




# Genome-wide population structure and admixture analysis reveals weak differentiation among Ugandan goat breeds

R. B. Onzima\*<sup>†</sup> , M. R. Upadhyay\*<sup>‡</sup>, R. Mukiibi<sup>§</sup>, E. Kanis\*, M. A. M. Groenen\* and R. P. M. A. Crooijmans\*

\*Wageningen University & Research, Animal Breeding and Genomics, P O Box 338, 6700AH Wageningen, The Netherlands. <sup>†</sup>National Agricultural Research Organization (NARO), P O Box 295, Entebbe, Uganda. <sup>‡</sup>Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden. <sup>§</sup>Department of Agriculture, Food and Nutritional Sciences (AFNS), Faculty of Agriculture, Life and Environmental Sciences, University of Alberta, 1416 College Plaza Edmonton, T6G 2C8 Alberta, Canada.

## Summary

Uganda has a large population of goats, predominantly from indigenous breeds reared in diverse production systems, whose existence is threatened by crossbreeding with exotic Boer goats. Knowledge about the genetic characteristics and relationships among these Ugandan goat breeds and the potential admixture with Boer goats is still limited. Using a medium-density single nucleotide polymorphism (SNP) panel, we assessed the genetic diversity, population structure and admixture in six goat breeds in Uganda: Boer, Karamojong, Kigezi, Mubende, Small East African and Sebei. All the animals had genotypes for about 46 105 SNPs after quality control. We found high proportions of polymorphic SNPs ranging from 0.885 (Kigezi) to 0.928 (Sebei). The overall mean observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosity across breeds was  $0.355 \pm 0.147$  and  $0.384 \pm 0.143$  respectively. Principal components, genetic distances and ADMIXTURE analyses revealed weak population substructuring among the breeds. Principal components separated Kigezi and weakly Small East African from other indigenous goats. Sebei and Karamojong were tightly entangled together, whereas Mubende occupied a more central position with high admixture from all other local breeds. The Boer breed showed a unique cluster from the Ugandan indigenous goat breeds. The results reflect common ancestry but also some level of geographical differentiation. ADMIXTURE and  $f_4$  statistics revealed gene flow from Boer and varying levels of genetic admixture among the breeds. Generally, moderate to high levels of genetic variability were observed. Our findings provide useful insights into maintaining genetic diversity and designing appropriate breeding programs to exploit within-breed diversity and heterozygote advantage in crossbreeding schemes.

**Keywords** breed composition, breed diversity, *Capra hircus*, heterozygosity, indigenous goats, population genetics

## Introduction

According to archaeo-zoological evidence, goats were among the first ungulates to be domesticated, about 10 000 years ago near the fertile crescent that spans from Eastern Anatolia to the Zagros Mountains in Northern Iran (Zeder & Hesse 2000; Naderi *et al.* 2008). Archaeological evidence suggests the rapid spread of goats from the centre

of domestication to Eurasia and Africa following human migrations and trade routes. Migration of goats into Africa occurred through three main entry routes: one along the Mediterranean coast, a second via the Red Sea hills region and a third through the Nile Valley via the Sinai peninsula and the Nile delta (Taberlet *et al.* 2008; Pereira *et al.* 2009; Gifford-Gonzalez & Hanotte 2011). Other movements have also been reported from the Near East into the Ethiopian highlands and central Sahara (Clutton-Brock 2000).

Today, goats are among the most important livestock species in developing countries. They are of significant socio-economic, nutritional and cultural importance in smallholder farming systems. Uganda has three major indigenous goat breeds (Mubende, Kigezi and Small East African goats) that are geographically isolated and are

Address for correspondence

R. B. Onzima, Wageningen University & Research, Animal Breeding and Genomics, P O Box 338, 6700AH Wageningen, The Netherlands.  
E-mail: robertonzima@gmail.com

Accepted for publication 18 November 2017

raised in diverse production systems (Mason & Maule 1960; Nsubuga 1996; MAAIF & UBOS 2009). Besides these three indigenous breeds, several indistinct ecotypes of Ugandan indigenous goat breeds exist, including Karamojong and Sebei. In the early 1990s, crossbreeding with Boer goats introduced from South Africa was initiated to improve the production characteristics of the Ugandan indigenous goats (Nsubuga 1996). Boer goats are widely used as a source of breeding stock to cross with the indigenous goats (Onzima *et al.* 2014). The choice of Boer goats was premised on the fact that they have a fast growth rate and exhibit better disease resistance than do other exotic goat breeds (Casey & Van Niekerk 1988). However, with uncoordinated breeding management, indiscriminate crossing may occur, increasing the risk of the disappearance of resilient and well-adapted indigenous breeds. The existence of the various breeds presents an enormous source of diversity in the current goat populations that needs to be characterized, conserved and utilized in a sustainable manner under the existing production systems. Genetic diversity in populations is important, as it provides the basis for natural as well as artificial selection (Qanbari & Simianer 2014).

In order to study diversity, molecular tools are essential as a valuable complement to the evaluation of phenotypes and production systems and, sometimes, as a proxy for phenotypic diversity of local breeds (Ajmone-Marsan *et al.* 2014). However, compared to other livestock species, African goats remain poorly studied, especially at the molecular level. Earlier studies in Africa, using mitochondrial and microsatellite DNA markers, indicate a lack of phylogeographic structure among the goat breeds (Alemu 2004; Chenyambuga *et al.* 2004; Okpeku *et al.* 2011; Hassen *et al.* 2012; Benjelloun *et al.* 2015). These studies were geared mainly towards assessing genetic diversity in an attempt to monitor genetic erosion and to identify conservation priorities. In Uganda, earlier genetic characterization of indigenous goats was carried out using a limited number of microsatellites (Chenyambuga *et al.* 2004; Muema *et al.* 2009). A drawback of microsatellite analysis is that it is difficult to integrate data across laboratories, due mainly to the inherent poor reproducibility of allele calling (FAO 2011). Therefore, a comparison of results from different studies that used microsatellites is complicated.

However, the advent of the GoatSNP50 BeadChip in 2014 (Tosser-Klopp *et al.* 2014) has changed the landscape and depth of genomic research in goats (Tosser-Klopp 2015) due to its robustness, low genotyping costs, automatic allele calling and ability to interrogate the goat genome at high resolution (Ajmone-Marsan *et al.* 2014). The Illumina GoatSNP50 BeadChip, which features 53 347 single nucleotide polymorphisms (SNPs), was developed from SNP loci detected by whole genome sequencing of six goat breeds (Tosser-Klopp *et al.* 2014; Tosser-Klopp 2015). The SNP chip has been used to study genetic diversity and

population structure of goats in various countries with indigenous goat breeds in locales such as Italy (Nicoloso *et al.* 2015), Spain (Manunza *et al.* 2016), South Africa (Lashmar *et al.* 2016; Mdladla *et al.* 2016; Visser *et al.* 2016), Ethiopia (Abegaz 2014; Mekuriaw 2016) and Australia (Kijas *et al.* 2013). Therefore, the objective of the current study was to: (i) assess the degree of genetic diversity in Ugandan goat breeds using SNPs; (ii) infer population structure and breed relationships; and (iii) investigate admixture among breeds, namely the influence of the commercial Boer breed in Ugandan goats. The information generated from this study can be used in management and conservation of Ugandan goat genetic resources and makes it possible to design effective strategies for breed improvement.

## Materials and methods

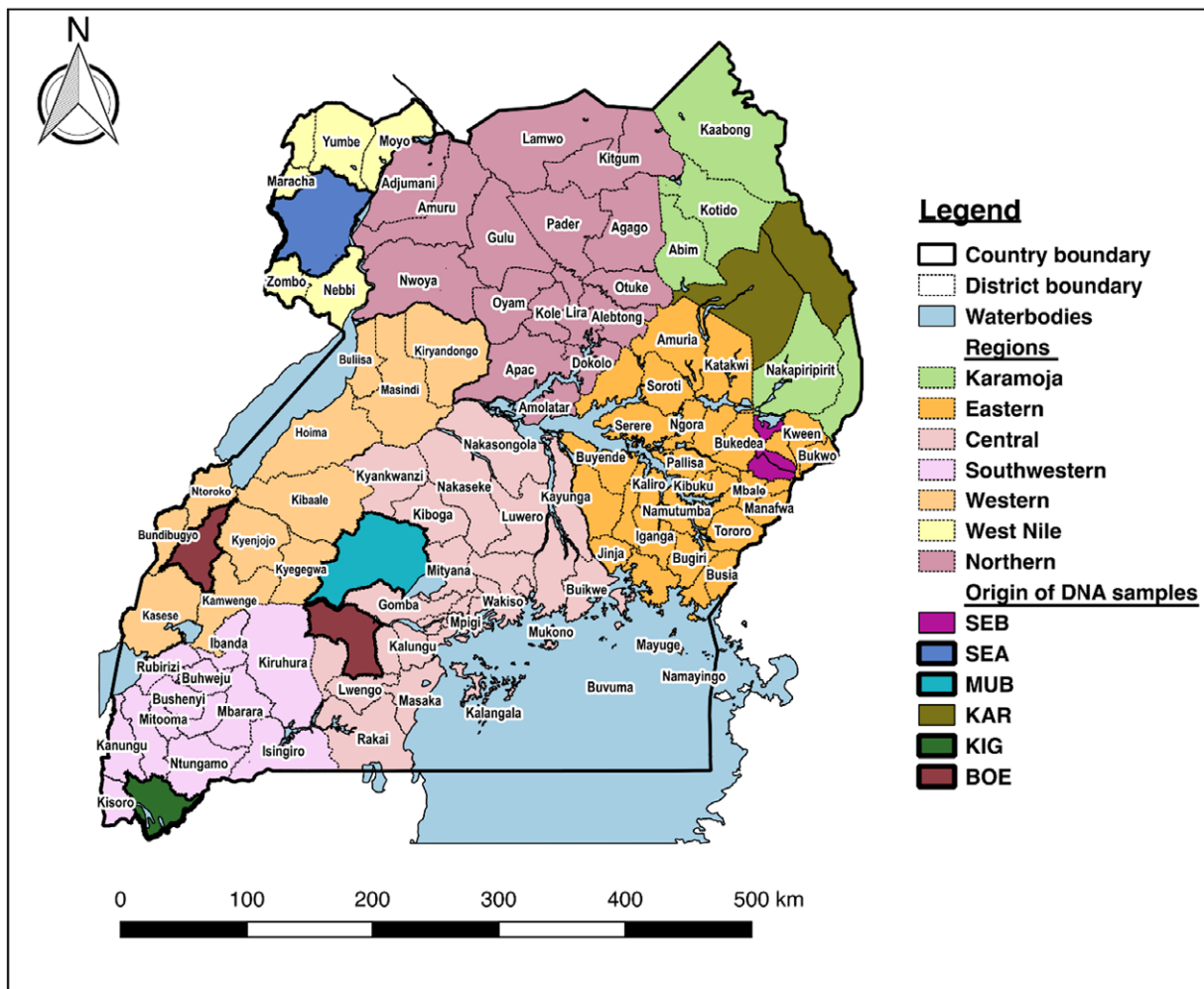
### Animal resources and sampling

A total of 144 animals from six goat breeds were included in this study. Five indigenous goat breeds (Mubende, Kigezi, Small East African, Karamojong and Sebei) were sampled from 79 smallholder farms/herds, and the exotic Boer goats were sampled from a commercial multiplication centre (Ssembeguya Estates) and a government breeding centre (Rubona Stock farm), which are sources of breeding stock for goat improvement in Uganda.

Sampling was carried out at selected geographical locations (Fig. 1) according to livestock statistics from the Livestock Census Report 2008 (MAAIF & UBOS 2009). The goat populations sampled originated from the following five agro-ecological zones of Uganda: Mubende goats from Mubende district in the mid-altitude farmlands and central wooded savanna, Kigezi from Kabale and Kisoro districts in the southwestern highlands, Small East African from Arua district in the short savanna grasslands, Karamojong from Moroto district in the northeastern semi-arid region, Sebei from Sironko in the eastern highlands and Boer goats from Ssembabule and Kabarole districts in the mid-altitude zone. Sampling was conducted to cover a wide distribution of individual animals across the selected production locations. Within a herd, we relied on the farmers' pedigree knowledge to select, as much as possible, unrelated individuals.

Ear punch tissue was collected from the 144 goats at smallholder farms for the indigenous breeds Mubende ( $n = 29$ ), Kigezi ( $n = 29$ ), Small East African ( $n = 29$ ), Karamojong ( $n = 15$ ) and Sebei ( $n = 29$ ) and at a commercial and a government breeding centre for Boer ( $n = 13$ ). The ear tissue samples were collected into vials containing a desiccant and stored within 12 h in a freezer at  $-4^{\circ}\text{C}$ .

The study was approved by the Ethics Committee of Uganda National Council of Science and Technology (UNCST; SBLS/REC/15/131).



**Figure 1** Map of Uganda showing geographical origin of the goat DNA samples analysed. Breed acronyms are defined as follows: BOE, Boer; KAR, Karamojong; KIG, Kigezi; MUB, Mubende; SEA, Small East African; SEB, Sebei.

### DNA extraction

Genomic DNA was extracted using the DNeasy blood and tissue kit (Qiagen®). Twenty DNA samples were randomly selected and analyzed on a 1% agarose gel for a preliminary estimate of the DNA quality and quantity. The final DNA quality and quantity were validated using the Qubit® dsDNA BR (Broad-Range) Assay Kit on the Qubit 2.0 fluorimeter (Invitrogen) prior to genotyping.

### Genotyping and quality control

DNA samples were genotyped with the Illumina GoatSNP50 BeadChip. The BeadChip, developed by the International Goat Genome Consortium (IGGC), features 53 347 SNPs across the whole goat genome with inter-SNP spacing of approximately 40 kb (Tosser-Klopp *et al.* 2014). Data were analyzed using GENOME STUDIO™ software v1.1 (Illumina, Inc.). Genomic locations of the SNPs and cluster files were provided by IGGC. Standard SNP genotype quality control

procedures were performed using PLINK v1.07 (Purcell *et al.* 2007). Individuals with a missing genotype call rate of greater than or equal to 10% were excluded from further analysis using the *-mind* function in PLINK. The remaining individuals were then subjected to SNP quality control. SNPs with a call rate of less than 0.95, minor allele frequency (MAF) of 0.05 or less and SNPs whose genotypes were not in Hardy Weinberg equilibrium ( $P < 0.001$ ) were excluded from downstream analysis. The dataset of SNPs used in the analysis is available from <https://www.animalgenome.org/repository/pub/WAGNL2017.1002/>.

### Data analysis

The observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosities for the respective populations were calculated using PLINK (Purcell *et al.* 2007). The population structuring and relatedness were estimated from the SNP genotypes using principal components analysis (PCA), available from the R package SNPRELATE (Zheng *et al.* 2012).

Additionally, population structure analyses were performed to infer the most likely number of ancestral populations using ADMIXTURE software version 1.23 (Alexander *et al.* 2009; Alexander & Lange 2011). To estimate the individual ancestry within the population, ADMIXTURE employs prior defined  $K$  values corresponding to the assumed number of ancestral populations. The procedure involves the use of maximum likelihood estimates on data from multiple loci to estimate individual ancestry within the population being considered. To determine the most optimal population structure, a cross-validation procedure was undertaken with hypothetical ADMIXTURE runs from  $K = 2$  to 7. Optimal partitioning of the population was achieved at the lowest cross-validation error.

Phylogenetic relationships between the goat breeds were inferred using the Neighbour-Net procedure in SPLITSTREE 4 software (Huson & Bryant 2006) based on Reynold's genetic distances, whereas individual relationships across all breeds were calculated using identity-by-state distances.

To further investigate admixture in Ugandan goat breeds, we performed a three-population ( $f_3$ ) test (Reich *et al.* 2009; Patterson *et al.* 2012) and a four-population ( $f_4$ ) test (Keinan *et al.* 2007; Patterson *et al.* 2012) implemented in TREEMIX (Pickrell & Pritchard 2012). These statistics are used to explain admixture history of the populations being investigated, particularly when correlations in allele frequencies do not conform to population evolution with a split tree (Reich *et al.* 2009; Patterson *et al.* 2012). To provide support for past admixture events between the populations, the THREEPOP program from TreeMix was used to calculate  $f_3$  (A;B,C) statistics for all possible combinations of three populations. Generally, if population A is a result of an admixture between two other populations B and C, the calculated  $z$ -score for each tested combination of three populations would have a significant negative value. A positive  $z$ -score may indicate either absence of admixture or substantial post-admixture drift resulting from the alleles in the population. Meanwhile, the FOURPOP program from TreeMix was used to calculate  $f_4$  (A;B; C,D) statistics for subsets of the population. The four-population test  $f_4$  (A;B; C,D) (Keinan *et al.* 2007; Patterson *et al.* 2012) was used to test if A,B and C,D were genetically distinct groups (clades) in the population tree. A significant

non-zero  $z$ -score indicates gene flow between A,B and C,D in the population tree (Keinan *et al.* 2007; Patterson *et al.* 2012; Makina *et al.* 2016). Larger values indicate strong evidence of gene flow in the tree.

## Results

### Level of SNP polymorphism within breeds

After quality control procedures on the 53 347 SNPs included in the SNP chip, 7242 SNPs were excluded (Table 1), which resulted in 46 105 loci available for downstream analysis. Of the SNPs excluded, 2093 showed a SNP call rate of less than 0.95, 3500 had MAFs less than 0.05 and 2817 significantly deviated from Hardy–Weinberg equilibrium ( $P < 0.001$ ). The highest number of SNPs that showed a MAF of less than 0.05 was found in Small East African ( $n = 7818$ ), whereas Sebei showed the lowest number of SNPs excluded ( $n = 6826$ ). All animals passed the quality criteria and were used in the analysis.

Boer (0.51) and Kigezi (0.43) showed the highest and lowest proportion of SNPs with MAF greater than or equal to 0.3 respectively (Fig. 2). The proportion of fixed loci (MAF = 0) was similar across the breeds, ranging from Karamojong (0.06) to Mubende (0.04). Despite only Boer being represented in the group of goat breeds used to develop this SNP array, 93.4% of the SNP markers across the six goat breeds was polymorphic (MAF  $\geq 0.05$ ) (Table 2). The highest proportion of polymorphic loci ( $P_N$ ) was found in Sebei (0.928) and the lowest in Kigezi (0.885); however, the differences in  $P_N$  were negligible across all the breeds.

### Breed genetic diversity

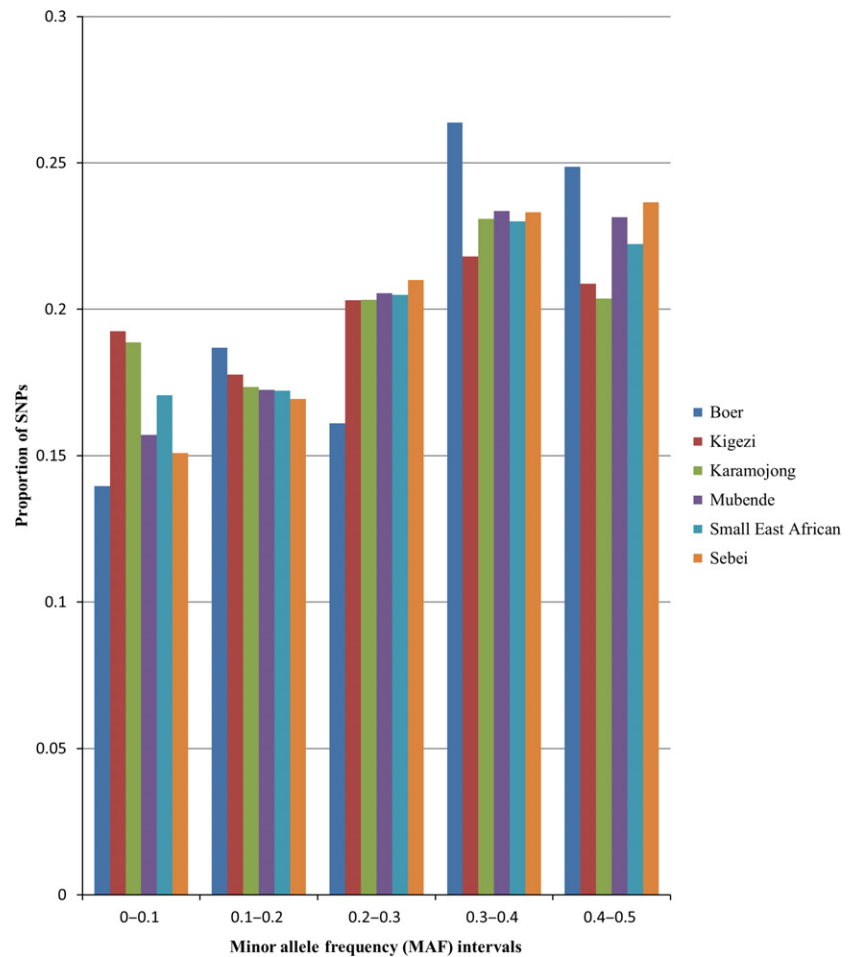
Genetic diversity was assessed within each breed (Table 2). The results indicate small differences in genetic diversity between the breeds. The lowest observed heterozygosity was found in Kigezi ( $H_O = 0.340 \pm 0.181$ ) and the highest in Boer ( $H_O = 0.377 \pm 0.193$ ), indicating higher diversity in Boer compared to Kigezi. Also, the MAFs across all loci were lowest in Kigezi (0.257) and highest in Boer (0.280). In general, the observed heterozygosity was slightly lower

Breed	$n$	Excluded SNPs <sup>1</sup>			Total	Remaining SNPs
		SNP CR < 0.95	MAF < 0.05	HWE		
Boer	13	2577	5280	208	7323	46 024
Karamojong	15	2640	5687	358	7803	45 544
Kigezi	29	1977	6139	543	7793	45 554
Mubende	29	2260	4922	767	7034	46 313
Small East African	29	2523	5669	589	7818	45 529
Sebei	29	2429	4824	469	6826	46 521
Merged	144	2093	3500	2817	7242	46 105

**Table 1** Number of animals and SNPs excluded and remaining after quality control procedures on genotype data.

$N$ , number of animals; CR, call rate; MAF, minor allele frequency; HWE, chi square test for Hardy–Weinberg equilibrium ( $P$ -value < 0.001).

<sup>1</sup>Some SNPs were excluded due to more than a single criterion.



**Figure 2** Distribution of SNPs by MAF intervals in each breed.

**Table 2** Population characteristics showing proportion of polymorphic SNPs ( $P_N$ ), mean minor allele frequency (MAF), expected ( $H_E$ ) and observed ( $H_O$ ) heterozygosity for the six goat populations.

Breed	$n$	$P_N$	MAF	$H_E \pm SD$	$H_O \pm SD$
Boer	13	0.901	0.280	$0.408 \pm 0.178$	$0.377 \pm 0.193$
Karamojong	15	0.893	0.271	$0.410 \pm 0.192$	$0.357 \pm 0.192$
Kigezi	29	0.885	0.257	$0.377 \pm 0.189$	$0.340 \pm 0.181$
Mubende	29	0.908	0.272	$0.391 \pm 0.179$	$0.355 \pm 0.178$
Small East African	29	0.894	0.266	$0.393 \pm 0.189$	$0.349 \pm 0.180$
Sebei	29	0.928	0.274	$0.395 \pm 0.178$	$0.365 \pm 0.176$
Merged	144	0.934	0.289	$0.384 \pm 0.143$	$0.355 \pm 0.147$

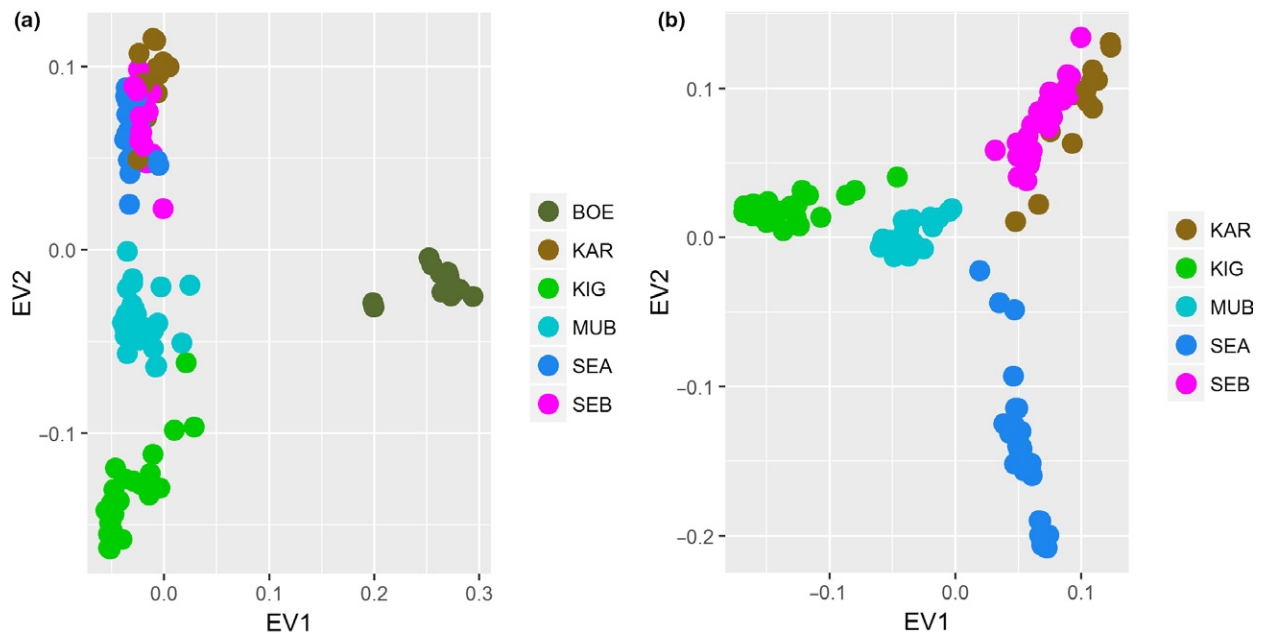
than the expected heterozygosity ( $H_O < H_E$ ), indicating a deficiency in heterozygosity across all the breeds.

### Population structure analysis

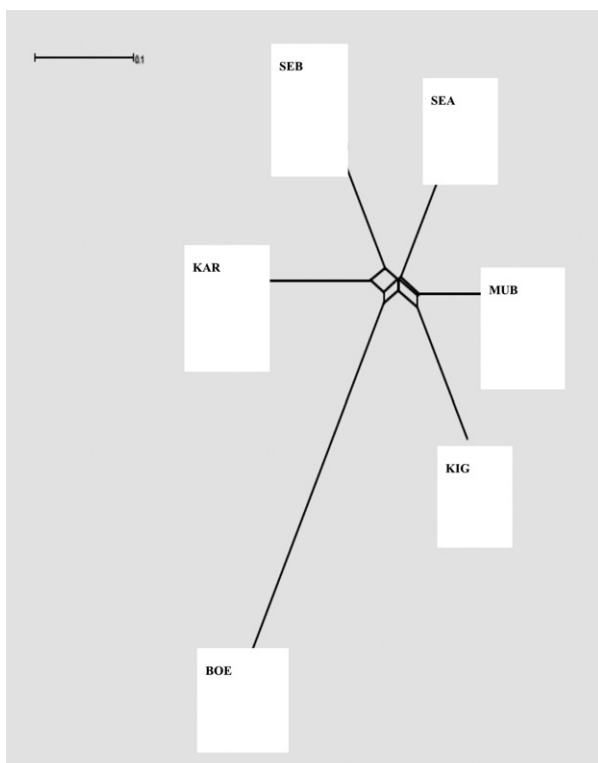
The first principal component, eigen vector 1 (EV1) shown in Fig. 3a, separated Boer from the Ugandan indigenous goat breeds and accounted for 10.7% of the total variance. The second principal component (EV2) accounted for 3.2% of the total variance and divided the Ugandan indigenous

goat breeds into two clusters: a distinct cluster comprising Kigezi and Mubende breeds and a combined breed cluster consisting of Sebei, Karamojong and Small East African breeds (Fig. 3). A more detailed analysis, in which the first principal component (EV1) explained 3.6% and the second principal component (EV2) accounted for 2.9% of total variance in the Ugandan indigenous goat breeds, showed a similar clustering pattern for all breeds except the Small East African goats, which formed a separate cluster (Fig. 3b). Mubende clustered among all other breeds, indicating possible admixture with the other goats.

Breed relationships were also assessed by computing genetic distances between each pair of individuals from the number of loci for which they differ. Based on the estimated genetic distances, a Neighbour-Net graph was computed to depict breed clustering (Fig. 4). The Ugandan indigenous breeds showed short branching, suggesting low differentiation between the breeds, whereas the exotic Boer goat breed showed a long branch, suggesting a well-differentiated and distinctive clade. Individuals belonging to the same breed mostly clustered together, as inferred by the identity-by-state distance-based neighbour-joining (NJ) tree (Fig. S1). Some Sebei and Karamojong individuals appeared



**Figure 3** Plot of two principal components showing genetic relationships among: (a) five Ugandan indigenous goats and Boer breeds and (b) five Ugandan indigenous goat breeds only. Goat populations analysed: BOE, Boer; KAR, Karamojong; KIG, Kigezi; MUB, Mubende; SEA, Small East African; SEB, Sebei.



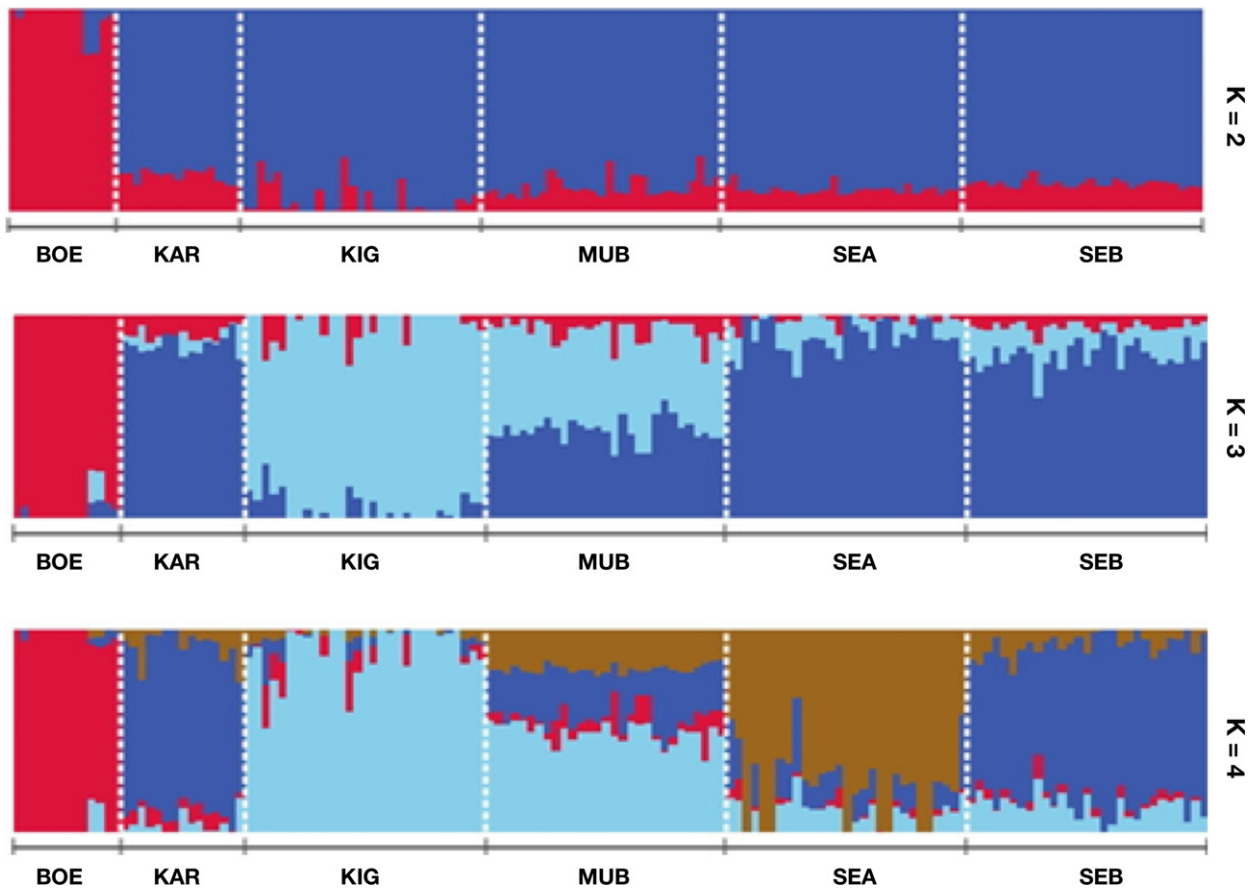
**Figure 4** Neighbour-Net graph based on Reynolds genetic distances depicting breed relationships among five Ugandan indigenous goat populations and one commercial goat breed. BO, Boer; KAR, Karamojong; KIG, Kigezi; MUB, Mubende; SEA, Small East African; SEB, Sebei.

to be entangled (or admixed), whereas Mubende was subdivided into two groups. The remaining breeds (i.e. Small East African, Kigezi and Boer) all formed tight groups.

#### Genetic admixture

Similar to the results from PCA, *ADMIXTURE* analysis at  $K = 2$  separated the Ugandan indigenous goats from the commercial Boer goats (Fig. 5). Additionally, at  $K = 3$  the analysis separated the populations into three subpopulations: Boer; Kigezi and Mubende; and Small East African, Karamojong and Sebei; it also indicated a considerable component of Boer and Kigezi in the subpopulations. Based on the least cross-validation error (Fig. S2),  $K = 4$  was identified as the optimal number of ancestral populations and indicated a Boer component in all five Ugandan indigenous goat breeds. On average, around 3, 5, 5, 1 and 1% of the Boer goat genome was shared with Karamojong, Kigezi, Mubende, Small East African and Sebei goats respectively (Table 3). The analysis also revealed a finer resolution of the Ugandan indigenous breeds, by which Kigezi and Small East African goats emerged as a distinct groups, Karamojong and Sebei remained tightly clustered together and the Mubende breed appeared to be the more admixed population, comprising Kigezi, Small East African and Karamojong/Sebei breeds (50%, 20% and 25% respectively). *ADMIXTURE* runs from  $K = 2$  to 4 revealed considerable admixture among the breeds, and gene flow from Boer was observed in all the indigenous Ugandan goat breeds.

To further confirm admixture among the goat breeds, we calculated  $f_3$  statistics for all possible three-population



**Figure 5** Population structure plots showing proportions of ancestral populations for each individual for  $K = 2$  to 4. BOE, Boer; KAR, Karamojong; KIG, Kigezi; MUB, Mubende; SEA, Small East African; SEB, Sebei.

**Table 3** Average breed composition of six goat populations given four clusters estimated by ADMIXTURE software.

Breed	$n$	Cluster 1 (SEA)	Cluster 2 (KAR/SEB)	Cluster 3 (BOE)	Cluster 4 (KIG)
Boer	13	$0.008 \pm 0.017$	$0.018 \pm 0.027$	$0.950 \pm 0.084$	$0.024 \pm 0.055$
Karamojong	15	$0.091 \pm 0.079$	$0.840 \pm 0.120$	$0.031 \pm 0.025$	$0.038 \pm 0.056$
Kigezi	29	$0.022 \pm 0.027$	$0.027 \pm 0.046$	$0.045 \pm 0.070$	$0.905 \pm 0.129$
Mubende	29	$0.202 \pm 0.022$	$0.245 \pm 0.066$	$0.049 \pm 0.053$	$0.503 \pm 0.055$
Small East African	29	$0.776 \pm 0.178$	$0.123 \pm 0.107$	$0.009 \pm 0.023$	$0.091 \pm 0.070$
Sebei	29	$0.077 \pm 0.059$	$0.784 \pm 0.121$	$0.012 \pm 0.021$	$0.127 \pm 0.058$

BOE, Boer; KAR, Karamojong; KIG, Kigezi; MUB, Mubende; SEA, Small East African; SEB, Sebei goats.

groups for all six breeds and  $f_4$  statistics for all possible subpopulations for sister and opposing sister groups. With the Ugandan goats in this study, we found only four significant  $f_3$  tests, all of which involved the Mubende breed (Table 4). This suggests admixture between Mubende and the other goat breeds.

Based on the  $f_4$  test statistic, combining Boer with any of the Ugandan indigenous goats resulted in the most significant values (Table 5). This suggests gene flow from Boer into Ugandan indigenous breeds. Similarly, the significant  $f_4$  statistics for subpopulations involving Karamojong and

Kigezi goats with the other Ugandan indigenous goats suggest gene flow from these breeds (Table S1).

## Discussion

In this study, we assessed genetic diversity, population structure and admixture in Ugandan indigenous goat breeds at a genome-wide scale using a moderately dense SNP panel. We further assessed the presence of admixture of an exotic goat breed (Boer) into the Ugandan breeds.

Population A (admixed)	Population B (source)	Population C (source)	$f_3$ statistic	SE	z-score
MUB	BOE	KIG	-0.0017	0.0002	-7.1891*
MUB	KIG	SEB	-0.0012	0.0001	-13.7787*
MUB	KAR	KIG	-0.0016	0.0001	-16.0069*
MUB	KIG	SEA	-0.0016	0.0001	-19.4077*

BOE, Boer; KAR, Karamojong; KIG, Kigezi; MUB, Mubende; SEA, Small East African; SEB, Sebei goats.

\*Significant  $f_3$  statistics ( $P < 0.05$ ).

### Genotypic data and level of polymorphism

The first objective of the study was to assess the level of polymorphism in Ugandan goat breeds using the GoatSNP50 BeadChip. Our results show that the proportion of polymorphic loci within Ugandan goat breeds ranges from 0.885 in Kigezi goats to 0.928 in Sebei (Table 2). This high level of genetic polymorphism indicates that most of the SNPs are segregating in the breeds under investigation. The level of polymorphism observed in this study is similar to those observed during SNP discovery and validation within breeds with similar numbers of animals genotyped (Tosser-Klopp *et al.* 2014). The GoatSNP50 BeadChip was developed using dairy and mixed breeds (Alpine, Saanen and Creole) and meat-type breeds (Boer, Katjang and Savanna). The chip was validated with 10 breeds from different backgrounds. The SNPs were segregating at greater than 78% in seven of the breeds, including Angora and Skopelos, which were not used during SNP discovery (Tosser-Klopp *et al.* 2014). Similar levels of polymorphism

in goat breeds have been reported elsewhere (Kijas *et al.* 2013; Nicoloso *et al.* 2015; Lashmar *et al.* 2016; Mdladla *et al.* 2016; Mekuriaw 2016). For example, Mdladla *et al.* (2016) reported levels of polymorphism ranging from 84.2 to 97.6% in nine South African indigenous goats, 96.8 to 99.7% in Italian goats (Nicoloso *et al.* 2015) and greater than 97% in Australian goat breeds (Kijas *et al.* 2013). The success of the chip can be attributed to the use of six goat breeds from different types, origins and production environments for SNP discovery. Therefore, similar high levels of polymorphism were envisaged across other breeds that were not used during the design of the SNP chip.

### Breed genetic diversity

The Ugandan indigenous goat and Boer breeds show a high degree of genetic diversity, as determined by the high heterozygosity values detected in this study. Our results revealed that the expected and observed heterozygosities

Population 1	Population 2	Population 3	Population 4	$f_4$ statistic	SE	z-score
BOE	KIG	KAR	MUB	0.0050	0.0002	20.5959*
BOE	MUB	KAR	KIG	0.0051	0.0003	20.4125*
BOE	KAR	MUB	SEB	0.0034	0.0002	16.4090*
BOE	KIG	KAR	SEB	0.0022	0.0002	14.3987*
BOE	KAR	KIG	SEB	0.0033	0.0003	12.5043*
BOE	MUB	KAR	SEB	0.0018	0.0002	12.1647*
BOE	MUB	KAR	SEA	0.0027	0.0003	10.8053*
BOE	KIG	KAR	SEA	0.0027	0.0003	10.2101*
BOE	KAR	SEA	SEB	0.0020	0.0002	9.2824*
BOE	SEA	KAR	SEB	0.0014	0.0002	8.2888*
BOE	KAR	MUB	SEA	0.0014	0.0002	6.0396*
BOE	SEA	KAR	MUB	0.0013	0.0002	5.6349*
BOE	SEA	KAR	KIG	0.0014	0.0003	5.0805*
BOE	SEB	MUB	SEA	0.0010	0.0002	4.4690*
BOE	KAR	KIG	SEA	0.0013	0.0003	4.4078*
BOE	SEB	KAR	KIG	-0.0011	0.0003	-4.0038*
BOE	MUB	SEA	SEB	-0.0009	0.0002	-4.1880*
BOE	SEB	KAR	MUB	-0.0016	0.0002	-6.4749*
BOE	KIG	MUB	SEA	-0.0023	0.0002	-9.8169*
BOE	MUB	KIG	SEA	-0.0024	0.0003	-9.4555*
BOE	KIG	MUB	SEB	-0.0028	0.0002	-12.9529*
BOE	MUB	KIG	SEB	-0.0033	0.0002	-14.3171*

BOE, Boer; KAR, Karamojong; KIG, Kigezi; MUB, Mubende; SEA, Small East African; SEB, Sebei goats.

\*Significant  $f_4$  statistics indicating presence of gene flow.

**Table 4** Summary of three-population tests with significant  $f_3$  statistics showing admixture in Ugandan goats with Kigezi as one of the source populations.

**Table 5** Summary of four-population tests showing some significant  $f_4$  statistics to detect admixture and gene flow within Ugandan indigenous and Boer goat breeds.



ranged from  $H_E = 0.377 \pm 0.189$  (Kigezi) to  $H_E = 0.410 \pm 0.192$  (Karamojong) and  $H_O = 0.340 \pm 0.181$  (Kigezi) to  $H_O = 0.377 \pm 0.193$  (Boer) respectively (Table 2). The heterozygosity values obtained in this study are comparable with those reported for indigenous goats in Ethiopia (Mekuriaaw 2016), South Africa (Mdladla *et al.* 2016), Egypt (Kim *et al.* 2016), Spain (Manunza *et al.* 2016) and Italy (Nicoloso *et al.* 2015) as well as for commercial goats from Canada and Australia (Brito *et al.* 2017); South Africa (Lashmar *et al.* 2016); and a variety of Angora goat populations in Argentina, France and South Africa (Visser *et al.* 2016).

The expected and observed heterozygosity for Boer ( $H_E = 0.408 \pm 0.178$ ;  $H_O = 0.377 \pm 0.193$  respectively) were slightly higher than those reported for Boer populations in Canada ( $H_E = 0.357$ ;  $H_O = 0.363$  respectively) (Brito *et al.* 2017) and Australia ( $H_E = 0.355$ ;  $H_O = 0.363$  respectively) (Kijas *et al.* 2013). These differences may be attributed to differences in effective population sizes, duration of isolation and selection practices in the different production systems.

We obtained the highest expected heterozygosity in Karamojong goats, which could be due to the pastoral production system used. Under communal production systems practiced by pastoral and smallholder farmers, there is an absence of structured artificial selection programs, with random mating and high admixing between populations and herds probably occurring. This favours an increase in genetic variability and reduction in inbreeding, which is a decisive factor in the success for conservation programs. A similar trend of genetic diversity was also reported for indigenous goat breeds in South Africa (Mdladla *et al.* 2016) and local goat breeds in Brazil (da Rocha *et al.* 2016). Similarly, a study investigating breed composition of Creole goats from 10 American countries found moderate to high heterozygosity values (Ginja *et al.* 2017).

The difference between the observed and expected heterozygosity was small and within a fraction of one standard error. As a general trend, the observed heterozygosity was lower than the expected heterozygosity ( $H_O < H_E$ ) within all breeds. Thus, these differences may be due to a Wahlund effect rather than inbreeding.

### Population structure and admixture

The results of the population structure and admixture analyses indicate that the five Ugandan goat breeds are weakly differentiated. This may be due to the recent establishment of these breeds from probably the same founder population or related populations, but to confirm this, an in-depth analysis of the breed history will be required. Another possible explanation for the low degree of differentiation could be continuous gene flow between the indigenous breeds.

The results of our population structure and admixture analyses are in agreement. The three methods were used

separate the breeds into clusters. The first principal component separates the Boer from the Ugandan indigenous goat breeds (Fig. 3). Furthermore, the presence of a distinct branching of Boer in the NJ tree suggests a differentiated gene pool (Fig. 4).

At  $K = 2$ , the ADMIXTURE analysis separates the breeds into two distinct clusters: the Boer and the Ugandan indigenous goats (Fig. 5). This observation is in agreement with the PCA results, which showed the same two major clusters. Optimal clustering is observed at  $K = 4$ , a value at which the cross-validation error is lowest (Fig. S2). At this optimal  $K$  value, ADMIXTURE analysis differentiates Ugandan indigenous goat breeds into distinct clusters of Small East African/Kigezi and Karamojong/Sebei tightly grouped together, whereas Mubende is admixed showing influences from the three groups. The differentiation of Kigezi and Small East African goats could be attributed to genetic drift but could also have resulted from selection and/or adaptation pressures.

The three- and four-population tests (Keinan *et al.* 2007; Reich *et al.* 2009; Patterson *et al.* 2012) were used to further qualify admixture and gene flow. Using the three-population ( $f_3$ ) test statistic, we found strong evidence of admixture in only four comparisons, all involving the Mubende goats as the admixed breed. The most significant  $z$ -score ( $-19.408$ ) was found with the Mubende  $f_3$  (Mubende/Kigezi/Small East African). All the significant  $f_3$  statistics were observed when Kigezi was one of the source populations, indicating that Kigezi might be contributing to the gene pool either through ancestral generations or crossbreeding with Mubende. The  $f_4$  statistics showed the most significant scores for the Boer and the Ugandan indigenous goats (Table S1). This is supported by the results of the ADMIXTURE analysis, as shown by the breed composition of the individual goats studied (Table S2). However, determining the extent of the admixture in the Ugandan goat populations requires further studies involving larger sample sizes and more ecotypes and breeds.

The indigenous Ugandan goat breeds showed clear differentiation according to their geographical regions. The results show a clear differentiation of Kigezi and Small East Africa, whereas Karamojong and Sebei remained tightly clustered together, which may be attributed to the contiguous territory of the breeds (Fig 1). Mubende is centrally located and prone to admixture with Kigezi, Karamojong/Sebei and Small East African. The lack of differentiation in some of the indigenous breeds is also confirmed by the presence of short branches in the NJ tree (Fig. 4), suggesting a high level of genetic similarity and low divergence, which may be attributed to local admixture between the breeds. The lack of differentiation in geographically diverse populations may also indicate common ancestry, short domestication history and lack of selection pressure, and the mobility of the goats may also play a role. The Boer goat breed was introduced into Uganda for

improving growth characteristics of the indigenous goats for meat production. The breed originates from South Africa, where it has been extensively selected for faster growth (Casey & Van Niekerk 1988).

Our results support admixture among all the breeds, although the highest admixture was observed in Mubende goats. This may be due to the fact that the breed is widely used across production systems as a preferred breed by the farmers due to its large live body weight (Onzima *et al.* 2016, 2017). ADMIXTURE analysis further revealed admixing between the Boer and the indigenous goats; however, the results suggest there is limited gene flow from Boer to the Ugandan indigenous goat populations due to crossbreeding.

## Conclusion

Overall, the results described in this study indicate high genetic variability of the Ugandan goat populations and sufficient genetic potential for further improvement of the breeds for heritable economic traits. The Ugandan indigenous goats are weakly differentiated, consisting of two breeds forming more uniform clusters (Kigezi and Small East African), two breeds clearly crossbred (Karamojong and Sebei) and Mubende showing signs of gene flow from all these goat populations. Nonetheless, there is rather limited Boer admixture in the Ugandan goat population. This knowledge can be exploited to devise strategies for sustainable utilization and maintenance of genetic diversity.

## Acknowledgements

We would like to acknowledge Bert Dibbits from Wageningen University & Research, Animal Breeding and Genomics for performing the quality and quantity measurements on the genomic DNA samples. We acknowledge Paul Kashaija from the National Livestock Resources Research Institute and Benda Katali from Kachwekano Zonal Agricultural Research and Development Institute for support in DNA extraction and field sampling respectively. We are grateful for funding from National Agricultural Research Organization (NARO) through a World Bank-supported project, Agricultural Technology and Agribusiness Advisory Services (ATAAS) (P109224). The smallholder farmers who allowed their animals to be sampled for the study are greatly honoured.

## Conflict of interest

The authors declare that they have no conflict of interest whatsoever to influence the outcome of the manuscript.

## Authors' contributions

RBO conceived and designed the study, collected the field samples and wrote the manuscript. MAMG, EK and RPMAC

supervised the study. MRU, RBO and RM participated in data analysis. All authors have read and edited the manuscript.

## References

- Abegaz S.G. (2014) *Design of community based breeding programs for two indigenous goat breeds in Ethiopia*. PhD Thesis, BOKU-University of Natural Resources and Life Sciences, Vienna, Austria.
- Ajmone-Marsan P., Colli L., Han J.L. *et al.* (2014) The characterization of goat genetic diversity: towards a genomic approach. *Small Ruminant Research* **121**, 58–72.
- Alemu T. (2004) *Genetic characterization of indigenous goat populations of Ethiopia using microsatellite DNA markers*. National Dairy Research Institute, Deemed University, Haryana, India.
- Alexander D.H. & Lange K. (2011) Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* **12**, 246.
- Alexander D.H., Novembre J. & Lange K. (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* **19**, 1655–64.
- Benjelloun B., Alberto F.J., Streeter I. *et al.* (2015) Characterizing neutral genomic diversity and selection signatures in indigenous populations of Moroccan goats (*Capra hircus*) using WGS data. *Frontiers in Genetics* **6**, 107.
- Brito L.F., Kijas J.W., Ventura R.V., Sargolzaei M., Porto-Neto L.R., Cánovas A., Feng Z., Jafarikia M. & Schenkel F.S. (2017) Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers. *BMC Genomics* **18**, 229.
- Casey N.H. & Van Niekerk W.A. (1988) The Boer goat I. Origin, adaptability, performance testing, reproduction and milk production. *Small Ruminant Research* **1**, 291–302.
- Chenyambuga S.W., Hanotte O., Hirbo J., Watts P.C., Kemp S.J., Kifaro G.C., Gwakisa P.S., Petersen P.H. & 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**, 445–52.
- Clutton-Brock J. (2000) Cattle, sheep, and goats south of the Sahara: an archaeozoological perspective. In: *The Origins and Development of African Livestock-Archaeology, Genetics, Linguistics and Ethnography* (Ed. by R.M. Blench & K.C. MacDonald), pp. 30–7. UCL Press, London, UK.
- FAO (2011) *Molecular Genetic Characterization of Animal Genetic Resources*. FAO Animal Health Production and Health Guidelines, No. 9. FAO, Rome.
- Gifford-Gonzalez D. & Hanotte O. (2011) Domesticating animals in africa: implications of genetic and archaeological findings. *Journal of World Prehistory* **24**, 1–23.
- Ginja C., Gama L.T., Martínez A. *et al.* (2017) Genetic diversity and patterns of population structure in Creole goats from the Americas. *Animal Genetics* **48**, 315–29.
- Hassen H., Lababidi S., Rischkowsky B., Baum M. & Tibbo M. (2012) Molecular characterization of Ethiopian indigenous goat populations. *Tropical Animal Health and Production* **44**, 1239–46.
- Huson D.H. & Bryant D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**, 254–67.

- Keinan A., Mullikin J.C., Patterson N. & Reich D. (2007) Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans. *Nature Genetics* **39**, 1251–5.
- Kijas J.W., Ortiz J.S., McCulloch R., James A., Brice B., Swain B., Tosser-Klopp G. & the International Goat Genome C. (2013) Genetic diversity and investigation of polledness in divergent goat populations using 52 088 SNPs. *Animal Genetics* **44**, 325–35.
- Kim E.S., Elbeltagy A.R., Aboul-Naga A.M., Rischkowsky B., Sayre B., Mwacharo J.M. & Rothschild M.F. (2016) Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity* **116**, 255–64.
- Lashmar S.F., Visser C. & Marle-Köster E.V. (2016) SNP-based genetic diversity of South African commercial dairy and fiber goat breeds. *Small Ruminant Research* **136**, 65–71.
- MAAIF & UBOS (2009) *The National Livestock Census Report 2008*. Ministry of Agriculture, Animal Industry and Fisheries (MAAIF), Entebbe/Uganda Bureau of Statistics, Kampala, Uganda.
- Makina S.O., Whitacre L.K., Decker J.E., Taylor J.F., MacNeil M.D., Scholtz M.M., van Marle-Köster E., Muchadeyi F.C., Makgahlela M.L. & Maiwashe A. (2016) Insight into the genetic composition of South African Sanga cattle using SNP data from cattle breeds worldwide. *Genetics Selection. Evolution* **48**, 88.
- Manunza A., Noce A., Serradilla J.M. *et al.* (2016) A genome-wide perspective about the diversity and demographic history of seven Spanish goat breeds. *Genetics Selection Evolution* **48**, 52.
- Mason I.L. & Maule J.P. (1960) *The Indigenous Livestock of Eastern and Southern Africa*. Commonwealth Bureau of Animal Breeding and Genetics. CAB (Commonwealth Agricultural Bureaux), Oxfordshire, UK.
- Mdladla K., Dzomba E.F., Huson H.J. & Muchadeyi F.C. (2016) Population genomic structure and linkage disequilibrium analysis of South African goat breeds using genome-wide SNP data. *Animal Genetics* **47**, 471–82.
- 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 University, Addis Ababa, Ethiopia.
- Muema E., Wakhungu J., Hanotte O. & Jianlin H. (2009) Genetic diversity and relationship of indigenous goats of Sub-Saharan Africa using microsatellite DNA markers. *Livestock Research for Rural Development* **21**, 28.
- Naderi S., Rezaei H.R., Pompanon F., Blum M.G.B., Negrini R. & Naghash H.R. (2008) The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 17659–64.
- Nicoloso L., Bomba L., Colli L. *et al.* (2015) Genetic diversity of Italian goat breeds assessed with a medium-density SNP chip. *Genetics Selection Evolution* **47**, 1–10.
- Nsubuga H.S.K. (1996) Small ruminant: goats and sheep in Uganda. In: *Small Ruminant Research and Development in Africa: Proceedings of the Third Biennial Conference of the African Small Ruminant Research Network*, p. 326. UICC, Kampala, Uganda, 5–9 December, 1994.
- Okpeku M., Peters S.O., Ozoje M.O., Adebambo O.A., Agaviezor B.O., O'Neill M.J. & Imumorin I.G. (2011) Preliminary analysis of microsatellite-based genetic diversity of goats in southern Nigeria. *Animal Genetic Resources* **49**, 33–41.
- Onzima R.B., Aheisibwe A.R., Benda K.K., Kanis E. & van Arendonk J.A.M. (2014) Economic analysis of cross breeding programs for indigenous goat breeds in Uganda. In: *10th World Congress on Genetics Applied to Livestock Production (WCGALP)*. Vancouver, BC, Canada.
- Onzima R.B., Mukibi R., Benda K.K., vanArendonk J.A.M., Komen H. & Kanis E. (2016) Participatory definition of breeding objectives and trait preferences by smallholder farmers for indigenous goats in Uganda. In: *Book of Abstracts of the 12th International Conference on Goats 'ICG 2016'*, Antalya, Turkey.
- Onzima R.B., Gizaw S., Kugonza D.R., van Arendonk J.A.M. & Kanis E. (2017) Production system and participatory identification of breeding objective traits for indigenous goat breeds of Uganda. *Small Ruminant Research*. <http://dx.doi.org/10.1016/j.smallrumres.2017.07.007>
- Patterson N., Moorjani P., Luo Y., Mallick S., Rohland N. & Zhan Y. (2012) Ancient admixture in human history. *Genetics* **192**, 1065–93.
- Pereira F., Queirós S., Gusmão L., Nijman I.J., Cuppen E. & Lenstra J.A. (2009) Tracing the history of goat pastoralism: new clues from mitochondrial and Y chromosome DNA in North Africa. *Molecular Biology and Evolution* **26**, 2765–73.
- Pickrell J.K. & Pritchard J.K. (2012) Inference of population splits and mixtures from genome-wide allele frequency data. *PLOS Genetics* **8**, e1002967.
- Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., Bender D., Maller J., Sklar P., Bakker P.I. & Daly M.J. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–75.
- Qanbari S. & Simianer H. (2014) Mapping signatures of positive selection in the genome of livestock. *Livestock Science* **166**, 133–43.
- Reich D., Thangaraj K., Patterson N., Price A.L. & Singh L. (2009) Reconstructing Indian population history. *Nature* **461**, 489–94.
- da Rocha L.L., Filho E.C.P., Filho M.A.G., Delgado J.V., Martínez A.M. & Ribeiro M.N. (2016) Impact of foreign goat breeds on the genetic structure of Brazilian indigenous goats and consequences to intra-breed genetic diversity. *Small Ruminant Research* **134**, 28–33.
- Taberlet P., Valentini A., Rezaei H.R., Naderi S., Pompanon F. & Negrini R. (2008) Are cattle, sheep, and goats endangered species? *Molecular Ecology* **17**, 275–84.
- Tosser-Klopp G. (2015) Update on the International Goat Genome Consortium Projects. In: *Plant and Animal Genome XXIII Conference*. San Diego, CA.
- Tosser-Klopp G., Bardou P., Bouchez O. *et al.* (2014) Design and characterization of a 52K SNP Chip for goats. *PLoS One* **9**, e86227.
- Visser C., Lashmar S.F., Van Marle-Köster E., Poli M.A. & Allain D. (2016) Genetic diversity and population structure in South African, French and Argentinian Angora goats from genome-wide SNP data. *PLoS One* **11**, e0154353.
- Zeder M.A. & Hesse B. (2000) The initial domestication of goats (*Capra hircus*) in the Zagros Mountains 10,000 years ago. *Science* **287**, 2254–7.
- Zheng X., Levine D., Shen J., Gogarten S.M., Laurie C. & Weir B.S. (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–8.

### Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

**Figure S1** Genetic relationships among five Ugandan indigenous goat breeds and one commercial goat breed; constructed using a neighbour-joining tree from identity-by-state (IBS) distances derived from 46 105 SNPs.

**Figure S2** Cross-validation error plot indicating the choice of the appropriate  $K$  value.

**Table S1** Significant  $f_4$  statistics for Ugandan goat breeds indicating gene flow in the breeds.

**Table S2** Breed composition of cluster 1, 2, 3 and 4, interpreted as representative of Small East African, Karamojong/Sebei, Boer and Kigezi ancestries respectively, estimated for each individual belonging to goat populations from Uganda.