Genetic characterization of Mozambican Nguni cattle and their relationship with indigenous populations of South Africa

F.J.M. King[1,2], C. Visser[2], C. Banga[3]

[1]Agricultural Research Institute of Mozambique, Directorate of Animal Science, P.O.Box 1410, Mozambique

[2]University of Pretoria, Department of Animal Sciences, Private Bag X20, Hatfield, 0028, South Africa

[3]Agricultural Research Council, Private Bag X2, Irene 0062, South Africa

Author's email address for correspondence: fejomak@gmail.com

Highlights

- South African and Mozambican indigenous beef cattle are poorly differentiated.
- All populations included show moderate genetic diversity and low levels of inbreeding.
- Mozambican Nguni and South African Nguni cattle share similar genetic ancestry.
- *KCNMB2* and *MYLK3* were identified using both F_{ST} and *Rsb* analyses, and indicate superior reproductive efficiency.

ABSTRACT

Knowledge of genetic variability among cattle populations is essential to gain insight into the adaptation mechanisms to different environments and to support the conservation of genetic resources. Individuals from Mozambican Nguni (MZ Nguni; n = 119), South African Nguni (SA Nguni; n = 150), South African Tuli (SA Tuli; n = 150), and South African Boran (SA Boran; n = 150) cattle populations were genotyped using the International Dairy and Beef SNP BeadChip version three (IDB) and the GeneSeek Genomic Profiler (GGP 80k) assays, to investigate their levels of genetic diversity and the relationships between these indigenous breeds. Levels of genetic diversity, assessed by expected heterozygosity (H_e), varied from 0.284 (SA Boran) to 0.324 (SA Tuli). Population structure, as well as principal component analysis (PCA), revealed tight clustering of the two Nguni populations, while the SA Tuli and SA Boran populations diverged, as expected, into two distinct clusters. Little genetic distance (0.031) was observed among MZ Nguni and SA Nguni, while SA Boran (a Zebu breed) was further removed from SA Tuli than from the other Sanga cattle populations. Runs of homozygosity (ROH) analysis revealed low inbreeding rates (with the average F_{ROH} per population ranging from 0.003 to 0.006). Short ROH segments (ROH ≤ 5 Mb) were more frequent

in all four populations than longer segments, suggesting more ancient inbreeding in these populations. The highest number of ROH (303) was observed in SA Tuli, while the lowest (56) was detected in SA Nguni. Analysis of both Wright's fixation index (F_{ST}) and ratio of extended haplotype homozygosity (*Rsb*) identified a total of 229 differentiated single nucleotide polymorphisms (SNP) to be under selection, in a comparison between the MZ Nguni cattle and South African cattle populations. Highly differentiated SNP ($F_{ST} \geq 0.26$ or $pRsb \geq 3$) indicated genes including *KLHL29*, *ZEB2*, *LAMC1*, *MYLK3*, and *KCNK5* that are implicated in several metabolic processes essential for adaptation and production traits.

Keywords: Boran, Admixture, SNPs, Dendrogram, Selection Signatures, Tuli

1. Introduction

Livestock is an important contributor to rural livelihoods in Southern Africa (Nyamushamba *et al*., 2017). Cattle provide a wide range of benefits including economic security through income from the sale of hides, milk, and meat (i.e. food security) and are commonly used in festivities and cultural rituals such as marriage and initiation ceremonies (Ibeagha-Awemu *et al*., 2019).

Archaeological and molecular evidence indicate that all modern breeds of cattle (indicine and taurine) originated from two subspecies of aurochs in two independent domestication events (Mwai *et al.,* 2015). Taurine cattle (*Bos taurus*) are presumed to have been domesticated in the Near East over 10,000 years ago and then brought to Africa through successive migrations (Van Marle-Köster *et al*., 2021). Zebu cattle (*Bos indicus*) emerged from the Indian subcontinent and reached the African continent around 4,000 years ago. However, its expansion to North and East Africa only occurred around 1,300 years ago, accompanying Arab migrations (Utsunomiya *et al*., 2019).

Recent genomic work suggests that taurine and indicine cattle were probably first hybridized 4,500 years ago in the Near East and then dispersed to West and East Africa (Van Marle-Köster *et al*., 2021). These highly resilient livestock populations were kept in what is now known as the Sahara and reached East Africa between approximately 4,000 and 3,000 years ago (Horsburgh *et al*., 2013) and Southern Africa about 2,000 years ago (Robbins *et al*., 2005; Robbins *et al*., 2008). Nowadays, the African continent is mainly inhabited by zebu and taurine cattle, as well as their various derivatives (hybrids) known as Sanga (*Bos taurus africanus*) cattle (Mwai *et al*., 2015).

The Southern region of Africa (SADC) has more than 64 million head of cattle, distributed across various agroecological zones and included in different production systems and this has resulted in considerable phenotypic variation. Many of these cattle populations are named according to their origin and physical characteristics (Mapiye *et al*., 2018). For instance, Mozambique (MZ) and South Africa (SA) share transboundary indigenous Sanga cattle, including the Mozambican Nguni (MZ Nguni; known as Landim) and South African Nguni (SA Nguni) breeds. However, their names may not be an indication of any genetic differentiation among them.

Indigenous cattle, including the South African Boran (SA Boran), South African Tuli (SA Tuli), South African Nguni (SA Nguni), and Mozambican Nguni (MZ Nguni) have been naturally selected to withstand food shortages, high temperatures, and high incidence of parasites and diseases characteristic of the Southern African region (Nyamushamba *et al*., 2017).

Even though the important role of these indigenous populations, their large phenotypic diversity, and superior adaptive capacity to harsh environments have been acknowledged (Mwai *et al*., 2015; Mapiye *et al*., 2019), knowledge of their genetic diversity and the relationships among them is still lacking.

Genetic markers including single nucleotide polymorphisms (SNP), by means of genotyping panels, have been successfully used in Spain (Cañas-Álvarez *et al*., 2015), Sudan (Bahbahani *et al*., 2018), Ireland (Kelleher *et al*., 2017), and Mozambique (King *et al*., 2021) to examine patterns of genome-wide genetic relatedness, population structure (and admixture), and genetic diversity parameters within and among several indigenous cattle populations. The genetic relationship between indigenous breeds of cattle from Mozambique and South Africa has, however, received limited research attention. Genetic characterisation of SA Boran, SA Tuli, SA Nguni, and MZ Nguni has been carried out previously, but was limited to within country studies (Bessa *et al*., 2009; Makina *et al*., 2016; van der Westhuizen *et al*., 2019; Mamogobo *et al*., 2020). The only known molecular study investigating the genetic relationship among indigenous breeds of cattle in Mozambique and South Africa is by Madilindi *et al.* (2020), using microsatellite markers and small sample sizes (n=30).

The purpose of this study was to investigate the genetic diversity and relationships of four indigenous cattle populations from Mozambique and South Africa using genome-wide SNP markers. This information could be beneficial in the joint planning of conservation programs in both countries, as their indigenous cattle populations are in decline due to indiscriminate cross-breeding with imported exotic breeds (Mapiye *et al*., 2019).

2. Materials and methods

2.1 Animals and sampling

Ethical approval for the sampling of the Mozambican Nguni (MZ Nguni) cattle population was obtained from the Ethics Committee of the University of Pretoria, Faculty of Natural and Agricultural Sciences, University of Pretoria (ECO25–18). Tail hairs were collected in the Gaza (n=33), Inhambane (n=60), and Maputo (n=26) provinces where the Nguni predominates. Permission to use the South African genotypic data was granted by the relevant breeder associations. Data included the genotypes of South African Nguni (SA Nguni; n = 150), South African Tuli (SA Tuli; n = 150) and South African Boran (SA Boran; n = 150).

2.2 Genotyping and quality control

The 119 MZ Nguni individuals were genotyped at the Weatherbys Scientific Laboratory (Ireland) using the International Dairy and Beef BeadChip version three (IDB) SNP panel, composed of 53,450 genome-wide SNP (Twomery *et al*., 2019). Genotypes for the three South African cattle populations were generated from the GeneSeek® Genomic Profiler^{TM} (GGP) 80K SNP genotyping panel, at the Agricultural Research Council's Biotechnology Platform (ARC-BTP). A set of 37 069 common SNPs were identified, and merged using PLINK v1.09 (Purcell *et al*., 2007). After excluding individuals exceeding 10% missing genotypes, removing SNP with minor allele frequencies (MAF) below 0.01 and a call rate of less than 0.90, 29 010 SNP and 559 individuals remained for further analysis.

2.3 Data analysis

Expected and observed heterozygosity (H_e and H_o) values and inbreeding rates (F_{IS}) were estimated with PLINK software (Purcell *et al*., 2007). Analysis of the principal components (PCA) was used to evaluate the within- and between-population diversity using GCTA version 1.24 (Yang *et al*., 2011). For the PCA, a genomic relationship matrix was created, from which the eigenvectors and eigenvalues were calculated. The eigenvectors for principal components 1 and 2 were then plotted against each other.

Genetic structure among the analyzed populations was confirmed by the 'find.clusters' function performing principal components discriminant analysis (DAPC) with Adegenet (Jombart and Collins, 2015), an R package version 3.3.2. This analysis ran consecutive clustering of K-means from K = 1 to K = 5 and the optimal cluster was chosen as the one with the lowest Bayesian Information Criterion (BIC) value (Jombart *et al*., 2010). Ancestry coefficients were calculated from K=2 through K=5 and, for each analysis, 100 iterations were performed and summarized using the

snmf function of the LEA package in R. A graphical representation of the admixture patterns was depicted with GENESIS version 0.2.6.

Genetic distance between the populations was computed following the method proposed by Nei (1972) and implemented in StAMPP version 1.6.1 software (Pembleton *et al*., 2013). The matrix of genetic distances based on pairwise F_{ST} values was utilized to calculate the neighbour-joining dendrogram via the neighbor-joining method. The dendrogram was then plotted using the APE package in R version 3.3.2 (Paradis *et al*., 2004).

Estimates of effective population size (N_e) were performed according to Corbin *et al*. (2012) with SNeP version 1.1 (Barbato *et al*., 2015). Values from 0 to 1000 Mb were used for minimum and maximum inter-SNP distances, respectively. Data from each population were grouped into 30 bins of 50 kb distances. The N_e was estimated from the average r^2 value in each of these bins.

For each of the four populations, runs of homozygosity (ROH) were defined in PLINK v1.9 (Purcell *et al*., 2007) by: (1) using 50 SNP for each sliding window, (2) including a maximum of one heterozygous SNP, and a maximum of two missing SNP, (3) allowing 1 SNP every 75 kb as minimum density and, (4) allowing up to 1 Mb between successive homozygous SNP. Runs of homozygosity based inbreeding (F_{ROH}) was computed as a fraction of the total ROH length (L_{ROH}) to the overall autosomal length covered by SNPs.

To detect regions that have undergone positive selection in the genome of the four cattle breeds, two inter-population-based statistics, namely Wright's fixation index (F_{ST}) and ratio of extended haplotype homozygosity (*Rsb*), were used. Genetic differentiation (F_{ST}) was computed using an unbiased estimator proposed by Weir & Cockerham (1984). Pairwise F_{ST} was computed using the $-F_{ST}$ functionality in PLINK v1.09 (Purcell *et al*., 2007) for each SNP comparing the MZ Nguni to each of the other populations in this study. Genomic regions representing the top 2 % of SNP ($F_{ST} \geq 0.25$), were considered as being under selection. To detect selection signals, Bonferroni corrected F_{ST} *P*-values were $-\log_{10}$ transformed and Manhattan plots were drawn using the qqman R package (Turner, 2014) with default settings. Bonferroni correction consisted of adjusting the critical significance level of 0.05, dividing it by the number of SNPs tested. This procedure was done with the STAT package in R (Bolar, 2019).

The ratio of Extended Haplotype Homozygosity (EHH) between populations (*Rsb*) approach was employed for a pairwise comparative analysis of EHH measures in each SNP between the Mozambican Nguni and each of the

three South African populations, using the rehh R package version 1.11. To detect loci that were under selection, the values of *Rsb* were converted into *pRsb* (*pRsb* = −log[1 − 2 × ($\Phi_{(Rsb)}$−0, 5)]), where the function Φ(x) represents the Gaussian cumulative distribution (Gautier and Vitalis, 2012). The *Rsb* scores were computed for each pairwise comparison of the studied breeds, and the SNPs for which *pRsb* ≥3 (*P* − value = 0.001) were investigated.

Genes within putative selection signature regions were identified in the National Center for Biotechnology Information (NCBI), using the Bovine reference assembly ARS-UCD1.2. The function and metabolic processes in which these genes are implicated were identified through Panther (Mi *et al*., 2013).

3. Results

3.1 Quality control and genetic diversity parameters

Quality control (Table S1, supplementary material) was performed on the common extracted SNP, and 8,059 SNP were excluded. Of these, 25% had a call rate below 0.90, 20% had MAF under 0.01 and 55% deviated considerably from Hardy-Weinberg equilibrium (p<0.001). The highest number of excluded SNP (6,043) was observed in the SA Boran population, whereas the lowest number (2,697) was observed in SA Tuli cattle. Overall 29,010 SNP (96.10%) and 559 individuals remained for downstream analyses.

The mean MAF varied between 0.205 (± 0.145) in the SA Boran and 0.242 (± 0.147) in the SA Tuli. The observed heterozygosity value (H_o = 0.298 ± 0.153) was lower than the expected heterozygosity (H_e = 0.313±0.158) value in the MZ Nguni, which also had the highest level of inbreeding (0.047±0.092). The SA Tuli population revealed the greatest levels of genetic variability (0.326±0.158). Negative (-0.025 ± 0.053) to low positive (0.047 ± 0.092) rates of inbreeding were detected among the four cattle populations (Table 1).

Table 1 Mean (±SD) estimated genetic diversity parameters and rates of inbreeding in the cattle populations

Population	N	Mean MAF±SD	Mean H_e±SD	Mean H_o±SD	Mean F_{IS}±SD
SA Boran	148	0.205±0.145	0.284±0.158	0.291±0.166	-0.025±0.053
SA Nguni	143	0.233±0.151	0.312±0.160	0.315±0.165	-0.007±0.044
SA Tuli	149	0.242±0.147	0.324±0.153	0.326±0.158	-0.009±0.052
MZ Nguni	119	0.234±0.150	0.313±0.158	0.298±0.153	0.047±0.092
Merged	559	0.226±0.147	0.306±0.157	0.293±0.150	-0.011±0.087

SD Standard deviation; **H_e** Expected heterozygosity; **H_o** Observed heterozygosity; **F_{IS}** inbreeding coefficient; **MAF** Minor Allele Frequency; **SA Boran** South African Boran; **SA Nguni** South African Nguni; **SA Tuli**; **MZ Nguni** Mozambican Nguni

The average MAF for the studied cattle populations showed little variation, with a range from 0.205 ± 0.145 for SA Boran cattle to 0.242 ± 0.147 for SA Tuli cattle. In all four cattle populations, the proportion of SNP in the lower MAF categories were higher compared to the higher MAF categories (Figure 1). The SA Boran had the least SNP with MAF exceeding 0.3.

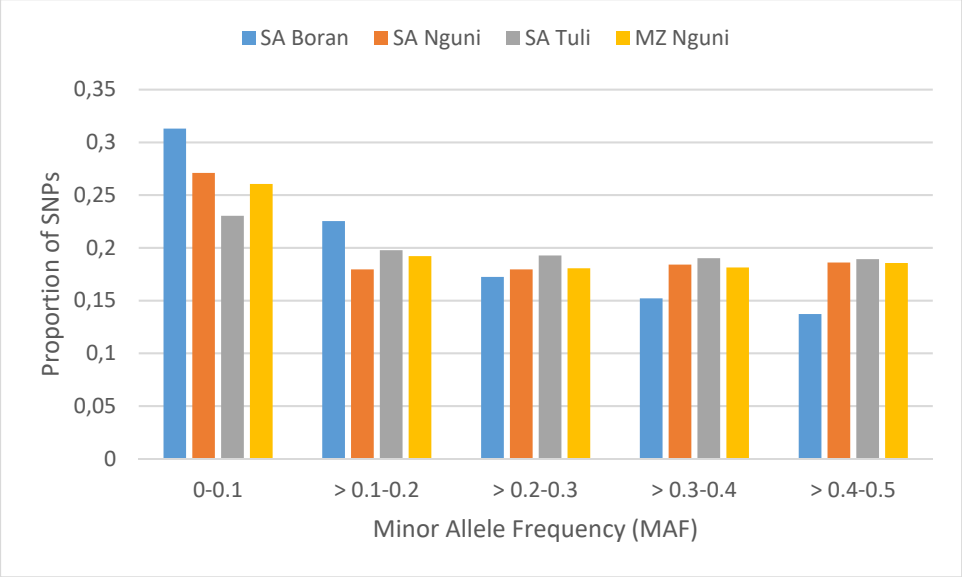

Figure 1 Distribution of SNP across MAF categories in each cattle population.

MAF: Minor Allele Frequency**; SA Boran** South African Boran; **SA Nguni** South African Nguni; **SA Tuli** South African Tuli**; MZ Nguni** Mozambican Nguni

3.2 Runs of homozygosity (ROH)

The overall number of ROH identified (Table 2) varied among populations, with the SA Tuli and SA Nguni populations displaying the highest (303) and lowest (56) number of ROH segments, respectively. Shorter ROHs were more abundant than longer ROH in all populations, suggesting more ancient inbreeding. The longest ROH segment (28.1 Mb) was found in the MZ Nguni population, which also had the largest mean number of SNP per ROH (93.92±50.87). The majority of ROH in all populations fell within the 1–5 Mb length category. The SA Tuli, followed by MZ Nguni and SA Boran, had the highest number of ROH segments in the largest length category (i.e. >15Mb). The estimated inbreeding coefficients were fairly uniform among the populations; SA Tuli exhibited the highest average F_{ROH} (0.006±0.003), whereas the lowest value was observed in the SA Nguni (0.003±0.003).

Table 2 Statistical parameters of ROH analyses in the four cattle populations

	SA Boran	**SA Nguni**	**SA Tuli**	**MZ Nguni**
Average length (Mb)	6.14±3.28	5.84±2.72	6.28±3.06	6.44±3.76
F_{ROH} (±SD)	0.004±0.003	0.003±0.003	0.006±0.003	0.004±0.004
F_{IS} (±SD)	-0.025±0.053	-0.007±0.044	-0.009±0.052	0.041±0.092
r (F_{ROH} – F_{IS})	0.430**	0.422**	0.462**	0.538**
Average number of SNP per ROH	88.56±43.76	85.71±40.32	92.07±42.80	93.92±50.87
N_e (12)	226	466	193	914
Maximum ROH length	24.52	15.77	19.93	28.10
Minimum ROH length	2.69	2.90	2.82	2.62
Number of ROH				
1–5 Mb	68	30	140	54
> 5–10 Mb	66	21	123	43
> 10-15 Mb	9	4	35	12
>15 Mb	4	1	5	4
Total	**147**	**56**	**303**	**113**

F_{ROH} ROH based inbreeding; **F_{IS}** Genomic based inbreeding; **r** Correlation; **N_e (12)** effective population size 12 generations ago; ****** ($P < 0.01$); **ROH** Runs of Homozygosity; **Mb** Megabyte; **SA Boran** South African Boran; **SA Nguni** South African Nguni; **SA Tuli** South African Tuli**; MZ Nguni** Mozambican Nguni

3.3 Admixture and population structure analyses

Principal component analysis (PCA) was performed to examine genetic relationships between the four populations of cattle (Figure 2). The first and second principal components explained 5.32% and 2.46% of the variability, respectively. The animals investigated in this study were allocated to three clusters. The first cluster comprised the two Nguni populations (MZ, and SA). The second and third clusters consisted of the SA Boran and SA Tuli populations, respectively. The SA Tuli and SA Boran populations were the most genetically distant (F_{ST} = 0.027).

Some individuals belonging to the MZ Nguni population were detected as outliers towards the SA Tuli and SA Boran populations, implying some genetic admixture with these populations or incorrect breed allocation.

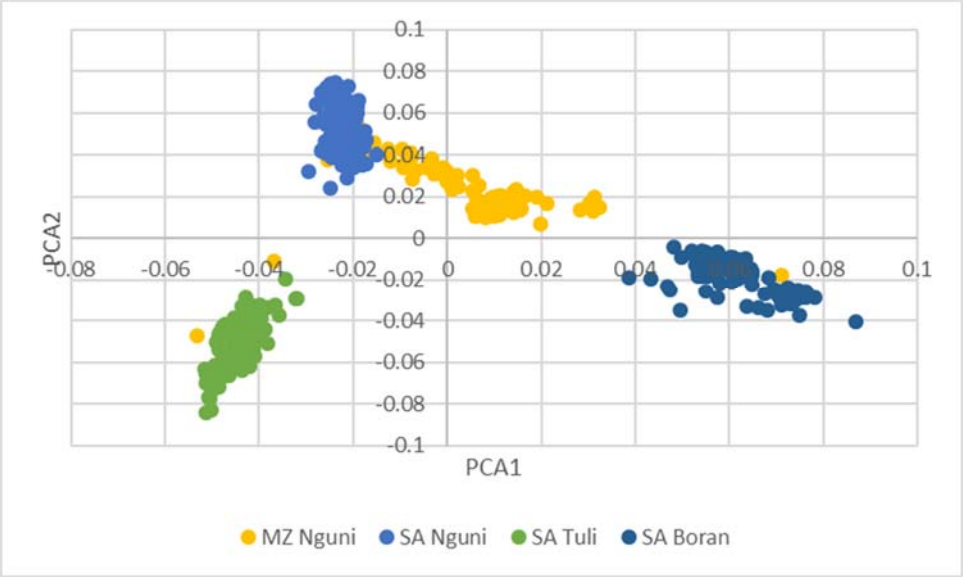

Figure 2 Genetic relationships between the four cattle populations.

SA Boran South African Boran; **SA Nguni** South African Nguni; **SA Tuli**; **MZ Nguni** Mozambican Nguni; **PCA** Principal component analysis

The most likely number of ancestral populations was estimated as three (K=3) based on the smallest Bayesian Information Criterion (BIC) value (Figure 3).

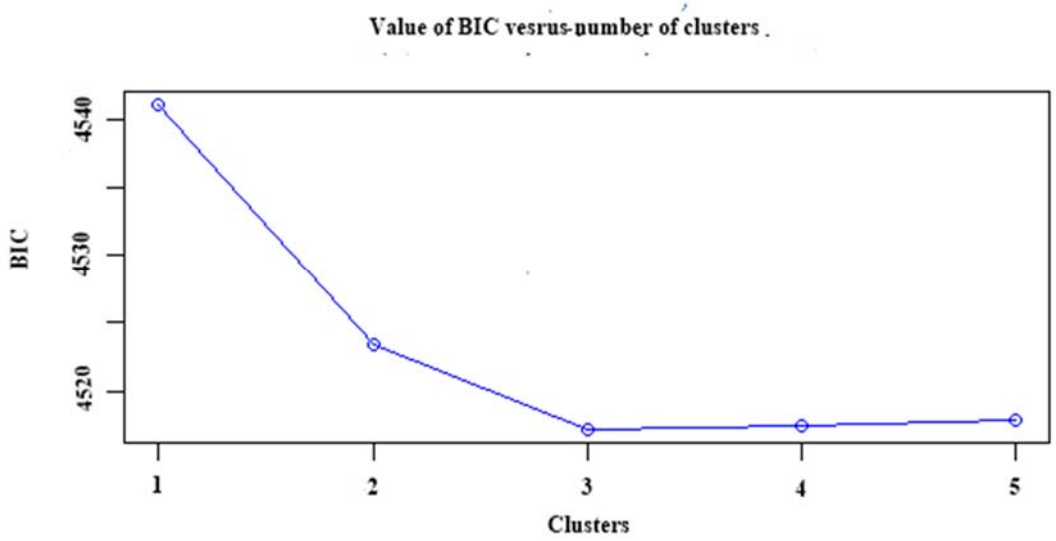

Figure 3 Bayesian Information Criterion (BIC) values versus number of clusters in the four cattle populations

The ADMIXTURE analysis supplied further evidence of the genomic similarity between the two Nguni populations, and the distinctness of this breed from the other breeds. At K = 3, which was proposed to be the most probable number of clusters, the SA Boran and SA Tuli populations formed distinct clusters. The MZ Nguni and SA Nguni populations shared a common ancestor (Figure 4).

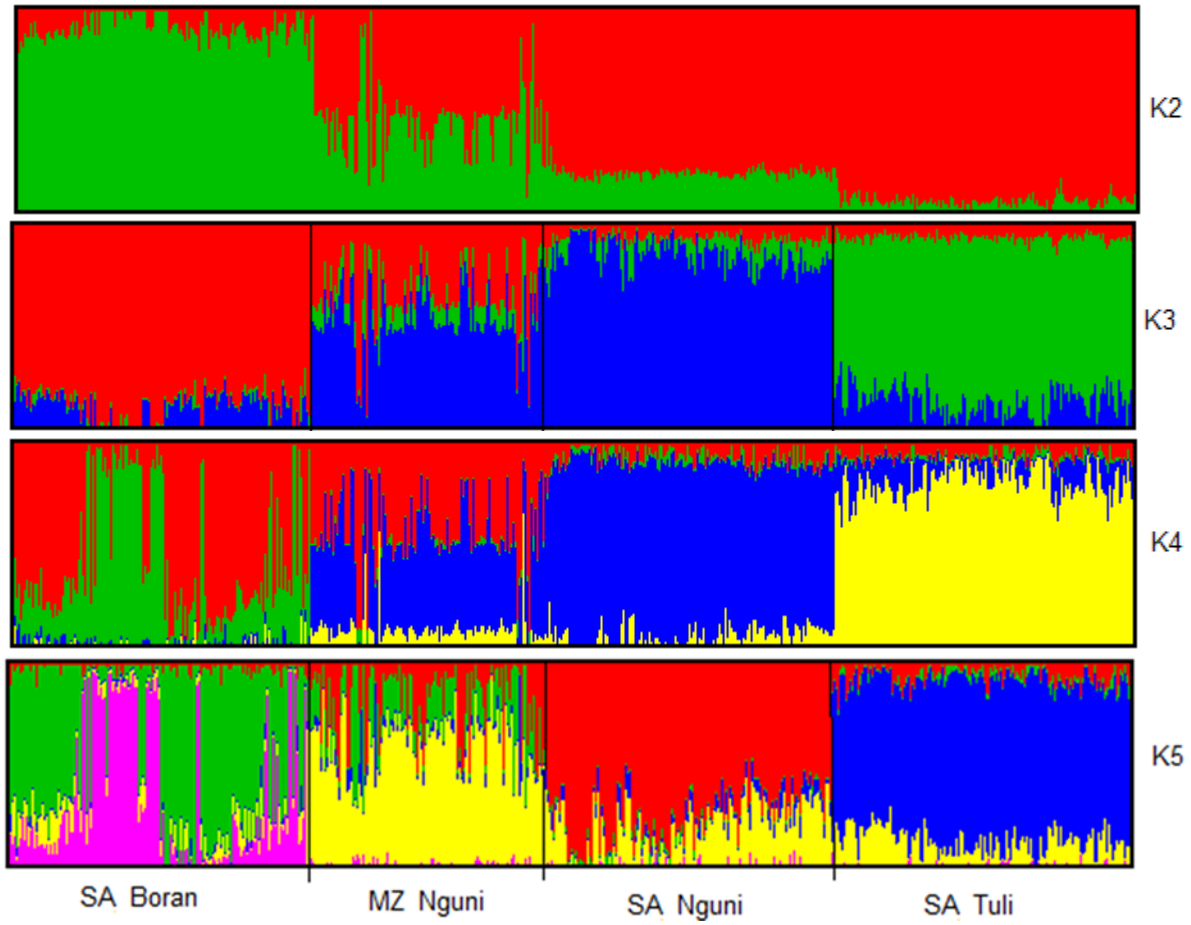

Figure 4 Admixture plot depicting the cluster allocation for the four cattle populations.

SA Boran South African Boran; **SA Nguni** South African Nguni; **SA Tuli** South African Tuli**; MZ Nguni** Mozambican Nguni

The SA populations were more uniform, with the within-population shared ancestry ranging between 81 % and 88 %, whereas the MZ Nguni population was more admixed and shared 55%, 35%, and 10% of their genome with SA Nguni, SA Boran, and SA Tuli populations, respectively (Table 3).

Table 3 Membership probability estimates for four studied cattle populations

Main genetic clusters (K = 3)			
Population	Cluster 1	Cluster 2	Cluster 3
SA Boran	**0.881**	0.019	0.100
SA Nguni	0.057	0.080	**0.863**
SA Tuli	0.058	**0.811**	0.131
MZ Nguni	0.351	0.103	**0.546**

SA Boran South African Boran; **SA Nguni** South African Nguni; **SA Tuli** South African Tuli**; MZ Nguni** Mozambican Nguni

Pairwise F_{ST} estimates revealed less differentiation between populations than expected. The smallest genetic distance was observed among the SA Nguni and MZ Nguni cattle and the largest between the SA Boran and SA Tuli populations. Of the SA breeds, the SA Nguni and SA Tuli Sanga breeds were the least related, with the indicine SA Boran being more related to the SA Tuli than the SA Nguni (Table 4).

Table 4 F_{ST}-based genetic differentiation among Mozambican and South African cattle populations

Population	**SA Boran**	**MZ Nguni**	**SA Nguni**	**SA Tuli**
SA Boran	***			
MZ Nguni	0.051	***		
SA Nguni	0.081	0.031	***	
SA Tuli	0.088	0.049	0.062	***

SA Boran South African Boran; **SA Nguni** South African Nguni; **SA Tuli** South African Tuli**; MZ Nguni** Mozambican Nguni

Figure 5 shows an individual-animal-based Neighbor-Joining (NJ) dendrogram for the 559 individuals. The groups formed in the NJ dendrogram were, in general, in agreement with the ADMIXTURE and PCA analyses. Most of the animals were grouped within their population in the dendrogram, although some of the MZ Nguni individuals were dispersed. Some MZ Nguni animals clustered with or towards the SA Nguni, SA Boran, and SA Tuli populations, indicating some genetic relationship.

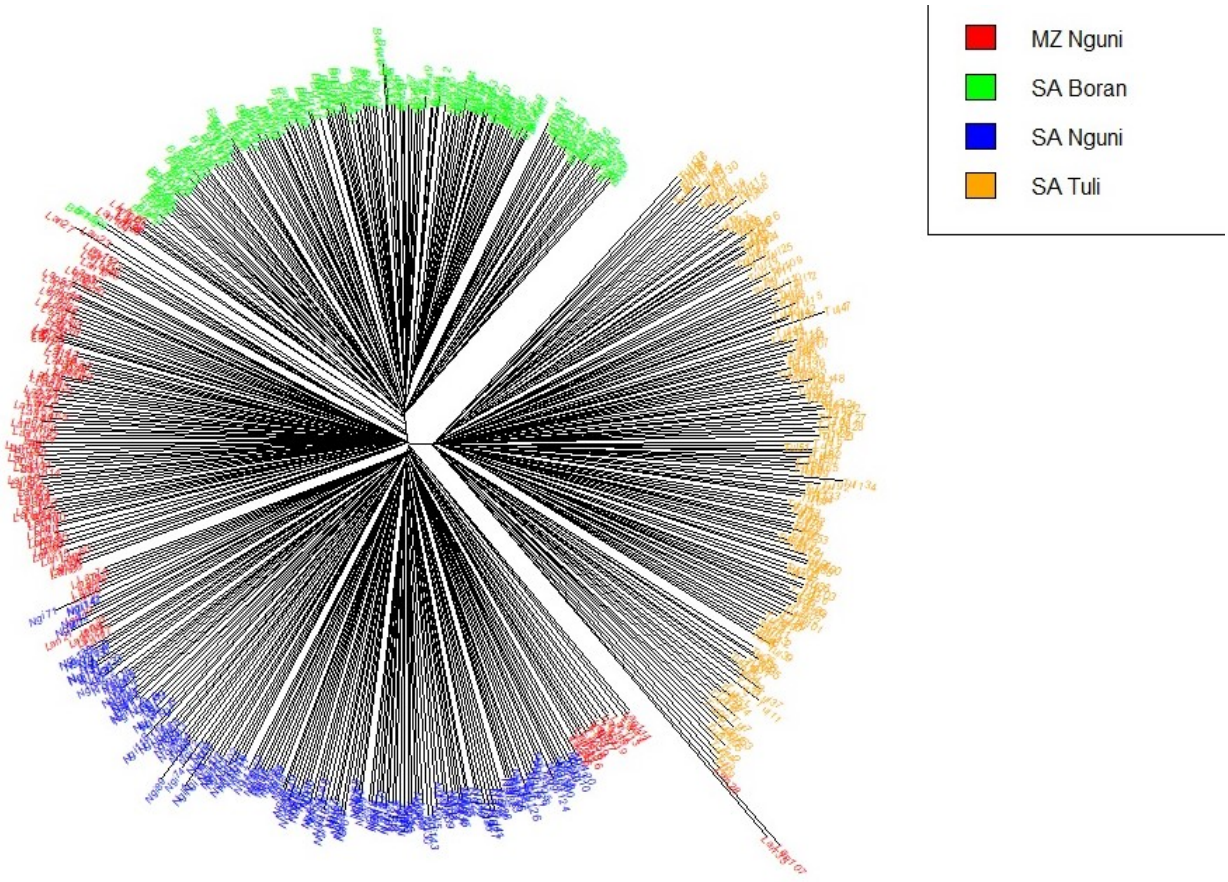

Figure 5 Unrooted NJ dendrogram depicting the relationships between animals from four indigenous Southern African cattle populations based on pairwise genetic distances. **SA Boran** South African Boran; **SA Nguni** South African Nguni; **SA Tuli** South African Tuli**; MZ Nguni** Mozambican Nguni

3.4 Signatures of selection identified using F_{ST}

Figure S1 (supplementary material) shows Manhattan plots of 72, 31, and 20 highly significantly differentiated SNP ($F_{ST} \geq 0.25$) between the MZ Nguni population and the SA Boran, SA Nguni, and SA Tuli populations, respectively. These SNP were located throughout the genome except for chromosomes (BTA) 15, 16, 20, 21, 26, 27, 28, and 29, which did not contain any selection signatures. One of the differentiated regions (on BTA4) was shared between the MZ Nguni vs SA Nguni and MZ Nguni vs SA Tuli breed comparisons. The MZ Nguni vs SA Boran pair had the most differentiated regions (72) while the MZ Nguni vs SA Tuli pair had the least (20). The most differentiated region ($F_{ST} = 0.47$) was detected between the MZ Nguni and SA Boran breeds on BTA11.

3.5 Signatures of selection detected with Rsb approach

The *Rsb* analysis found 21 significant SNP (*pRsb* ≥ 3) across ten chromosomes in the MZ Nguni vs SA Boran comparison. In the MZ Nguni vs SA Nguni comparison, a total of 36 SNPs located across eighteen chromosomes were significantly differentiated. Finally, in the MZ Nguni vs SA Tuli comparison, 49 significant SNPs were found spread over twenty-one chromosomes. One of these candidate regions was shared between the MZ Nguni vs SA Nguni and MZ Nguni vs SA Boran breed comparisons. SNP with scores above the mean (-log10 *p*-value ≥ 4) were found on BTA2 and BTA20 in the MZ Nguni vs SA Boran comparison, on BTA10 for the MZ Nguni vs SA Nguni comparison, and on BTA14 for the MZ Nguni vs SA Tuli comparison (Figure S2, supplementary material).

Using Panther analysis (Mi *et al*., 2013), a number of genes associated with important functions and pathways in cattle were identified. Corresponding genes were identified by both *Rsb* and F_{ST} analyses on chromosomes BTA1, BTA4, BTA8, BTA9, and BTA18. These genes are related to several biological functions, such as growth and feed efficiency (*TRNAC-GCA*), muscle movements and calving ease (*KCNMB2* and *MYLK3*), as well as oxytocin signaling pathway, and milking speed (*MYLK3*) (Table 5).

Other genes that warrant further investigation include *ZEB2*, which has been associated with polledness and congenital malformations, and *LAMC1*, which has been associated with parasites resistance in cattle. Genes related to riboflavin transport (e.g. *SLC52A3*), lipid metabolism and adipose tissue development (e.g. *NGFR*), cell differentiation and embryo development (e.g. *NGFR*, *FGD5*), renal potassium transport (e.g. *KCNK5*) and ciliogenesis (e.g. *CEP83*) were also identified.

Table 5 Genes under selection in three pairwise comparisons between Southern African cattle populations

Breed pair	BTA	Position (bp)		Method	Gene name
		Start	End		
MZ Nguni vs SA Boran	2	52174189	52264481	*Rsb*	*ZEB2*
	4	13061908	13441077	F_{ST}	*DYNC1I1*
	11	75240502	75572709	F_{ST}	*KLHL29*
	16	64062643	64183739	*Rsb*	*LAMC1*
	18	9138128	10154233	F_{ST}, *Rsb*	***CDH13***
	19	37093168	37112123	F_{ST}	*NGFR*
	22	57381762	57505298	F_{ST}	*FGD5*
MZ Nguni vs SA Nguni	1	88158068	88671406	F_{ST}, *Rsb*	***KCNMB2***

	4	17065366	17065438	F_{ST}, *Rsb*	***TRNAC-GCA***
	5	24070150	24216286	F_{ST}	*CEP83*
	8	33330545	33331036	F_{ST}, *Rsb*	***LOC782926***
	8	33969443	33971875	F_{ST}, Rsb	***LOC100141071***
	15	76852567	77063233	*Rsb*	*CSTPP1*
	18	15032402	15089534	F_{ST}, *Rsb*	***MYLK3***
	21	12595555	12855213	*Rsb*	*MCTP2*
MZ Nguni vs SA Tuli	4	17065366	17065438	F_{ST}, *Rsb*	***TRNAC-GCA***
	8	33330545	33331036	F_{ST}, *Rsb*	***LOC782926***
	8	33969443	33971875	F_{ST}, *Rsb*	***LOC100141071***
	9	17257479	17257550	F_{ST}, *Rsb*	***TRNAA-UGC***
	10	29455840	29970835	F_{ST}	*FMN1*
	13	60290424	60307559	F_{ST}	*SLC52A3*
	14	69265272	69301597	*Rsb*	*NDUFAF6*
	18	15032402	15089534	F_{ST}, *Rsb*	***MYLK3***
	23	13060758	13101202	F_{ST}	*KCNK5*

Bold indicate coincident genes

SA Boran South African Boran; **SA Nguni** South African Nguni; **SA Tuli** South African Tuli**; MZ Nguni** Mozambican Nguni

4. Discussion

Knowledge of genetic diversity and relationships amongst indigenous cattle populations is crucial for successful animal production, especially in the challenging environments of Southern Africa, and in the face of climate change (Nyamushamba *et al*., 2017). The genetic diversity of a population is essential for its genetic improvement, particularly for adaptive traits and those associated with sustainable production (Groeneveld *et al*., 2010; Mapiye *et al*., 2019). This work examined the genetic relationships and diversity between four cattle populations from Mozambique and South Africa.

The average MAF for the four populations was 0.226±0.147, with the highest proportion of SNP showing low MAF values (≤ 0.1). These findings correspond with values found in other indicine and indicine-hybrid (Sanga) breeds (Qwabe *et al*., 2013; Perez O'Brien *et al*., 2014; Lashmar *et al*., 2018; Gebrehiwot *et al*., 2021). Unlike indicine cattle (*Bos indicus*), taurine breeds (*Bos taurus*) generally have a higher percentage of SNP in the higher MAF

categories as they are discovery and international transboundary breeds (Qwabe *et al*., 2013; Zwane *et al*., 2016; Lashmar *et al*., 2018). The lower values reported in indicine cattle have been associated with possible ascertainment bias due to the design and development process of commercial SNP panels (Perez O'Brien *et al*., 2014; Bejarano *et al*., 2018).

Heterozygosity is an indicator of genetic diversity within and between populations, which is important in the design of breeding programs and conservation strategies for indigenous cattle populations. Moderate genetic diversity was detected in the four studied cattle populations. These values were higher than those observed in South African Nguni cattle populations (0.23–0.24) by Zwane *et al*. (2016). They were, however, lower than expected heterozygosity values (0.40) reported in Sukuma, Tarime, and Maasai in Tanzanian indigenous zebu populations (Msalya *et al*., 2017).

Similar estimates of observed and expected heterozygosity were reported for Sanga cattle from South Africa (Makina *et al*., 2014) and for indigenous Ethiopian cattle (Edea *et al*., 2012). Among the four studied populations, SA Tuli cattle showed the highest degree of genetic variability (H_e = 0.32) whereas SA Boran revealed the lowest. The lower H_e values of SA Boran cattle could reflect its smaller population size due to its late introduction to the country (Abin *et al*., 2016). High heterozygosity levels are usually related to long-term natural selection for adaptation and the admixture history of different populations (Gororo *et al*., 2018). However, uncontrolled breeding practiced in pastoral and agro-pastoral production systems was reported to be responsible for the high genetic variability present in Barentu and Awgaro indigenous cattle populations of Eritrea (Goitom *et al*., 2019), and this could also be a contributing factor in the current findings.

The principal component analysis (PCA) indicated a strong genetic relationship between MZ Nguni and SA Nguni, as opposed to the SA Tuli and SA Boran. The close relationship between the MZ Nguni and SA Nguni could be explained by the common origin of these two cattle populations. Nguni cattle originated from North Africa and migrated southwards, settling in different areas including South Africa and Mozambique where distinctive cattle ecotypes were developed (Bester *et al*., 2003; Mwai *et al*., 2015). Furthermore, SA Nguni cattle were re-introduced to Mozambique after the civil war which had greatly reduced the national herd (King *et al*., 2021). The neighbor-joining dendrogram supported the PCA results as many MZ Nguni animals were blended with the SA Nguni, SA Boran, and SA Tuli populations, indicating some genetic proximity.

The results of the Admixture analyses supported the strong genetic relationship between the MZ Nguni and the SA populations. The MZ Nguni's genetic material is shared with all three South African populations, with the highest levels of similarity observed with the SA Nguni population. The genetic links observed among the populations suggest common genetic origin and a high degree of gene flow among populations (Radhika *et al*., 2018). MZ Nguni cattle have undergone a significant genetic bottleneck owing to the civil war (1977–1992), which significantly reduced their numbers. This period was followed by introgressions, as the Mozambican government imported cattle from neighboring countries, especially Zimbabwe and South Africa (King *et al*., 2021). Furthermore, gene flow may have been possible as a result of cattle migrations and trading, as well as indiscriminate cross-breeding practiced in communal management systems in both countries.

Generally, effective population size estimates in the present study are comparable with those of other indigenous cattle breeds in South Africa (Makina *et al*., 2015; Abin *et al*., 2016). The small effective population size of the SA Tuli could reflect the small founding population of this breed. The SA Tuli, which was developed and improved from Tswana cattle in Zimbabwe, was only introduced to South Africa in 1976 (Glennels, 2019). Among the studied populations, SA Tuli cattle had a relatively higher genomic inbreeding coefficient compared to the other three populations. Additionally, the SA Tuli has a relatively small population size of approximately 7,000 animals, and might have undergone bottlenecks as suggested by the high number of ROH segments in the genome. The small number of long ROH segments found in this study indicates that none of the populations were subjected to recent inbreeding. The lower ROH abundance found in Nguni populations suggests that a relatively larger effective population size was preserved over generations.

Signatures of selection containing genes with important biological functions were identified using both *Rsb* and F_{ST} approaches. Overall, this study detected 229 signatures of selection, most of which were identified by a single methodology. Only eleven signatures of selection were identified by both the *Rsb* and F_{ST} approach. This may be due to differences in the parameters used in each methodology, thus detecting distinct traces left in the genome over time (Tang *et al*., 2007).

The *LAMC1* gene was found to be significantly differentiated in the MZ Nguni vs SA Boran comparison by *Rsb* analysis (Table 5). This gene was previously associated with resistance to *H. contortus* infection in Red Maasai and Dorper sheep breeds under natural infection conditions (Benavides *et al*., 2015). A region identified by *Rsb*

analysis in the MZ Nguni vs SA Boran comparison (Figure S2, supplementary material) included the *ZEB2* gene. Recent studies indicate the association of this gene with polledness and congenital malformations in French Charolais cattle (Wiedemar *et al*., 2014).

Other important genes in the selected regions are directly or indirectly involved in reproduction, including fertilization and calving ease in cattle. These genes were identified by both *Rsb* and F_{ST} analyses and included *KCNMB2* and *MYLK3* in the MZ Nguni vs SA Nguni comparison (Buzanskas *et al*., 2017; Fraser *et al*., 2020). Indigenous Southern African cattle breeds are adapted to survive and reproduce without assistance in the extensive, harsh environments found in the region (Mwai *et al*., 2015). These findings support the existence of selection signatures at loci involved in reproduction which likely occurred in the course of adaptation of these cattle populations to southern African conditions.

5. Conclusions

Genomic studies examining the population structure of Mozambican indigenous cattle are scarce. This work is the first extensive study aiming to evaluate the genetic structure of MZ Nguni and its relationships with South African indigenous cattle using genome-wide SNP markers. The results of this research indicate low genetic differentiation among the four populations, which could be a consequence of both common ancestry and high gene flow rates. Most importantly, the study confirmed that the MZ Nguni and the SA Nguni have similar genetic ancestry. Furthermore, candidate genes were detected which may contribute to a better understanding of the genetic adaptation to various selection pressures. Finally, these findings point to the need for a common genetic resource management program between Mozambique and South Africa to conserve indigenous livestock in the region.

Acknowledgements

We would like to thank the Mozambican Ministry of Agriculture and Rural Development and all the cattle farmers involved in this study, for their cooperation. The authors wish to acknowledge the permission from the respective South African breed societies to use the genotypes for this study.

Funding

We thank the Mozambican Ministry of Science, Technology and Higher Education for the funding provided, without which it would not have been possible to carry out this work. South African beef cattle genotypes used in the analyses were generated within the SA Beef Genomic Program funded by the Technology Innovation Agency.

References

Abin, S., Theron, H.E., van Marle-Köster, E., 2016. Short communication: Population structure and genetic trends for indigenous African beef cattle breeds in South Africa. S. Afr. J. Anim. Sci. 46, 152-156. http://dx.doi.org/10.4314/sajas.v46i2.5.

Bahbahani, H., Salim, B., Almathen, F., Al Enezi, F., Mwacharo, J.M., Hanotte, O., 2018. Signatures of positive selection in African Butana and Kenana dairy zebu cattle. PLoS One 13(1), e0190446. https://doi.org/10.1371/journal.pone.0190446.

Barbato, M., Orozco-terWengel, P., Tapio, M., Bruford, M.W., 2015. SNeP: A tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. Front. Genet. 6,109-122. https://doi.org/10.3389/fgene.2015.00109.

Bejarano, D., Martínez, R., Manrique, C., Parra, L.M., Rocha, J.F., Gómez, Y., Abuabara, Y., Gallego, J., 2018. Linkage disequilibrium levels and allele frequency distribution in Blanco Orejinegro and Romosinuano Creole cattle using medium density SNP chip data. Genet. Mol. Biol. 41, 2, 426-433. http://dx.doi.org/10.1590/1678-4685-GMB-2016-0310.

Benavides, M.V., Sonstegard, T.S., Kemp, S., Mugambi, J.M., Gibson, J.P., Baker, R.L., Hanotte, O., Marshall, K., Van Tassell, C., 2015. Identification of novel loci associated with gastrointestinal parasite resistance in a Red Maasai x Dorper backcross Population. PLoS One 10(4), e0122797. http://dx.doi.org/10.1371/journal.pone.0122797.

Bessa, I., Pinheiro, I., Matola, M., Dzama, K., Rocha, A., Alexandrino, P., 2009. Genetic diversity and relationships among indigenous Mozambican cattle breeds. S. Afr. J. Anim. Sci. 39: 61-72. http://dx.doi.org/10.4314/sajas.v39i1.43548.

Bester, J., Matjuda, L.E., Rust, J.M., Fourie, H.J., 2003. The Nguni: a case study. In: Vilakati, D., Morupisi, C., Setshwaelo, L., Wollny, C., von Lossau, A., Drews, A. (eds.), Community based management of animal genetic resources. FAO (Food and Agriculture Organization of the United Nations), Rome, Italy. pp. 45–68.

Bolar, K., 2019. Interactive document for working with basic statistical analysis. https://cran.r-project.org/web/packages/STAT/STAT.pdf (accessed 03 May 2022).

Buzanskas, M.E., do Amaral Grossi, D., Ventura, R.V., Schenkel, F.S., Chud, T.C.S., Stafuzza, N.B., Rola, L.D., Meirelles, S.L.C., Mokry, F.B., Mudadu, M.A., Higa, R.H., da Silva, M., de Alencar, M.M., Regitano, L.C.A., Munari, D.P., 2017. Candidate genes for male and female reproductive traits in Canchim beef cattle. J. Anim. Sci. Biotech. 8, 67. https://doi.org/10.1186/s40104-017-0199-8.

Cañas-Álvarez, J.J., González-Rodríguez, A., Munilla, S., Varona, L., Díaz, C., Baro, J.A., Altarriba, J., Molina, A.A., Piedrafita, J., 2015. Genetic diversity and divergence among Spanish beef cattle breeds assessed by a bovine high-density SNP chip. J. Anim. Sci. 93, 5164–5174. https://doi.org/10.2527/jas.2015-9271.

Corbin, L.J., Liu A., Bishop S., Woolliams, J., 2012. Estimation of historical effective population size using linkage disequilibria with marker data. J. Anim. Breed. Genet. 129(4), 257–270. http://dx.doi.org/10.1111/j.1439-0388.2012.01003.x.

Edea, Z., Dadi, H., Kim, S.W., Dessie, T., Kim, K.S., 2012. Comparison of SNP variation and distribution in indigenous Ethiopian and Korean Cattle (Hanwoo) populations. Genomics Inform. 10(3), 200–205. https://doi.org/10.5808/GI.2012.10.3.200.

Fraser, L., Paukszto, Ł., Mańkowska, A., Brym, P., Gilun, P., Jastrzębski, J.P., Pareek, C.S., Kumar, D., Pierzchała, M., 2020. Regulatory potential of long non-coding RNAs (lncRNAs) in boar Spermatozoa with good and poor Freezability. Life (Basel). 21, 10(11), 300. doi: 10.3390/life10110300.

Gautier, M., Vitalis, R., 2012. Rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics 28 (8), 1176–1177. https://doi.org/10.1093/bioinformatics/bts115.

Gebrehiwot, N.Z., Strucken, E.M., Marshall, K., Aliloo, H., Gibson, J.P., 2021. SNP panels for the estimation of dairy breed proportion and parentage assignment in African crossbred dairy cattle. Genet. Sel. Evol. 53, 21. http://dx.doi.org/10.1186/s12711-021-00615-4.

Glennels, K., 2019. Tuli Cattle, Beef production in South Africa. http://southafrica.co.za/tuli-cattle.html (accessed 13 May 2020).

Goitom, S., Gicheha, M.G., Njonge, F.K., Kiplangat, N., 2019. Genome-wide genetic diversity, population structure and admixture analysis in Eritrean Indigenous Cattle. S. Afr. J. Anim. Sci, 49 (6). https://doi.org/10.4314/sajas.v49i6.12.

Gororo, E., Makuza, S.M., Chatiza, F.P., Chidzwondo, F., Sanyika, T.W., 2018. Genetic diversity in Zimbabwean Sanga cattle breeds using microsatellite markers. S. Afr. J. Anim. Sci. 48(1), 128-141. https://doi.org/10.4314/sajas.v48i1.15.

Groeneveld, L.F., Lenstra, J.A., Eding, H., Toro, M.A., Scherf, B., Pilling, D., 2010. Genetic diversity in farm animals – a review. Anim. genet. 41(1), 6-31. https://doi.org/10.1111/j.1365-2052.2010.02038.x.

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

Ibeagha-Awemu, E.M., Peters, S.O., Bemji, M.N., Adeleke, M.A., Do, D.N., 2019. Leveraging available resources and stakeholder involvement for improved productivity of African livestock in the era of genomic breeding. Front. Genet. 10, 357. http://dx.doi.org/10.3389/fgene.2019.00357.

Jombart, T., Collins, C., 2015. A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0. http://adegenet.r-forge.r-project.org/files/ tutorial-dapc.pdf (accessed 13 March 2021).

Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11, 94. https://doi.org/10.1186/1471-2156-11-94.

Kelleher, M.M., Berry, D.P., Kearney, J.F., Mcparland, S., Buckley, F., Purfield, D.C., 2017. Inference of population structure of purebred dairy and beef cattle using high-density genotype data. Animal 11, 15-23. https://doi.org/10.1017/S1751731116001099.

King, F.J.M., Banga, C.B., Visser, C., 2021. Genetic diversity and population structure of three native cattle populations in Mozambique. Trop. Anim. Health Prod. 53, 117. https://doi.org/10.1007/s11250-021-02562-0.

Lashmar, S., Visser, C., van Marle-Koster, E., Muchadeyi, F., 2018. Genomic diversity and autozygosity within the SA Drakensberger beef cattle breed. Livest. Sci. 212, 111–119. https://doi.org/10.1016/j.livsci.2018.04.006.

Madilindi, M.A., Banga, C.B., Bhebhe, E., Sanarana, Y.P., Nxumalo, K.S., Taela, M.G., Magagula, B.S., Mapholi, N.O., 2020. Genetic diversity and relationships among three Southern African Nguni cattle populations. Trop. Anim. Health Prod. 52, 753–762. https://link.springer.com/article/10.1007/s11250-019-02066-y.

Makina, S.O., Muchadeyi, F.C., van Marle-Köster, E., MacNeil, M.D., Maiwashe, A., 2014. Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel. Front. Genet. 5, 333. https://doi.org/10.3389/fgene.2014.00333.

Makina, S.O., Taylor, J.F., Van Marle-Köster, E., Muchadeyi, F.C., Makgahlela, M., MacNeil, M.D., Maiwashe, A., 2015. Extent of linkage disequilibrium and effective population size in four South African Sanga cattle breeds. Front. Genet. 6, 1–12. https://dx.doi.org/10.3389%2Ffgene.2015.00337.

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. Genet. Select. Evol. 48, 88. https://doi.org/10.1186/s12711-016-0266-1.

Mamogobo, M.D., Mapholi, N.O., Nephawe, K.A., Nedambale, T. L., Mpofu, T. J., Sanarana, Y. P., Mtileni, B. J., 2020. Genetic characterisation of non-descript cattle populations in communal areas of South Africa. Anim. Prod. Sci., 61, 84–91. https://doi.org/10.1071/AN20030.

Mapiye, C., Chikwanha, O.C., Chimonyo, M., Dzama, K., 2019. Strategies for sustainable use of indigenous cattle genetic resources in Southern Africa. A review. Diversity. 11, 214 https://doi.org/10.3390/d11110214.

Mi, H., Muruganujan, A., Thomas, P.D., 2013. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res. 41(Database issue), D377-D386. https://dx.doi.org/10.1093%2Fnar%2Fgks1118.

Msalya, G., Kim, E.S., Laisser, E.L., Kipanyula, M.J., Karimuribo, E.D., Kusiluka, L.J., Chenyambuga, S.W., Rothschild, M.F., 2017. Determination of genetic structure and signatures of selection in three strains of tanzania shorthorn zebu, boran and Friesian cattle by genome-wide SNP analyses. PLoS One 12, e0171088. http://dx.doi.org/10.1371/journal.pone.0171088.

Mwai, O., Hanotte, O., Kwon, Y.J., Cho, S., 2015. African indigenous cattle: unique genetic resources in a rapidly changing world. Asian Austral. J. Anim. Sci. 28, 911–921. https://doi.org/10.5713/ajas.15.0002R.

Nei, M., 1972. Genetic distance between populations. Am. Nat. 106, 283-292. https://doi.org/10.1086/282771.

Nyamushamba, G., Mapiye, C., Tada, O., Halimani, T., Muchenje, V., 2017. Conservation of indigenous cattle genetic resources in Southern Africa's smallholder areas: Turning threats into opportunities – A review. Asian Austral. J. Anim. Sci. 30, 603–622. https://doi.org/10.5713/ajas.16.0024.

Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290. https://doi.org/10.1093/bioinformatics/btg412.

Pembleton, L.W., Cogan, N.O.I., Forster, J.W., 2013. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Mol. Ecol. Resour. 13, 946–952. https://doi.org/10.1111/1755-0998.12129.

Perez O'Brien, A.M., Utsunomiya, Y.T., Meszaros, G., Bickhart, D.M., Liu, G.E., Van Tassell, C.P., Sonstegard, T.S., Da Silva, M.V., Garcia, J.F., Solkner, J., 2014. Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. Genet. Sel. Evol. 46 (1), 19. http://dx.doi.org/10.1186/1297-9686-46-19.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575. https://doi.org/10.1086/519795.

Qwabe, S.O., van Marle-Köster, E., Maiwashe, A., Muchadeyi, F.C., 2013. Evaluation of the BovineSNP50 genotyping array in four South African cattle populations. S. Afr. J. Anim. Sci. 43, 64-67. http://dx.doi.org/10.4314/sajas.v43i1.7.

Radhika, G., Aravindakshan, T.V., Jinty, S., Ramya, K., 2018. Evaluation of genetic diversity, population structure, and relationship between legendary Vechur cattle and crossbred cattle of Kerala State, India. Anim. Biotechnol. 29, 50–58. https://doi.org/10.1080/10495398.2017.1297719.

Robbins, L.H., Campbell, A.C., Murphy, M.L., Brook, G.A., Liang, F., Skaggs, S.A., Srivastava, P., Mabuse, A.A., Badenhorst, S., 2008. Recent archaeological research at Toteng, Botswana: Early domesticated livestock in the Kalahari. J. Afr. Archaeol. 6, 131–149. http://dx.doi.org/10.3213/1612-1651-10106.

Robbins, L.H., Campbell, A.C., Murphy, M.L., Brook, G.A., Srivastava, P., Badenhorst, S., 2005. The advent of herding in Southern Africa: Early AMS dates on domestic livestock from the Kalahari Desert. Curr. Anthropol. 46, 671- 677. http://dx.doi.org/10.1086/432748.

Tang, K., Thornton, K.R., Stoneking, M., 2007. A new approach for using genome scans to detect recent positive selection in the human genome. PLoS Biol. 5 (7), e171. https://doi.org/10.1371/journal.pbio.0050171.

Turner, S.D., 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J. Open Source Softw. 3(25), 731. https://doi.org/10.21105/joss.00731.

Twomery, A.J., Berry, D.P., Evans, R.D., Doherty, M.L., Graham, D.A., Purfiled, D.C., 2019. Genome-wide association study of endo-parasite phenotypes using imputed whole-genome sequence data in dairy and beef cattle. Genet. Sel. Evol. 51, 15. http://dx.doi.org/10.1186/s12711-019-0457-7.

Utsunomiya, Y.T., Milanesi, M., Fortes, M.R.S., Porto-Neto, L.R., Utsunomiya, A.T.H., Silva, M.V.G.B., Garcia, J.F., Ajmone-Marsan, P., 2019. Genomic clues of the evolutionary history of *Bos indicus* cattle. Anim. Genet. 50 (6), 557-568. https://doi.org/10.1111/age.12836.

van der Westhuizen, L., MacNeil, M.D., Scholtz, M.M., Neser, F.W.C., Makgahlela, M.L., van Wyk, J.B., 2019. Genetic variability and relationships in nine South African cattle breeds using microsatellite markers. Trop. Anim. Health Prod. 52, 177–184. https://doi.org/10.1007/s11250-019-02003-z.

Van Marle-Köster, E, Visser, C., Sealy, J., Frantz, L., 2021. Capitalizing on the Potential of South African Indigenous Beef Cattle Breeds: A Review. Sustainability 13, 4388. https://doi.org/10.3390/su13084388.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370. https://doi.org/10.2307/2408641.

Wiedemar, N., Tetens, J., Jagannathan, V., Menoud, A., Neuenschwander, S., Bruggmann, R., Thaller, G., Drogemuller, C., 2014. Independent polled mutations leading to complex gene expression differences in cattle. PlosONE9(3),e93435. doi:10.1371/journal.pone.0093435.

Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M., 2011. GCTA: A tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88, 76–82. https://dx.doi.org/10.1016%2Fj.ajhg.2010.11.011.

Zwane, A.A., Maiwashe, A., Makgahlela, M.L., Choudhury, A., Taylor, J.F., van Marle-Köster, E., 2016. Genome-wide identification of breed-informative single nucleotide polymorphisms in three South African indigenous cattle breeds. S. Afr. J. Anim. Sci. 46, 302-312. http://dx.doi.org/10.4314/sajas.v46i3.10.