representing six haplogroups
183 observed in goats and five wild ancestors.

184



185

186 **Figure 3.** Median-joining network for the 389 mtDNA haplotypes of indigenous goats
187 represented by different colours. The area of the circle is proportional to haplotype
188 frequency.

189 To provide a wider resolution of the phylogenetic relationship between the Tanzanian
190 goat populations with those of other regions of the African continent, an MJ network was
191 constructed using sequences from Tanzanian goats in this study and goat populations
192 from across nine African countries retrieved from the GenBank database (Figure 4).
193 Three clusters were again formed with global haplogroup A comprising of individuals
194 from all countries. Haplogroup B was comprised of individuals from Tanzanian, South
195 African, and Namibian goat populations. Haplogroup G was occupied by individuals
196 from Tanzania, Egypt, Ethiopia, and Kenya. The MJ network further revealed that
197 Tanzanian goats shared many haplotypes with Kenyan goats and a few with goats from
198 South Africa, Namibia, and Mozambique. There was no sharing of haplotypes observed

between individuals from Tanzanian goat populations with individuals from North and West Africa (Nigeria, Egypt and Algeria).

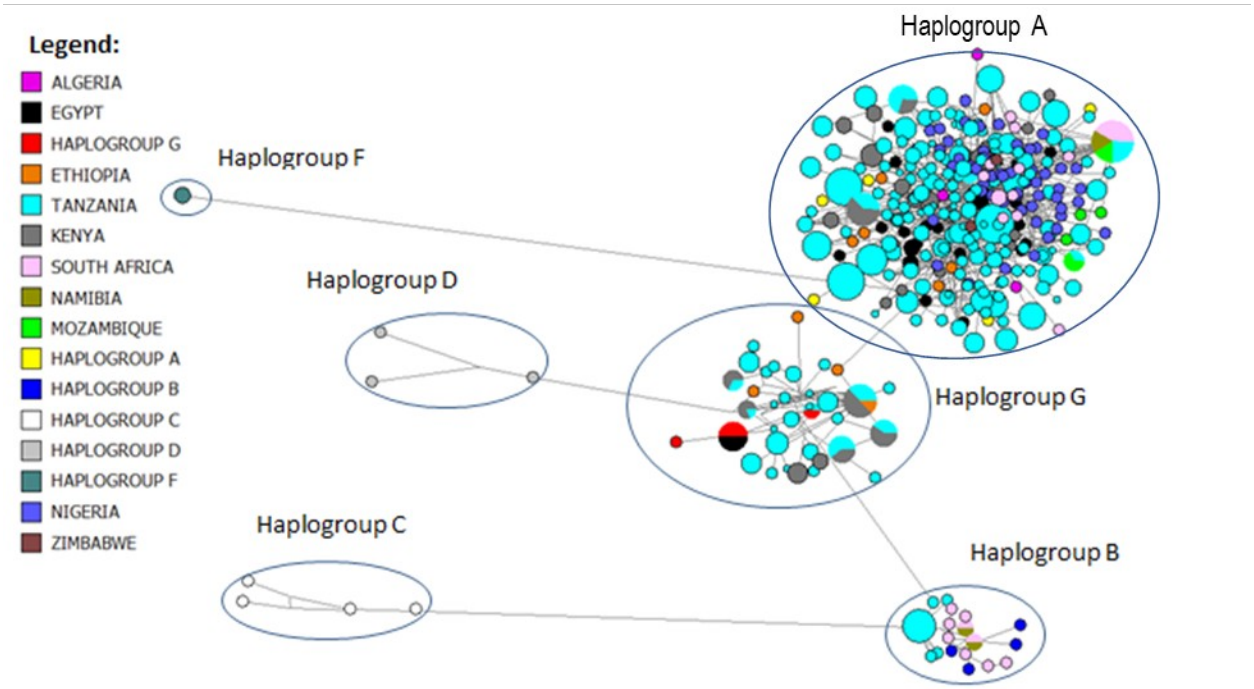


Figure 4. Median-joining network based on the haplotypes of HV1 control region of indigenous Tanzanian goats and goats from nine different African countries. The different colours are related to the geographical origin and the area of the circle is proportional to haplotype frequency.

Population Structure

The AMOVA results shown in Table 3 below revealed that a non-significant ($P<0.05$) proportion (2.8%) of the total genetic variation occurred among the goat populations while 97.2% of the total genetic variation was observed within the Tanzanian goat populations. Comparison with goats from different regions of the African continent revealed significant genetic variation at 12%. Structure analysis grouped the animals

212 into four gene pools (Figure 5) with the largest gene pool and the smallest gene pools
213 comprising 38.3% and 7.8% of all the animals from the 12 goat populations studied
214 respectively (Table 4). All the gene pools were present in each of the population and
215 only Newala and Lindi populations had more than half of their individual goats belonging
216 to one gene pool and the rest of the populations.

217

Table 3. Results of AMOVA based on the analysis of the complete mtDNA d-loop in 12 Tanzanian goat populations

	Tanzanian goat populations		**African regions' goat populations**[1]	
Source of variation	**Among populations**	**Within populations**	**Among regions**	**Within regions**
DF	11	1242	4	327
SS	393.94	11060.45	198.901	1943.625
Variance	0.259	8.91	0.84407	5.94381
% of Variation	2.83	97.17	12.43	87.57
P value	0.063	0.0028*	0.048*	0.019*

* Statistically significant ($P < 0.05$)

DF – Degree of freedom; SS – Sum of squares

[1] Four geographic regions defined as; Tanzania, East Africa, North Africa, West Africa, and South Africa

218

219

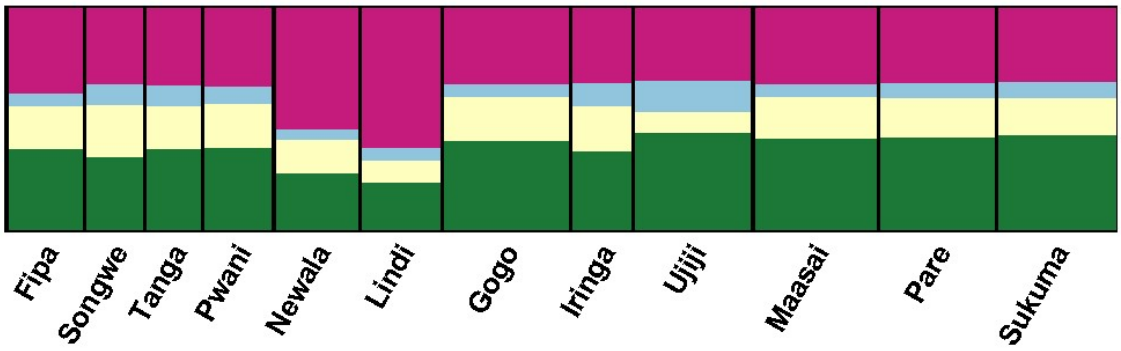


Figure 5. Structure plot of 12 indigenous goat populations of Tanzania.

.

Table 4. Population assignment to four different gene pools

Population	Gene pool I	Gene pool II	Gene pool III	Gene pool IV	N
Fipa	0.3663	0.1934	0.0603	0.3800	44
Songwe	0.3321	0.2344	0.0944	0.3391	34
Tanga	0.3660	0.1953	0.0916	0.3471	33
Pwani	0.3733	0.1992	0.0764	0.3511	40
Iringa	0.3563	0.2050	0.1042	0.3345	35
Maasai	0.4155	0.1853	0.0570	0.3422	71
Newala	0.2542	0.1529	0.0488	0.5441	49
Lindi	0.2160	0.0960	0.0599	0.6281	46
Gogo	0.4052	0.1944	0.0615	0.3389	73
Sukuma	0.4304	0.1679	0.0700	0.3318	67
Pare	0.4168	0.1807	0.0659	0.3366	67
Ujiji	0.4376	0.0947	0.1424	0.3253	68
Overall	0.3639	0.1748	0.0777	0.3833	627

Population Demographic History

Past population expansion events were inferred based on the pattern of the mismatch distributions and neutrality test estimates (Fu's Fs and Tajima's D statistics) presented in Table 5. The Tajima's D values for the goat populations were negative and not

229 significantly different from zero for all populations while Fu's estimates were not
230 statistically significant except for Gogo and Maasai whose Fu's F values were negative
231 and significant. For all populations, the mismatch distributions showed ragged and
232 bimodal pattern (Figures 6). Furthermore, the SSD and Happending's raggedness index
233 (r) computed to ascertain the goodness of fit of the mismatch distributions varied among
234 populations (Table 5). Estimates for SSD and raggedness index values were positive
235 and non-significant for all populations.

236

Table 5. Population demographic parameters estimated from the analysis of the complete mtDNA D-loop in 12 Tanzanian goat populations

Population	N	SSD	Raggedness index (r)	Tajima's D	Fu's Fs
Fipa	44	0.021	0.010	0.132	-0.778
Songwe	34	0.026	0.013	-0.167	0.814
Tanga	33	0.033	0.014	-0.134	0.300
Pwani	40	0.017	0.006	-0.484	-2.461
Iringa	35	0.070	0.012	-0.6	-0.80
Maasai	68	0.011	0.005	-0.771	-18.972 (0.004)
Newala	49	0.023	0.011	-0.253	2.184
Lindi	46	0.011	0.014	-0.635	1.458
Gogo	67	0.015	0.005	-0.648	-14.543 (0.022)
Sukuma	73	0.011	0.006	-0.627	-5.644
Pare	72	0.008	0.004	-0.547	-6.788
Ujiji	67	0.017	0.009	0.126	--8.623
Haplogroup A	522	0.018	0.008	-1.364	-24.06(0.004)
Haplogroup B	22	0.016	0.008	-1.589	-1.044
Haplogroup G	83	0.019	0.007	-1.165	-14.069 (0.00)
Mean	346	0.012	0.006	-0.499	-11.40 (0.06)

N, sample sizes; S, segregating sites; SSD, sum of squared deviations.

237

238

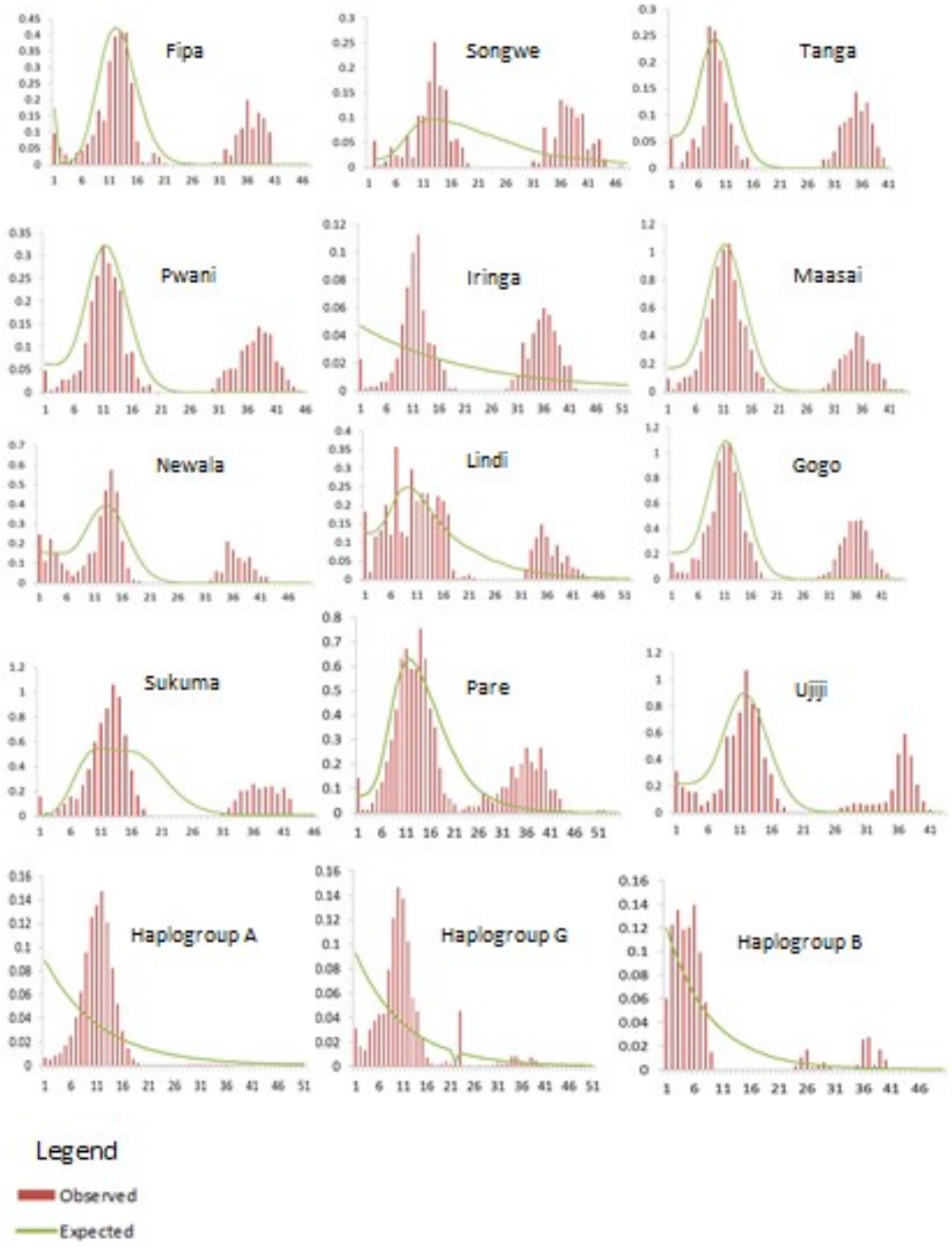


Figure 6. Mismatch distribution graphs for each of the 12 Tanzanian indigenous goat populations and Haplogroup A, B and G analyzed in this study. The x axis shows the number of pairwise differences, the y axis shows the frequency of the pairwise comparisons.

Discussion

Genetic diversity and maternal origin

To understand the origin and diversity of goats in Tanzania, the genetic analysis of the maternally inherited mtDNA was conducted. In this study, high genetic diversity is depicted by high within-population haplotype and nucleotide diversity. Other studies reported haplotype diversity ranging from 0.95 to 1 for other African goat populations (Awotunde et al, 2015; Kibegwa et al, 2015; Kivila et al, 2018). The high genetic diversity observed in this study may partly be attributed to reasons mentioned previously including high mutation rate of the control region, multiple maternal wild ancestors (Naderi et al, 2007), and capture of a large part of the wild diversity during domestication (Benjelloun et al, 2011). The overall ratio of transitions: transversions (28.7:1) revealed a heavy transition bias even higher than the 17:1 and 16.7 ratios reported before in domestic goats (Luikart et al., 2001; Joshi et al., 2004). The presence of a large number of indels within the Lindi and Mtwara populations and a large number of ambiguous nucleotides in the Ujiji population compared to other populations points to a possibility of having unique gene pools and unique genetic features in those populations which need further investigation.

Using the available goat mtDNA haplogroup classification system (Naderi et al. (2007), three haplogroups (A, B, and G) were found among the indigenous Tanzanian goat populations. The predominance of haplogroup A was consistent with the world scenario described in previous studies (Liu et al. (2009); Sultana et al. (2003); Joshi et al. (2004) in all goat breeds in the world. Results obtained in this study further support the concept of multiple maternal origins of domestic goats (Luikart et al., 2001; Joshi et al., 2004).

Classification of the Tanzanian indigenous goats into three distinct haplogroups could be interpreted as evidence that they come from three separate genetically distinct maternal populations, or from one origin which is an extremely large population containing three highly divergent maternal lineages. There was no specific distribution pattern of goat populations in the different agro-ecological zones in haplogroup A since each population was represented in the haplogroup. Haplogroup B on the other hand was dominated by Ujiji goats while Newala, Fipa, Pare and Maasai goats were completely absent. Haplogroup G had members from all populations except the Ujiji goat population. This again indicates that Ujiji goats might possess some uniqueness due to limited gene flow from other populations caused by reproductive isolation. Kigoma region where Ujiji goats are predominantly raised is in the western part of the country which for a long time was geographically isolated from other regions due to its biophysical barriers and poor infrastructure and therefore less movement of people from other goat keeping communities into the region.

Tanzanian goat sequences in this study were compared with goat sequences from nine other countries representing different regions of Africa. This was done considering the possibility of intermixing or of sharing similar migration routes from the point of initial domestication which is considered to be South West Asia (Payne and Wilson, 1999). This could, therefore, enable tracing the path of the arrival of domestic goats into East Africa and Tanzania in particular. Haplogroup A which is said to have originated from Eastern Anatolia is the most diverse, ancient, and widely distributed in the world (Naderi et al., 2007) which is why they are found in all goat populations regardless of the geographical origin. Haplogroup G has been reported in Egypt, Ethiopia and Kenya and

its presence in Tanzanian goats could indicate the same maternal origins and/or gene flow between the goat populations from the countries. Egypt is reported to be one of the historical entry points of domesticated animals into the African continent; these results seem to indicate that one of the routes of introduction of goats into Tanzania was through southward movement from Egypt through Sudan, Ethiopia, and Kenya. Similar observations were made by Tarekegn et al. (2018) when tracing the route of dispersal of goats into Cameroon and Ethiopia from their origin of domestication. Haplogroup B is mostly found in whole Asia and Sub-Saharan Africa and its presence in other parts of the world is through human migration (Naderi et al. 2007). The presence of Haplogroup B in goats from Tanzania and from South Africa, Namibia, Mozambique and its absence in the Ethiopian and Kenyan goats might imply introduction of goats into southern African countries via a marime diffusion route through the Indian Ocean from Asia. Then, the Haplogroup was introduced into Tanzania through gene flow facilitated by historic annual long-distance migrations of the Nguni people from South Africa to Tanzania. Haplogroups C, D, and F have been reported in Europe (Naderi et al., 2007) and the absence of these Haplogroups in the analyzed samples may indicate that the indigenous goats of Tanzania do not share a recent common ancestor with the European goat breeds and that there is no or little introgression of European germplasm into Tanzanian indigenous goats. This was expected given that mtDNA is maternally inherited and the few crossbreeding programs involving European dairy goat breeds in Tanzania use breeding bucks or semen. Additionally, such programs are implemented in areas with intensive goat production systems in which only dairy goats are raised in small numbers.

Population structure and phylogenetic relationship

The AMOVA results revealed high variation within the Tanzanian goat populations but very low and insignificant among population variation. The lack of genetic structure observed in the indigenous goats of Tanzania is further supported by population structure analysis. The structure analysis which separated the goats into four different gene pools did not show any clear pattern of differentiation among all populations since all populations were represented in each gene pool. However, as already explained above, Lindi and Newala populations showed a relatively higher level of differentiation since the majority (more than 54%) of individuals in those populations belonged to only one gene pool. Mtwara and Lindi regions where Newala and Lindi goats are found are in the Southern part of the country and until recently there was limited movement of pastoralists from other parts of the country into the regions due to the prevalence of cattle trypanosomosis. In another study using microsatellite markers, Nguluma et al. (2017) reported a higher proportion (8%) of the between-population variation for indigenous goats in Tanzania. The low genetic variation observed between Tanzanian goat populations could be attributed to the intermixing of animals across geographical regions due to pastoralism (Mwambene et al, 2014; Tenga et al, 2008), trade between people of different regions, and cultural influences like dowry payments. High within-population variation observed in the present study could be contributed by uncontrolled mating and lack of selective breeding practiced by smallholder farmers in Tanzania as also observed by Tarekegn et al. (2018) for Cameroonian goats. Lack of proper and structured breeding programs with clear breeding strategies leads to lack of population genetic structure observed in the present study. The movement of livestock keepers

with their animals in search of water and pasture has been one of the main defining characteristics of the pastoral production system in which most of the goat production in Tanzania takes place. Even in the Southern regions of Tanzania particularly Mtwara and Lindi where livestock production was relatively less common have experienced a rapid increase in the number of goats partly due to the recent high influx of pastoralists from other regions of Tanzania who have been forced to look for grazing areas bringing into the regions new goat genotypes (Mwambene et al, 2014). The high within-population diversity observed for the Tanzanian goats could be very useful when planning and implementing community-based breeding programs in which farmers select and breed animals within their herds.

Significant population sub-structuring was observed when goat populations from different regions of the African continents were considered in the analysis consistent with what was reported for Sub-Saharan African goat breeds with 14% of inter-population variation using microsatellites markers (Chenyambuga et al, 2002). This is due to physical distance and lack of interaction between some regions of Africa like between East Africa and West Africa.

Demographic History

The bimodal pattern of distribution observed for each population indicates that the populations were in equilibrium or stable (Rogers and Harpending, 1992; Hartl, 2004) consistent with the neutrality statistics of Tajima's D and F's which were non significant. A similar demographic pattern has been observed in Ethiopian (Tarekegn et al, 2018) and Somalian indigenous goats (Al-Araimi et al, 2017), but not in Nigerian goats (Awotunde et al, 2015; Okpeku et al, 2016). On the contrary the Harpending's r and

SSD values did supported population expansion. Geographical barriers and small migration rates between populations could have caused an increase in the mean of the mismatch and thereby making the distribution of expanding populations as bimodal as observed by Rana et al. (2013) for Indian breeds. A separate analysis for Haplogroups revealed a smooth unimodal distribution for Haplogroup A indicating population expansion consistent with some previous studies (Joshi et al., 2004; Hou et al., 2008; Zhao et al., 2011) but contrary to other studies (Getinet et al, 2018; Kibegwa et al, 2015) who reported unimodal peaks for Haplogroup A. Bimodal peaks though not very distinct, were observed for Haplogroup B and G similar to what was observed previously (ibid) for Ethiopian and Kenyan goats implying a mutation-drift equilibrium for the haplogroups.

Conclusion

This is the first study that investigates the genetic diversity within and between the indigenous goat populations using samples from all the agro-ecological zones where goat production is practiced in Tanzania.The goats have high genetic diversity and come from three maternal origins A, B and G with the majority of them originating from Haplogroup A. There are four major gene pools within the indigenous goats of Tanzania which are admixed and very low genetic variation between populations. Popuation expansion occurred in Haplogroup A and the individual populations but due to small migration rates between them the mismatch distributions appeared as stable populations. High genetic diversity observed within indigenous goat populations presents an abundant resource for selective breeding in the different agro-ecological

regions of the country. Information generated in this study provides a valuable tool for conservation strategies and the data herein indicate that for many of the popuations, the inherent genetic diversity has been sucesfully maintained. The adaptive traits and other unique features in these populations need to be well studied, understood, and preserved in the breed improvement programs as a strategy for conservation of animal genetic resources.

ACKNOWLEDGMENTS

This project was supported by the BecA-ILRI Hub through the Africa Biosciences Challenge Fund (ABCF) program. The first and second authors were under the ABCF program and they are grateful to the Tanzania Livestock Research Institute (TALIRI) for granting them a study leave. The ABCF Program is funded by the Australian Department for Foreign Affairs and Trade (DFAT) through the BecA-CSIRO partnership; the Syngenta Foundation for Sustainable Agriculture (SFSA); the Bill & Melinda Gates Foundation (BMGF); the UK Department for International Development (DFID); and the Swedish International Development Cooperation Agency (Sida). Authors greatly appreciate the owners of the animals used in this study and the district extension office in the sampling areas for supporting the sampling process.

CONFLICT OF INTEREST

There are no conflicts of interest between authors, and permission from each author has been granted.

AUTHOR CONTRIBUTIONS

Athumani Nguluma conducted sample collection, laboratory and data analyses, and interpretation as well as writing of the manuscript. Rose Loina conducted sample collection and laboratory analysis. Martina Kyalo supervised sample analyses in the laboratory, data analyses and interpretation and proof-read and refined the manuscript. Roger Pelle devised the study and supervised its execution and assisted with writing the manuscript. Sebastian Chenyambuga and Zabron Nziku assisted in interpretation of the data, manuscript proof-reading and refining.

DATA AVAILABILITY STATEMENT

Mitochondrial sequence data generated as part of this project are stored in FASTA format and will be uploaded to the NCBI/GenBank nucleotide sequence repository. Other mtDNA sequences incorporated into the analysis were downloaded from this source and can be retrieved as per the relevant citations.

REFERENCES

1. Awotunde EO, Bemji MN, Olowofeso O, James IJ, Ajayi OO, Adebambo AO. Mitochondrial DNA Sequence Analyses and Phylogenetic Relationships Among Two Nigerian Goat Breeds and the South African Kalahari Red, Animal Biotechnology 2015; 26(3): 180–187,
2. Al-Araimi NA, Al-Atiyat RM, Gaafar OM, Vasconcelos R, Luzuriaga-Neira A, Eisa MO, et al. Maternal genetic diversity and phylogeography of native Arabian goats. Livestock Science 2017a; 206: 88–94.
3. Benjelloun B, Pompanon F, Ben Bati M, et al. Mitochondrial DNA polymorphism in Moroccan goats. Small Rumin Res 2011; 98: 201–205.

4. Chenyambuga S W, Watts P C, Hirbo J, Kemp S J, Hanotte O. 2002. Analysis of genetic diversity and relationships of Tanzanian local goat populations using microsatellite DNA markers. Tanzania Journal of Agricultural Science, 5: 29–38
5. Chenyambuga S W, Komwihangilo D M and Jackson M 2012: Production performance and desirable traits of Small East African goats in semi-arid areas of Central Tanzania. *Livestock Research for Rural Development. Volume 24, Article #118.* Retrieved April 30, 2020, from http://www.lrrd.org/lrrd24/7/chen24118.htm
6. Devendra C, Burns M. Goat Production in the Tropics (Revised Edn.). Technical Communication Bureaux of Animal Breeding and Genetics, Commonwealth Agricultural Bureau, England, 1983; pp. 183;
7. FAO. 1992. The management of global animal genetic Resources. Proceedings of an Expert Consultation, Rome, Italy, April 1992. Edited by J. Hodges. Animal Production and Health Paper No.104. Rome.
8. FAO/UNEP. 1998. Primary guidelines for development of national farm animal genetic resources management plans. Rome.
9. Hartl D L. 2004. The Origin of malaria: mixed messages from genetic diversity. Nature Reviews, 2: 15-22
10. Hou, L., J. Wang, P. Li, J. Li, and F. Xing. 2008. MtDNA D-loop polymorphism and phylogeny of sympatric goat populations in lower valley of Yellow river. Chin. J. Vet. Sci. 28:1474–1479.
11. Joshi M B, Rout P K, Mandal A K, Tyler-Smith C, Singh L, Thangaraj K. 2004. Phylogeography and origin of Indian domestic goats. Mol Biol Evol, 21: 454–462.

12. Kibegwa F M, Githui K E, Jung'a J O, Badamana M S, Nyamu M N. 2015. Mitochondrial DNA variation of indigenous goats in Narok and Isiolo counties of Kenya. J Anim Breed Genet, 133: 238–247.
13. Liu Y P, Cao S X, Chen S Y, Yao Y G, Liu T Z. Genetic diversity of Chinese domestic goat based on the mitochondrial DNA Sequence Variation. J Anim Breed Genet 2009; 126: (1) 80–89.
14. Luikart G, Gielly L, Excoffier L, Vigne J D, Bouvet J, Taberlet P. Multiple maternal origins and weak phylogeographic structure in domestic goats. Proc Natl Acad Sci USA, 2001; 98: 5927–5932.
15. Ministry of Livestock and Fisheries. Tanzania Livestock Master Plan (2017/2018 – 2021/2022). Available at https://www.mifugouvuvi.go.tz/uploads/projects/1553601793-TANZANIA %20LIVESTOCK%20MASTER%20PLAN.pdf
16. Msanga Y N, Mbaga S H, Msechu J K. Farm Animals Breeds and Strains of Tanzania. In: Proc of SUA-MU ENRECA Project Workshop on Farm Animal Genetic Resources 2001; 6th August 2001, Morogoro, Tanzania, pp. 36-49.
17. Mwambene P L, Mbwile R P, Hoeggel F U, Kimbi E C, Materu J, Mwaiganju A and Madoffe S. 2014. Assessing dynamics of forced livestock movements, livelihoods and future development options for pastoralists/agro-pastoralists in Ruvuma and Lindi Regions, in the Southern Tanzania. *Livestock Research for Rural Development. Volume 26, Article #4.* Retrieved from http://www.lrrd.org/lrrd26/1/mwam26004.htm.

18. Naderi S, Rezaei H R, Taberlet P, et al. Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. PLoSONE 2007; 2: (10), e1012.
19. National Bureau of Statistics. 2020. The National Sample Census of Agriculture, Livestock and Fisheries for Agricultural Year 2019/20. Accessed from https://www.nbs.go.tz/index.php/en/census-surveys/agriculture-statistics/579-press-release-key-findings-on-the-national-sample-census-of-agriculture-livestock-and-fisheries-for-agricultural-year-2019-20.
20. Nguluma, A. S., Huang, Y., Zhao, Y., Chen, L., Msalya, G., Lyimo, C. 2018. Assessment of genetic variation among four populations of Small East African goats using microsatellite markers. S. Afr. J. Anim. Sci. 48, 117–127. doi: 10.4314/sajas.v48i1.14
21. Okpeku M, Peters SO, Imumorin IG, Caires KC, Sharma VK, Wheto M, et al. Mitochondrial DNA hypervariable region 1diversity in Nigerian goats. Animal Genetic Resources, 59: 47–54. https://doi.org/10.1017/S2078633616000102.
22. Rana, S; Singh, S; Dureja, V; Joshi. Mitochondrial D-loop analysis reveals single maternal lineage for coastal region goats of India. 2013. DHR International Journal Of Biomedical and Life Sciences (DHR-IJBLS), 4(1): 2278-8301.
23. Rogers A R, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol, 9: 552–569.
24. Schneider S, Excoffier L. Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. Genetics 1999; 152: 1079-1089.

499 25. Slatkin M, Hudson R R. Pairwise comparisons of mitochondrial DNA sequences
500 in stable and exponentially growing populations. Genetics 1991; 129: 555-562.
501 26. Syrstad O. and Ruane J. 1998. Prospects and strategies for genetic
502 improvement of the dairy potential of tropical cattle by selection. Tropical Animal
503 Health and Production, 30: 257 – 268
504 27. Tajima H. Statistical method for testing the neutral mutation hypothesis by DNA
505 polymorphism. Genetics 1989; 123: 585-595.
506 28. Tarekegn GM, Tesfaye K, Mwai OA, Djikeng A, Dessie T, et al. 2018.
507 Mitochondrial DNA variation reveals maternal origins and demographic dynamics
508 of Ethiopian indigenous goats. Ecol Evol 2018; 8(3):1–11.
509 https://doi.org/10.1002/ece3.3710.
510 29. Tarekegn GM, Wouobeng P, Jaures KS, Mrode R, Edea Z, Liu B. 2019.
511 Genome-wide diversity and demographic dynamics of Cameroon goats and their
512 divergence from east African, north African, and Asian conspecifics. PLoS ONE
513 14(4): e0214843. https://doi.org/10.1371/journal. pone.0214843
514 30. Tenga R, Mattee A, Mdoe N, Mnenwa R, Mvungi S and Walsh M 2008. A study
515 on options for pastoralists to secure their livelihoods in Tanzania. Current policy,
516 legal and economic issues. 139pp.