Check for updates

Genetic diversity and population structure of non-descript cattle in South African smallholder systems

M. P. Ramoroka[1,2]*, M. D. MacNeil[1,2,3], F. W. C. Neser[1], S. F. Lashmar[1] and M. L. Makgahlela[1,2]

[1]Department of Animal Science, University of the Free State, Bloemfontein, South Africa, [2]Agricultural Research Council, Animal Production, Irene, South Africa, [3]Delta G, Miles, MT, United States

The genetic background and characteristics of South African smallholder cattle populations remain largely unknown. These cattle exhibit remarkable adaptability to challenging environments with minimal inputs from farmers, making them a valuable genetic resource for sustainable farming. This study aimed to genetically characterize non-descript cattle kept in smallholding systems using single-nucleotide polymorphism (SNP) markers. A total of 188 non-descript smallholder beef cattle were sampled from seven South African provinces; Eastern Cape (n = 27), Free State (n = 28), Gauteng (n = 13), KwaZulu-Natal (n = 29), Limpopo (n = 34), North West (n = 44) and Northern Cape (n = 10). In addition, samples were obtained from Afrikaner (n = 42), Bonsmara (BON; n = 46), Boran (n = 20), Brahman (n = 96), Drakensberger (n = 25), Hereford (n = 31), Holstein (HOL; n = 29), Nguni (n = 59) and Shorthorn (n = 35) to serve as reference populations. Quality control of the original SNP data removed less informative animals and SNPs, which resulted in a final data set consisting of 185 animals and 119,392 SNPs. Principal coordinate analysis, ancestry, and genomic diversity statistics revealed moderate to high levels of diversity within smallholder cattle and substantial relationship with commercial beef cattle (i.e., Afrikaner, Bonsmara, Brahman, Drakensberger, Hereford, Holstein and Nguni). In North West province, there was tendency towards greater influence of Bonsmara, whereas in KwaZulu Natal the cattle were more closely related to Holstein. The smallholder populations were shown not to be unique, likely due to indiscriminate hybridization with the commercial breeds. Among the provinces, estimates of observed heterozygosity (H_O) ranged from 0.328 ± 0.001 to 0.395 ± 0.001, while expected heterozygosity (H_E) ranged from 0.326 ± 0.001 to 0.389 ± 0.000. Inbreeding levels were low, with (mean $\pm$ standard error) per-province inbreeding coefficients (F_{IS}) ranging from -0.023 ± 0.009 to 0.133 ± 0.0254. The low F_{ROH} (<0.05) across all populations indicate a more diverse population, which is less likely to express deleterious recessive traits. Estimates of the population differentiation fixation index (F_{ST}) indicated greater genetic distance between animals from KwaZulu natal and Gauteng provinces ($F_{ST} = 0.083$) and less distance between the animals from Eastern Cape and Free State provinces ($F_{ST} = 0.010$), suggesting a closer genetic relationship probably as a result of the proximity of the latter provinces and hence trans-boundary use of bulls. These findings suggest indiscriminate

crossbreeding in smallholder cattle within and across the provinces of South Africa. The results provide foundational information for the transfer of technology for targeted breeding programs to smallholder farmers.

1 Introduction

In South Africa, smallholder farmers manage approximately 40% of the national cattle population, playing a crucial role in food security and sustaining rural livelihoods (Köhler-Rollefson, 2004). These farmers predominantly rely on resilient indigenous breeds from the Sanga group, such as the Afrikaner, Nguni, and Drakensberger, which are well-adapted to the environment with harsh conditions, including nutritional, parasitic, and pathogenic challenges (Scholtz, 2010; Nyamushamba et al., 2017). The hardiness of these breeds makes them essential to smallholder production systems, contributing to both local food security and rural economic stability (Kumar, 2019; Nyamushamba et al., 2017). Furthermore, indigenous cattle help preserve genetic diversity and may hold the potential to improve commercial breeds in the future (Mapiye et al., 2019).

Unregulated crossbreeding between indigenous and exotic breeds has been a longstanding practice, resulting in the widespread emergence of non-descript cattle, which now constitute up to 66.4% of smallholder herds (Bessa et al., 2009; Mapiye et al., 2009; Scholtz et al., 2008). This shift poses significant risks to the genetic integrity of traditional breeds and threatens to erase valuable genetic resources (Khombe, 2002; Nyamushamba et al., 2017). Despite their prevalence, the genetic diversity and adaptive traits of non-descript cattle remain poorly understood, limiting their potential for strategic breeding and conservation (Scholtz et al., 2008). Inadequate knowledge of breed composition and the relationships within and between populations could lead to the loss of local genetic resources and increased risks of inbreeding depression (Rege et al., 2007; Rewe et al., 2009).

Traditionally, non-descript cattle populations were characterized by their phenotypic traits such as body size and coat color (Teneva et al., 2009). However, recent technological advances in the field of genomics, such as SNP genotyping and whole-genome sequencing, facilitated their genetic characterization, offering a more precise means of assessing genetic diversity, demographic history, and phylogenetic relationships with other local and international breeds (Gama and Bressan, 2011; Shirasawa et al., 2012; Cendron et al., 2024). This study aims to genetically characterize the non-descript cattle populations of South African smallholder farms, providing crucial insights into their genetic history and enabling the identification of population variation and breed origin.
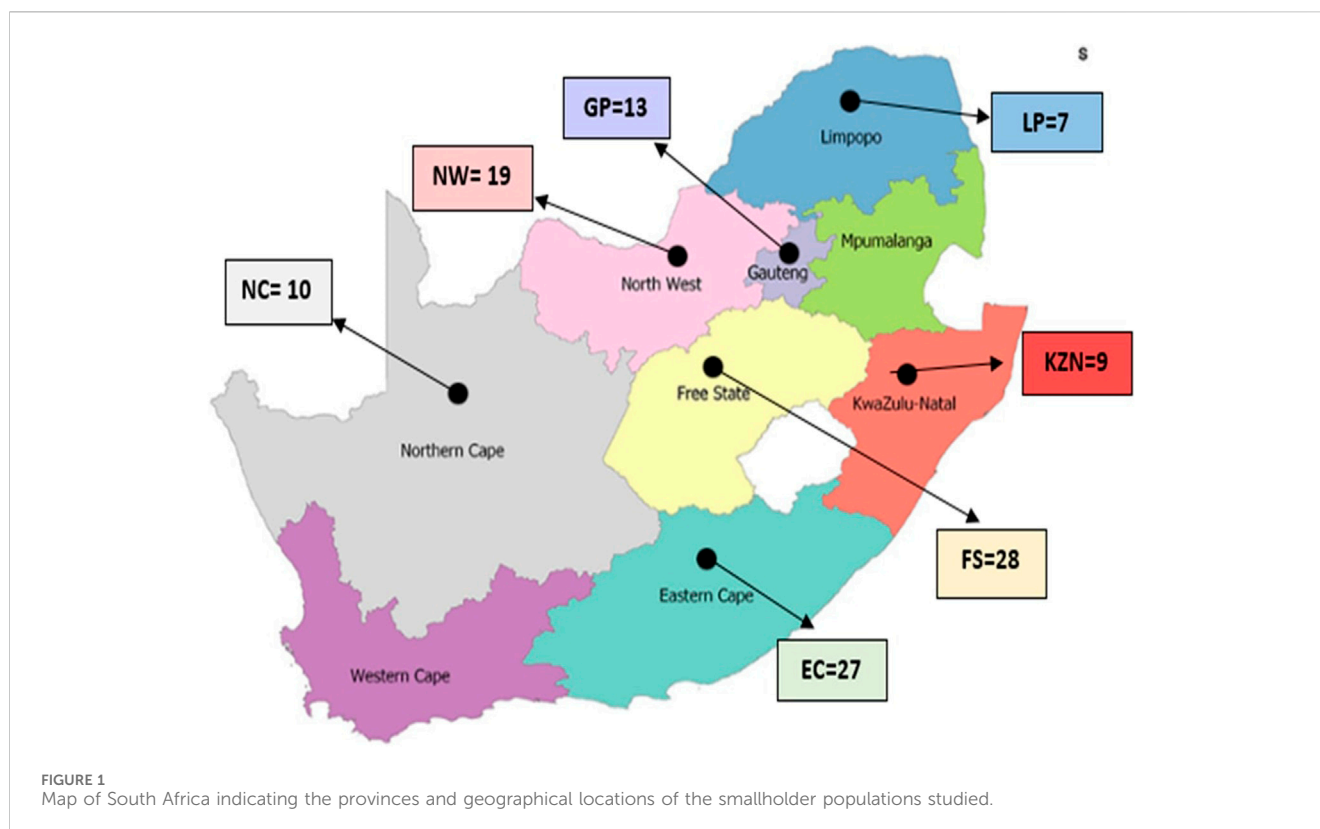
2 Materials and methods

The study was approved by the ethics committee of the Agricultural Research Council-Animal Production (Ethics number: APIC18/03).

2.1 Sampling and genotyping

A total of 188 hair samples were collected from non-descript smallholder (SHD) cattle from rural areas in the Eastern Cape (EC; n = 27), Free State (FS; n = 28), Gauteng (GP; n = 13), KwaZulu-Natal (KZN; n = 29), Limpopo (LP; n = 34), North West (NW; n = 44) and Northern Cape (NC; n = 10) provinces. Animals were randomly selected from several SHD cattle farmers within each province. Villages were identified based on consultations with local agricultural extension officers to target areas with a strong presence of non-descript cattle populations. Multiple villages per region were included to ensure a broad representation of the non-descript cattle. Farmer participation was voluntary, and efforts were made to include a diverse range of cattle, specifically targeting older animals ($\geq$10 years of age), to provide a representative sample of the SHD cattle population. Samples were not collected in the Western Cape (WC) and Mpumalanga (MP) due to the scarcity of non-descript cattle and logistical challenges, respectively. Figure 1 illustrates the geographical areas in SA where samples were collected (specific sampling sites are represented by black marks).

Genotyping was conducted at the ARC-Biotechnology Platform. The GeneSeek® Genomic Profiler™ (GGP) bovine 150 K SNP panel (featuring over 130, 000 SNPs for *Bos taurus* x *Bos indicus*) was used to obtain the genotypes from SHD populations. Genotyped samples were imaged with the Illumina® iScan Reader and a Genome Studio plug-in. The raw Illumina® Genome Studio v2.0 output files were converted into PLINK v1.9 software (Purcell et al., 2007) input files for quality check and further downstream analysis. Genotypes for Afrikaner (AFR; n = 42), Bonsmara (BON; n = 46), Brahman (BRA; n = 96), Drakensberg (DRB; n = 25), Holstein (HOL; n = 29) and Nguni (NGU; n = 59) breeds were obtained from Beef Genomics Program (BGP). Additionally, Boran (BOR; n = 20), Hereford (HFD; n = 31), and Shorthorn (SHT; n = 35) were obtained from the Web-interfaced next-generation Database dedicated to genetic Diversity Exploration (WIDDE (http://widde.toulouse.inra.fr/widde/widde/main.do?module=cattle). These reference datasets had been genotyped using the Illumina® Bovine SNP50 (BOR and HRD) and GeneSeek® Genomic Profiler™ (GGP) bovine 150 K SNP panels (AFR, BON, BRA, DRB, HOL, and NGU). In preparation for the joint analysis, the SNPs common to both the 50K and 150 K reference datasets were extracted. Since the SHD dataset was also genotyped using the bovine 150 K SNP array, the common SNPs were similarly extracted from that dataset as well. A total of 49,564 common SNPs were identified between the 50K and 150 K datasets. The reference and SHD population datasets were merged using the--bmerge command in PLINK v1.9 software (Purcell et al., 2007).

FIGURE 1
Map of South Africa indicating the provinces and geographical locations of the smallholder populations studied.

2.2 Data editing

Data were edited with PLINK v1.9 (Purcell et al., 2007) to discard low-quality markers and samples. Individuals with sample call rate less than 90% were removed as were non-autosomal, unmapped, and duplicated SNPs. In addition, individual SNP were required to have call rate greater than 95%, minor allele frequency (MAF) greater than 5%, and to exhibit Hardy-Weinberg equilibrium (P > 0.0001). After editing, the final SHD data set consisted of 185 individuals and 119,392 SNPs was retained for statistical analysis.

2.3 Analysis of molecular variance and genetic differentiation

The minor allele frequency (MAF), observed heterozygosity (H_O), expected heterozygosity (H_E), and inbreeding coefficient (F_{IS}) were estimated for the study populations using the PLINK v1.9 software (Purcell et al., 2007; www.cog-genomics.org/PLINK/1.9/). The R-package detectRUNS (Biscarini et al., 2018) was employed to detect the runs of homozygosity (ROH), which are indicative of elevated homozygosity due to mating of closely related individuals (Zeitler and Gilbert, 2024). Several parameters were used to define ROH: (i) a minimum ROH length of 500 kb, (ii) a homozygous overlap window proportion of 0.05, (iii) a minimum of 100 consecutive SNPs per ROH, (iv) a minimum SNP density of one SNP per 100 kb, (v) a maximum gap of 100 kb between consecutive homozygous SNPs, and (vi) a tolerance of up to two SNPs with missing genotypes and a maximum of one heterozygous SNP within an ROH, as described by Xu et al. (2019). The ROH were analyzed to estimate the molecular inbreeding coefficient based on ROH (F_{ROH}) for each animal with R v4.2.2 (R Core Team, 2019). To detect differentiation among provinces, among individuals within provinces, and within individuals in the studied provinces, an analysis of molecular variance was conducted and pairwise genetic distances (F_{ST}) were calculated using ARLEQUIN v3.5. 2 software (Excoffier et al., 2005). The F_{ST} statistics were used to define the degree of non-random association of alleles within individuals and measure the genetic distance between the populations (Caballero and Toro, 2002).

2.4 Population structure analysis

Principal Coordinates Analysis (PCoA) was performed and a phylogenetic tree (Nei distance matrix) was generated in R v4.2.2 (R Core Team, 2019) to investigate the relationship among individuals in the SHD using the 'ape' R-package (Paradis et al., 2004; Popescu et al., 2012). The ADMIXTURE software (version 2.0; Alexander et al., 2009) was used to detect the number of ancestral populations the most appropriate number of founder populations (K) determined by cross-validation. Genesis software version 0.2.3 (Buchmann and Hazelhurst, 2012) was then utilized to visualize the ADMIXTURE results.

3 Results

3.1 Genetic diversity and differentiation

The genetic diversity parameters for the SHD population are presented in Table 1. The mean and standard error of minor allele

frequency (MAF) ranged from 0.247 ± 0.001 in animals from LP to 0.301 ± 0.001 in animals from the FS indicating a slightly greater presence of minor alleles in the FS and hence, a higher probability of heterozygous genotypes. This was supported by the higher level of observed heterozygosity (H_O) in the FS (0.395 ± 0.001) compared to LP (0.328 ± 0.001). Except for animals from the EC province, all animals from each of the other provinces exhibited gains in diversity (i.e., $H_O > H_E$), and corresponding low levels of inbreeding, with F_{IS} values ranging from −0.023 ± 0.009 in FS to 0.133 ± 0.025 in LP. The observed F_{ROH} ranged from 0.020 ± 0.004 in NW to 0.047 ± 0.006 in KZN. The F_{ROH} across provinces were further visualized in Figure 2 wherein animals from EC and KZN displayed the widest range of F_{ROH} values, with some individuals having very high inbreeding coefficients (up to 0.20), indicating substantial variation in inbreeding levels. Animals from FS, GP, LP, NC and NW showed narrow ranges of F_{ROH} values, suggesting relatively low to moderate levels of inbreeding.

The AMOVA revealed that 94% of the variation in genotypes was present within the individual animals (Table 2). Only 1% of the variation was explained by individuals within provinces, and the difference between provinces explained 5% of the variation. The high percentage of variation within individuals indicates their genetic diversity, which can be expected in outbreeding species where individuals are not highly related.

Estimates of pairwise F_{ST} are presented in Table 3. The results indicated greater genetic differentiation (pairwise F_{ST} = 0.083) between animals from the non-adjacent KZN and GP provinces compared to those from the adjacent EC and FS (pairwise FST = 0.010) provinces, suggesting some potential differentiation resulting from geographic distance.

The phylogenetic tree (Figure 3) provides valuable insights into the genetic composition and relationships between the non-descript cattle and established commercial and indigenous breeds. The non-descript cattle in the FS were more closely tied

TABLE 1 Summary of minor allele frequency (MAF), observed (H_O) and expected (H_E) heterozygosity, inbreeding coefficient (F_{IS}) and inbreeding coefficient based on ROH (F_{ROH}) for the provincial non-descript cattle.

Populations	Number of individuals	MAF ±se	H_O ± se	H_E ± se	F_{IS} ± se	F_{ROH} ± se
EC	27	0.293 ± 0.001	0.375 ± 0.001	0.382 ± 0.000	0.029 ± 0.015	0.033 ± 0.003
FS	28	0.301 ± 0.001	0.395 ± 0.001	0.389 ± 0.000	−0.023 ± 0.009	0.023 ± 0.000
GP	13	0.256 ± 0.001	0.358 ± 0.001	0.338 ± 0.001	0.073 ± 0.005	0.027 ± 0.000
KZN	29	0.258 ± 0.001	0.347 ± 0.001	0.341 ± 0.001	0.102 ± 0.022	0.047 ± 0.006
LP	34	0.247 ± 0.001	0.328 ± 0.001	0.326 ± 0.001	0.133 ± 0.025	0.0039 ± 0.004
NC	10	0.270 ± 0.001	0.368 ± 0.001	0.355 ± 0.001	0.053 ± 0.015	0.024 ± 0.002
NW	44	0.259 ± 0.001	0.355 ± 0.001	0.342 ± 0.001	0.084 ± 0.007	0.020 ± 0.004
All	185	0.271 ± 0.001	0.364 ± 0.001	0.346 ± 0.001	0.064 ± 0.014	0.025 ± 0.001

Eastern Cape (EC), Free State (FS), KwaZulu-Natal (KZN), Northern Cape (NC), Gauteng (GP), Limpopo (LP) and North West (NW).
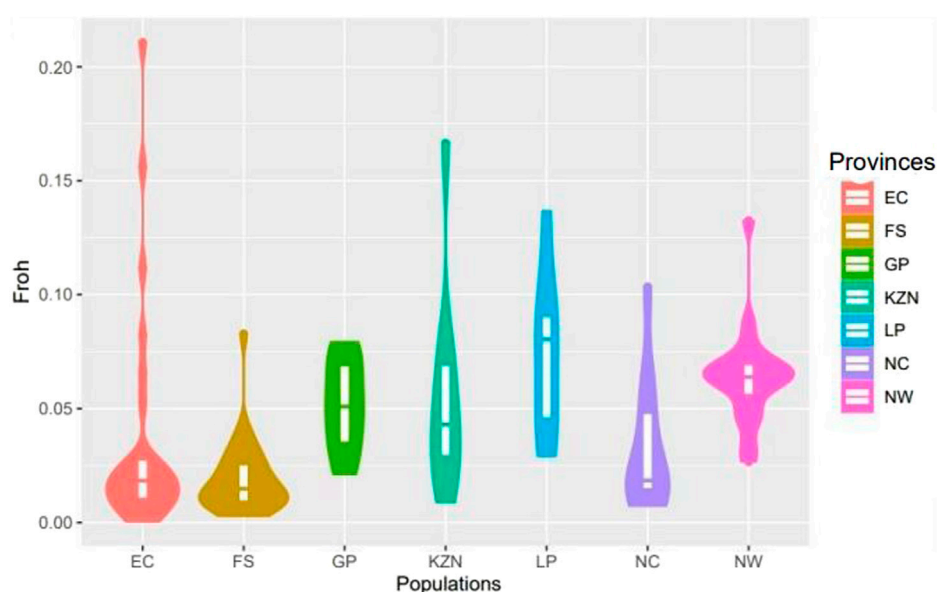

FIGURE 2
Violin plot of the inbreeding coefficients based on runs of homozygosity (FROH) in smallholder cattle populations.

TABLE 2 The analysis of molecular variance among smallholder populations.

Source of variance	Sum of squares	Variance components	Percentage variation
Among populations	188,366	593	5
Among individuals within populations	1,366,660	114	1
Within individuals	1,427,982	12,310	94
Total	2,983,007	13,017	

TABLE 3 Genetic differentiation (F_{ST}) among pairs of smallholder populations.

	EC	FS	GP	KZN	LP	NC	NW
EC	**0**						
FS	*0.010	**0**					
GP	0.051	0.046	**0**				
KZN	0.019	0.034	**0.083	**0**			
LP	0.036	0.043	0.027	0.053	**0**		
NC	0.014	0.016	0.047	0.038	0.027	**0**	
NW	0.045	0.043	0.070	0.070	0.062	0.024	**0**

SHD, population: Eastern Cape (EC), Free State (FS), KwaZulu-Natal (KZN), Northern Cape (NC), Gauteng (GP), Limpopo (LP) and North West (NW).

Bold value indicates: *Lowest pairwise FST value (i.e. least differentiated); **Highest pairwise FST value (i.e. most differentiated).

FIGURE 3
Phylogenetic tree between reference and smallholder population: Eastern Cape (EC), Free State (FS), KwaZulu-Natal (KZN), Northern Cape (NC), Gauteng (GP), Limpopo (LP) & North West (NW).
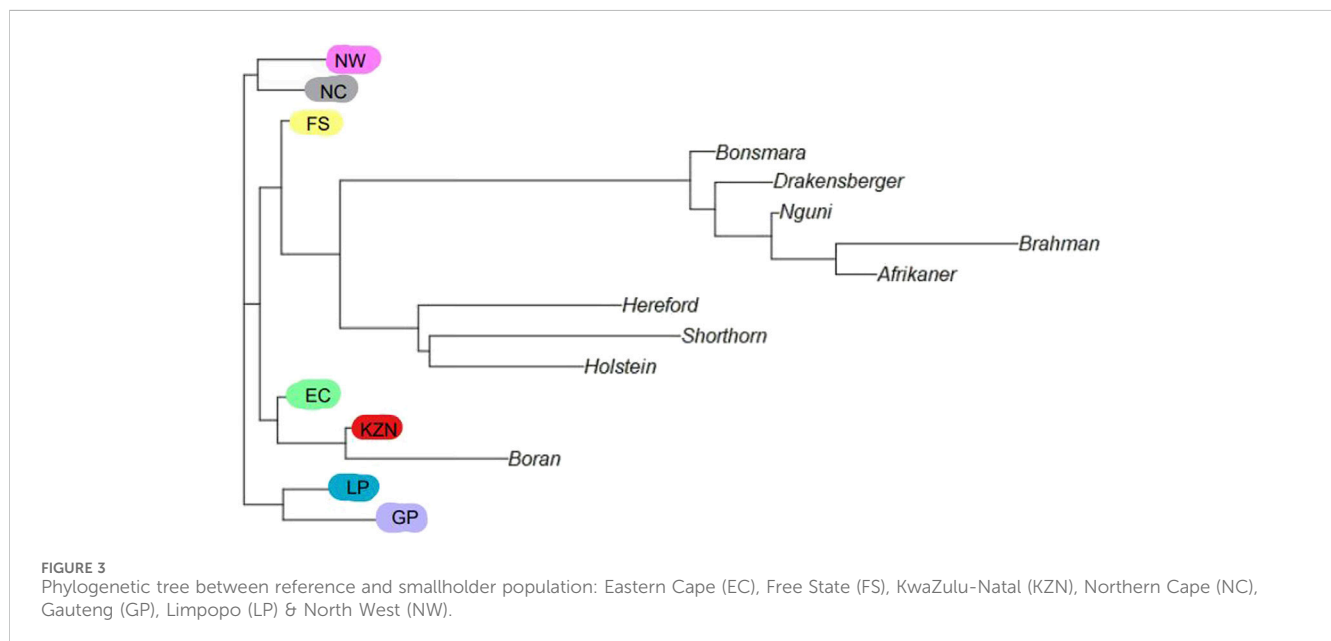
to the recognized breeds indicating that those non-descript cattle were more influenced by them than in the other provinces. The NW, and NC populations are quite distinct from the commercial breeds. While, KZN and EC show a notable connection to the Boran breed. The cattle in LP and GP form a distinct branch, separated from both commercial and indigenous breeds.

3.2 Genomic relatedness and ancestry estimation

The PCoA was used to determine the genomic relationships between all of the SHD cattle sampled within each location and across all of the provinces. The only provinces that showed more defined clustering (i.e., high relatedness between cattle within the
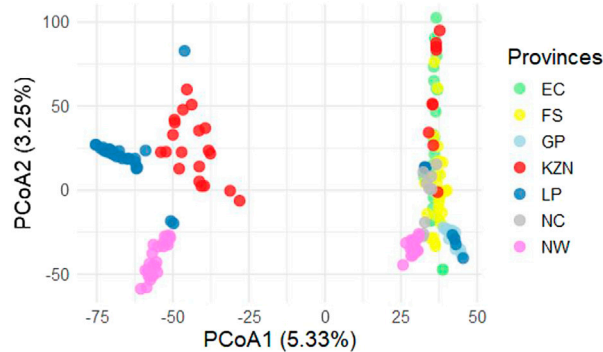
FIGURE 4
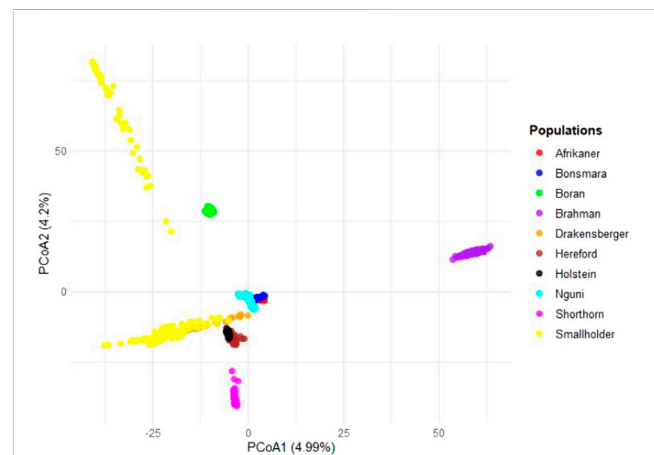Principal coordinates analysis (PCoA) for Smallholder cattle population (SHD). Eastern Cape (EC), Free State (FS), KwaZulu-Natal (KZN), Northern Cape (NC), Gauteng (GP), Limpopo (LP) & North West (NW).

province), were GP, LP, and NW; with GP and LP forming a combined cluster (i.e., showing strong genomic linkages between cattle from these provinces) (Figure 4). The SHD cattle from the remaining provinces (EC, FS, KZN, and NC) were interspersed, indicating low relatedness within provinces, however, some strong relatedness between individual animals across provinces. The first coordinate accounts for 5.33% of the total genetic variation in the population andhas an eigenvalue of 48.23, indicating that it captures a substantial proportion of the total variance in the SHD cattle. This large eigenvalue suggests that first coordinate represents a major underlying pattern or relationship within the SHD cattle, potentially linked to a dominant trait or combination of traits influencing the overall population structure. The second coordinate, with an eigenvalue of 22.53, explains notably less variance than the first. However, this substantial eigenvalue suggests that the second coordinate may capture a distinct biological or environmental factor contributing to population diversity, potentially reflecting regional or adaptive traits. The third through sixth coordinates, with eigenvalues of 10.85, 10.18, 7.65 and 7.25, respectively, reflect further elucidated structural details. These components likely represent additional, more subtle sources of variation within the SHD cattle, potentially associated with traits or specific environmental interactions. Together, these components provide a better view of the underlying genetic diversity across the provinces of South Africa with those that are more divergent along the first coordinate being less genetically related. In this case, cattle from LP and NW have distinct genetic differences compared to the others. The second coordinate represents 3.25% of the genetic variation. The EC, KZN, FS, and NC populations appear to group relatively close to each other indicating that they are genetically similar. This close clustering is interpreted to suggest these populations share a common genetic background.

Secondarily, the SHD cattle were combined into a single group and compared to recognized commercial breeds (Figure 5) to identify possible genomic relationships between non-descript cattle and the well-recognized breeds of cattle. Results showed strong genomic relatedness to the larger Sanga (*B. taurus africanus*) cluster, indicating a predominance of Sanga genetics within the non-descript cattle sampled.
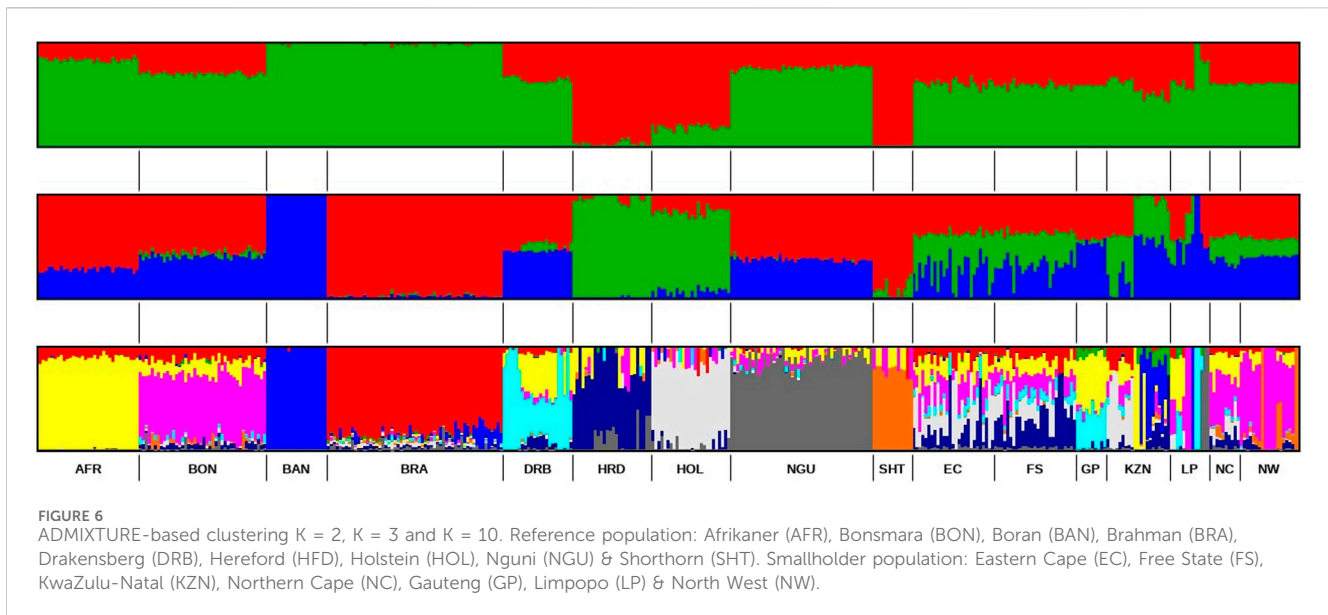

FIGURE 5
Principal coordinates analysis (PCoA) for smallholder and commercial cattle breeds.

The ancestral composition of the studied SHD population was further investigated using ADMIXTURE software (Alexander et al., 2009), which aimed to detect the maximum possible number of ancestral populations that contribute to the observed ADMIXTURE within individuals (Khanyile et al., 2015). Based on the population structure analysis (Figure 6) at K = 2, the SHD populations exhibited shared European taurine and Sanga breeds ancestry. At K = 3, the results revealed an additional ancestral contribution from *B. indicus*.

The most probable number of distinct ancestral populations was determined to be 10. The analysis revealed that the SHD cattle from EC, FS, NC, and KZN shared ancestry with Afrikaner, Bonsmara, Brahman, Drakensberg, Hereford, Holstein, and Nguni populations. Whereas, cattle from GP and LP were shown to have similar genetic backgrounds to Afrikaner, Bonsmara, and Drakensberg. Finally, the cattle held by smallholder farmers in NW were found to have Afrikaner, Bosmara, Brahman, Hereford, and Shorthorn ancestry.

Bonsmara is the most prevalent commercial beef breed in South Africa. Thus, it is not surprising that Bonsmara was at least moderately represented in the cattle of smallholder farmers across all of the provinces. Its greatest representation was in the

FIGURE 6
ADMIXTURE-based clustering K = 2, K = 3 and K = 10. Reference population: Afrikaner (AFR), Bonsmara (BON), Boran (BAN), Brahman (BRA), Drakensberg (DRB), Hereford (HFD), Holstein (HOL), Nguni (NGU) & Shorthorn (SHT). Smallholder population: Eastern Cape (EC), Free State (FS), KwaZulu-Natal (KZN), Northern Cape (NC), Gauteng (GP), Limpopo (LP) & North West (NW).

TABLE 4 Shared ancestry (%) between the SHD cattle populations and existing beef breeds.

	AFR	BON	BAN	BRA	DRB	HRD	HOL	NGU	SHT
EC	15	18	1	11	6	16	31	1	1
FS	16	17	0	10	9	30	16	1	1
GP	42	24	0	2	30	1	0	0	0
KZN	16	19	22	7	1	12	17	0	7
LP	7	48	4	10	13	0	6	12	0
NC	23	40	0	6	3	10	12	1	3
NW	7	67	0	3	1	4	1	0	19

Reference population: Afrikaner (AFR), Bonsmara (BON), Boran (BAN), Brahman (BRA), Drakensberg (DRB), Hereford (HFD), Holstein (HOL), Nguni (NGU) and Shorthorn (SHT). Smallholder, population: Eastern Cape (EC), Free State (FS), KwaZulu-Natal (KZN), Northern Cape (NC), Gauteng (GP), Limpopo (LP) and North West (NW).

NW cattle population wherein it exhibited 67% shared ancestry followed by the LP and NC populations with 39% and 40% shared ancestry, respectively (Table 4). While the cattle held by smallholder farmers in GP shared 30% ancestry with the Drakensberger and 42% with the Afrikaner. In EC and KZN the cattle of smallholders showed 31% and 17% shared ancestry with the Holstein, respectively. In FS the cattle held by smallholder farmers exhibited 30% ancestry with Hereford, along with 15% contributions from both the Afrikaner and Bonsmara breeds.

4 Discussion

In this study, we performed a genomic characterization of non-descript cattle in South African smallholder production systems investigating their genomic composition, level of genetic diversity, and the relatedness within and between the populations in different provinces of South Africa. The purported ability of indigenous breeds to adapt to a changing environment may be a result of their genetic diversity, which further assists with improvement and conservation efforts (Chokoe et al., 2020). Information on both

internal (variation in genetic makeup within a single population) and external (variation in genetic makeup between different populations) genetic diversity and population structure becomes crucial when designing genetic improvement and conservation strategies at the national and international levels (Tittensor et al., 2014). To maximize the productivity of smallholder systems, Ojango et al. (2014) argued that the use of genomic data to assess genetic diversity and population structure could produce optimal recommendations considered by farmers. Additionally, it provides smallholder industries in developing regions an opportunity to participate in genetic enhancement programs (Maake, 2020).

According to Hlophe (2011), genetic variation within populations is crucial to allow individuals to adapt to changing environments. Furthermore, Thornton (2010) suggested that the observed genetic variation presents an opportunity to implement genetic improvement programs for smallholder farmers. The SHD cattle had MAF averaging 0.27 across all provinces. Makina et al. (2015) observed MAF for overall chromosomes was 0.25 for Afrikaner, 0.26 for Nguni, 0.27 for Drakensberger and 0.26 for Bonsmara; results which are similar to those found for the

smallholder cattle in this study. According to Botero et al. (2014), the potential for the animal to adapt to environmental changes in the future will decline as the level of heterozygosity decreases and alleles are lost. The cattle sampled in this study have demonstrated the potential for long-term survival (over 10 years) and adaptation to harsh environmental conditions. The genetic diversity within the cattle held by smallholder farmers was measured based on observed (H_O) and expected (H_E) heterozygosity and the inbreeding coefficient (F_{IS}). The average mean H_E observed in the current study was lower than reported for stud animals of the SA Drakensberger (H_E = 0.36) and the composite Bonsmara (H_E = 0.37; Van Marle-Köster et al., 2021), as well as for non-descript dairy cattle in South Africa (H_E- = 0.4; Maake, 2020). However, it was higher than the H_E of 0.19 observed for non-descript cattle in Bangladeshi (Bhuiyan et al., 2021) and the 0.34 reported for Nguni cattle (Van Marle-Köster et al., 2021). Since the observed heterozygosity (Ho) is slightly greater than the expected heterozygosity (H_E) in SHD populations ($H_O > H_E$), this suggests a gain in diversity. Thus, the cattle held by smallholder farmers have a relatively high level of genetic diversity, which could indicate a more diverse gene pool or a different breeding history which can be important for long-term adaptability and resilience. However, the difference between H_O and H_E is very small, indicating only a minor gain in diversity, and the populations are likely close to Hardy-Weinberg equilibrium. According to Sharma et al. (2016), the difference between H_O and H_E is due to non-random mating among the individuals of the population. Higher heterozygosity contributes to greater genetic diversity, which enhances the capacity of cattle to adapt to fluctuating environmental conditions (Kanaka et al., 2023). The increased heterozygosity observed in the smallholder cattle provides a genetic basis for the advantages of heterosis. Heterosis refers to the phenomenon where crossbred individuals exhibit superior performance, particularly in traits like growth rate, fertility, and adaptability (Falconer and Mackay, 1996). This effect is important in low-input, smallholder systems where environmental stressors such as limited feed, extreme climates, and disease pressures are prevalent. The study's results showed low inbreeding levels (F_{IS} and F_{ROH}), supporting the hypothesis that extensive crossbreeding reduces the risk of inbreeding, as suggested by Radhika et al. (2018). The inbreeding coefficient derived from ROH (F_{ROH}) is, on average, 0.025 ± 0.001 across all of the cattle from smallholder farmers. Since F_{ROH} represents the probability of regions being identical by descent (Kardos et al., 2018), the combination of low F_{IS} and F_{ROH} values suggests that SHD populations have large effective population sizes. However, the cattle from EC and KZN include several outliers with higher F_{ROH} values, indicating that certain individuals are substantially inbred. The variability in F_{ROH} highlights opportunities to exploit heterosis by identifying and crossing genetically diverse individuals. In contrast, farmers whose cattle have less range in F_{ROH}, such as those in the FS, GP, LP, NW and NC, may benefit from introducing new genetic material to mitigate the risk of future genetic bottlenecks and maintain genetic diversity. Moreover, heterosis can mitigate the deleterious effects of inbreeding, a common concern in isolated or small populations (Söderquist et al., 2020). The introduction of genes from diverse breeds, including Sanga and taurine cattle,

creates a buffer against the negative effects of inbreeding, contributing to the overall genetic health and adaptability.

The SHD cattle had a moderate level of genetic diversity across the provinces of South Africa, which may be an indication of extensive crossbreeding. According to Ojango et al. (2014), the admixture of diverse populations and natural selection for adaptation is supported by high levels of genetic diversity. Smallholder animals are genetically diverse and can adapt to various production systems, and the value of diversity in this sector has been strongly emphasized (Morrison, 2007; Mdladla, 2016). Understanding the genetic diversity of the SHD cattle population is crucial because it helps identify the range of genetic traits within the population. This diversity is important for achieving maximum productivity, as it allows for the selection of cattle with desirable traits such as disease resistance, growth rate, and adaptability to local environments. It ensures that breeding programs maintain a healthy and sustainable population, prevent inbreeding, and support the long-term viability of cattle populations in smallholding systems. The evaluation of genetic diversity ensures that specific breeds are attained for rural development initiatives and genetic improvement programs (Chagunda et al., 2018).

The genetic variation is distributed within and between SHD cattle populations. The highest percentage variance (94%) of the SHD population was contributed by (and could be explained by) the genetic diversity within individuals (i.e., the genome-level diversity per animal) with only small percentages (5% and 1%, respectively) accounting for among-population (i.e., across provinces) and among-individuals within populations (i.e., within provinces). This sub-division of the explained variation was in concordance with observations by Groeneveld et al. (2010). The variation among SHD populations could be attributed to their geographical area, farmer's preferences, socio-cultural aspects, natural processes of mutation, and adaptation to the different ecological zones of SA. The variation within province specific SHD populations is supported by the population structure analysis and indicates a well-diverse population with a low risk of inbreeding. Although positive inbreeding was observed in SHD cattle populations (EC, GP, KZN, LP, NC, and NW), the AMOVA results show that the population can be improved through breeding selection.

In this study, the genetic relatedness (PCoA), ancestry (ADMIXTURE), and genetic diversity estimators concordantly supported the genetic diversity in and amongst South African SHD cattle populations. According to Pritchard et al. (2000), population structure examines the variation within a population by considering factors such as individual differences, migration patterns, and the occurrence of crossbreeding. The population structure of individuals from EC, FS, KZN, and NC explains that most of the animals are raised in communal areas with no formalized breeding programs, where individuals mate naturally and randomly (Scholtz et al., 2008). The majority of smallholder farmers prefer to breed with bulls from indigenous breeds due to their low maintenance requirements and ability to adapt to local conditions, survive, and reproduce (Nyamushamba et al., 2017). Most cattle in the rural areas (smallholder farmers) are non-descript and crossbreds of local indigenous breeds in SA (Palmer and Ainslie, 2006). The SHD cattle populations in this study showed shared ancestry with predominantly Sanga breeds (*B. taurus africanus*) (Afrikaner, Bonsmara, Drakensberg, and Nguni) but also taurine

(Hereford, and Holstein) and indicine (Brahman) breeds. The findings of this study align with those reported by Chagunda et al. (2018), who observed similar patterns in different indicine breeds among crossbred cattle in smallholder herds in Rwanda and Tanzania (Mujibi et al., 2019), as well as in India (Ahmed et al., 2019). In the North West (NW) province, Bonsmara cattle are the most readily available breed for smallholder (SHD) and emerging cattle farmers, who primarily use them in unstructured crossbreeding practices. The breed's popularity is driven by its ability to produce superior replacement heifers, regardless of whether they are bred with a Bonsmara bull or another breed. This availability and adaptability make Bonsmara a preferred choice, contributing to its widespread use and success in smallholder farming systems. Since the establishment of the breed society in 1964, the Bonsmara has become the most popular cattle breed in South Africa (Steyl, 2024), demonstrating significant growth over the past 55 years. The appearance of the Holstein breed in the smallholder dairy (SHD) cattle populations of KZN, EC, and FS is attributable to the regional distribution of major milk producers in SA. These producers are categorized into four regions based on the prevailing production systems and the markets they serve, namely, KZN, WC, FS, and EC (Gertenbach, 2006). While, indigenous SA cattle breeds, such as the Nguni and Afrikaner, exhibit low milk productivity. Consequently, many SHD farmers prefer to crossbreed these indigenous breeds with specialized dairy breeds like the Holstein (Marshall et al., 2019). Additionally, cattle farmers have garnered interest from indigenous breeds due to their capacity to produce and reproduce under challenging environmental conditions and their natural immunity against ticks (Mapholi et al., 2014). This genetic admixture facilitates heterosis, which can lead to improved fitness traits such as disease resistance, heat tolerance, and reproductive efficiency. Similarly, Morrison. (2007) emphasized the value of heterosis in enhancing the adaptability of cattle in smallholder systems, where natural selection pressures favor resilient animals.

The current study provided evidence of a great variety of genomic contributions from ten inferred ancestral populations (K = 10) to the genetic composition of the SHD populations, and shared ancestry with recognized commercial breeds. The results indicated low genetic differentiation between SHD cattle populations with high levels of admixed within individuals, providing evidence of gene flow and shared genomic heritage with commercial beef cattle breeds (mainly indigenous Sanga breeds). The sampled SHD cattle showed a smallholder-farmer preference for Afrikaner, Bonsmara, Brahman, Hereford, Holstein, Drakensberger, and Nguni breeds. Although the sample size in this study was limited, and cannot comprehensively represent the cattle genomic make-up of the entire provinces, population genetics estimators indicated moderate to high levels of genetic diversity, and low inbreeding incidence, for cattle in this production system. The pairwise F_{ST} matrix indicated the smallest differentiation and, hence, a close genetic linkage between the EC and FS populations compared to other pairwise provincial comparisons. This observation indicates that cattle farmers from these two provinces could be sourcing cattle from the same place, as these provinces (regions) are neighbors, and transportation costs can be kept at a minimum (Mamogobo et al., 2020). On the contrary,

the considerable genetic divergence between KZN and GP SHD populations suggests negligible gene flow that may be attributed to the physically distant between them. Geographic isolation may result in genetic divergence over time as each population adjusts to its own micro-environment. Furthermore, the coastal KZN is known to be dairy-focused (with the Holstein being the largest dairy breeds utilized). The study confirms that the genomic composition of the SA SHD cattle population is characterized by ADMIXTURE, with varying contributions from a gene pool of recognized breeds, facilitated by crossbreeding. Furthermore, the breed preference for crossbreeding is province or region-specific but across provinces, there is a tendency toward the predominant utilization of indigenous breeds, especially the Afrikaner, Bonsmara, Drakensberger, and Nguni, as well as the Brahman.

5 Conclusion

Genetic diversity and demographic structure of non-descript cattle that were distributed across seven provinces of SA in the non-commercial smallholder production systems were elucidated. This study shows that genomics plays a role in "breed-type" assignment by providing a detailed understanding of the genetic composition of cattle populations. By analyzing genomic data, it becomes possible to accurately assign cattle to specific breeds or ecotypes based on their genetic markers. In almost every province the non-descript cattle defy clear categorization, with no single breed dominating their genetic composition. Instead, these cattle are highly heterogeneous, reflecting unstructured crossbreeding among indigenous and commercial breeds. The high level of heterozygosity and low level of inbreeding observed suggests an influence of heterosis, which may contribute to the adaptability and resilience of these cattle in diverse and challenging environments. The lack of a predominant breed across provinces indicates much of the genetic variation exists among individual animals. These findings emphasize the potential of using the existing genetic diversity through informed breeding programs, focusing on enhancing key adaptive traits such as growth, disease resistance, and environmental resilience. Leveraging heterosis and genetic diversity could further optimize productivity while minimizing the risks of inbreeding. By implementing programs designed to regulate selective crossbreeding practices in smallholder systems, this issue can be mitigated, ensuring both improved cattle performance and the conservation of valuable genetic resources for future generations. By implementing programs designed to regulate selective crossbreeding practices in smallholder systems, this issue can be mitigated, ensuring both improved cattle performance and the conservation of valuable genetic resources for future generations. Introduction of community based breeding programs can help farmer to grouping and managing cattle according to their genomic profiles to enhance breeding strategies by enabling more targeted selection for desirable traits. This approach can lead to improved productivity, disease resistance, and adaptability to

local conditions, ultimately maximizing the efficiency of resource use in SHD systems.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The animal studies were approved by Agricultural Research Council-Animal Production Animal Ethics Committee (APIC18/ 03). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent was obtained from the owners for the participation of their animals in this study.

Author contributions

MR: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Visualization, Writing–original draft, Writing–review and editing, Validation. MDM: Conceptualization, Supervision, Writing–review and editing. FN: Conceptualization, Supervision, Writing–review and editing. SL: Formal Analysis, Supervision, Writing–review and editing. MLM: Conceptualization, Data curation, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing–review and editing.

Funding

Acknowledgments

Conflict of interest

Author MDM was employed by Delta G.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2025.1535730/ full#supplementary-material

References

Ahmad, S. F., Panigrahi, M., Chhotaray, S., Pal, D., Parida, S., Bhushan, B., et al. (2019). Ancestry informative markers derived from discriminant analysis of principal components provide important insights into the composition of crossbred cattle. *Genomics* 112, 1726–1733. doi:10.1016/j.ygeno.2019.10.008

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi:10.1101/gr. 094052.109

Bessa, I., Pinheiro, I., Matola, M., Dzama, K., Rocha, A., and Alexandrino, P. (2009). Genetic diversity and relationships among indigenous Mozambican cattle breeds. *South Afr. J. Animal Sci.* 39, 61–72. doi:10.4314/sajas.v39i1.43548

Bhuiyan, M. S. A., Lee, S.-H., Hossain, S. M. J., Deb, G. K., Afroz, M. F., Lee, S.-H., et al. (2021). Unraveling the genetic diversity and population structure of Bangladeshi indigenous cattle populations using 50K SNP markers. *Animals* 11 (8), 2381. doi:10. 3390/ani11082381

Biscarini, F., Cozzi, P., Gaspa, G., and Marras, G. (2018). detectRUNS: detect runs of homozygosity and runs of heterozygosity in diploid genomes. *R. package*.

Botero, C. A., Gardner, B., Kirby, K. R., Bulbulia, J., Gavin, M. C., and Gray, R. D. (2014). The ecology of religious beliefs. *Proc. Natl. Acad. Sci. (PNAS)* 111, 16784–16789. doi:10.1073/pnas.1408701111

Buchmann, R., and Hazelhurst, S. (2012). *Genesis manual*. Johannesburg, South Africa: University of Witwatersrand.

Caballero, A., and Toro, M. A. (2002). Analysis of genetic diversity for the management of conserved subdivided populations. *Conserv. Genet.* 3, 289–299. doi:10.1023/A:1019956205473

Cendron, F., Ledesma-Rodríguez, A., Mastrangelo, S., Sardina, M. T., Díaz-Herrera, D. F., Uffo Reinosa, O., et al. (2024). Genome-wide analysis of the Siboney de Cuba cattle breed: genetic characterization and framing with cattle breeds worldwide. *Front. Genet.* 15, 1302580. doi:10.3389/fgene.2024.1302580

Chagunda, M., Dusingizimana, T., Kamana, O., Cheruiyot, E., Mujibi, F. D., and Mwai, O. (2018). Use of high density single nucleotide polymorphism (SNP) arrays to assess genetic diversity and population structure of dairy cattle in smallholder dairy systems: the case of Girinka Programme in Rwanda. *Front. Genet.* 9, 438. doi:10.3389/fgene.2018.00438

Chokoe, T. C., Mdladla-Hadebe, K., Muchadeyi, F., Dzomba, E., Matelele, T., Mphahlele, T., et al. (2020). Genetic diversity of South African indigenous goat population from four provinces using genome-wide SNP data. *Sustainability* 12, 10361. doi:10.3390/su122410361

Excoffier, L., Laval, G., and Schneider, S. (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinforma. Online* 1, 47–50. doi:10.1177/117693430500100003

Falconer, D. S., and Mackay, T. F. C. (1996). *Introduction to quantitative genetics*. 4th ed. doi:10.1093/genetics/167.4.1529

Gama, L., and Bressan, M. (2011). Biotechnology applications for the sustainable management of goat genetic resources. *Small Ruminant Res.* 98, 133–146. doi:10.1016/j.smallrumres.2011.03.031

Gertenbach, W. D. (2006). Dairy farming in South Africa: production regions and systems. *South Afr. Milk. Producers' Organ. (MPO) Rep*.

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

Hlophe, S. R. (2011). Genetic variation between and within six selected South African sheep breeds using random amplified polymorphic DNA and protein markers. *Univ. Zulul. Institutional Repos.* Available at: http://hdl.handle.net/10530/611 (Accessed September 18, 2024).

Kanaka, K. K., Sukhija, N., Goli, R. C., Singh, S., Ganguly, I., Dixit, S. P., et al. (2023). On the concepts and measures of diversity in the genomics era *Curr. Plant Biol.* 33, 100278. doi:10.1016/j.cpb.2023.100278

Kardos, M., Nietlisbach, P., and Hedrick, P. W. (2018). How should we compare diferent genomic estimates of the strength of inbreeding depression? *Proc. Natl. Acad. Sci. U. S. A.* 115, E2492–E2493. doi:10.1073/pnas.1714475115

Khanyile, K. S., Dzomba, E. F., and Muchadeyi, F. C. (2015). Population genetic structure, linkage disequilibrium and effective population size of conserved and extensively raised village chicken populations of Southern Africa. *Front. Genet.* 6 (13), 13–11. doi:10.3389/fgene.2015.00013

Khombe, C. T. (2002). *Genetic improvement of indigenous cattle breeds in Zimbabwe: a case study of the Mashona group breeding scheme*. Nairobi, Kenya: AGTR Case Study. International Livestock Research Institute.

Köhler-Rollefson, I. (2004). Farm Animal Genetic Resources Safeguarding National Assets for Food Security and Trade' Deutsche Gesellschaft für Technische Zusammenarbeit. *Food Agric. Organ. U. N. Tech. Centre Agric. Rural Coop. ACP-EU*, 56.

Kumar, S. (2019). Importance of indigenous cattle as well as indigenous technical knowledge for management of indigenous cattle. *Acta Sci. Veterinary Sci.* 1 (4), 23–25.

Maake, M. E. (2020). *Genetic population genetic structure of smallholder dairy cattle herds in South Africa using SNP markers*. MSc Thesis (South Africa: University of Limpopo).

Makina, S. O., Taylor, J. F., van Marle-Köster, E., Muchadeyi, F. C., Makgahlela, M. L., MacNeil, M. D., et al. (2015). Extent of linkage disequilibrium and effective population size in four South African Sanga cattle breeds. *Front. Genet.* 6, 337. doi:10.3389/fgene.2015.00337

Mamogobo, M. D., Mapholi, N. O., Nephawe, K. A., Nedambale, T. L., Mpofu, T. J., Sanarana, Y. P., et al. (2020). Genetic characterisation of non-descript cattle populations in communal areas of South Africa. *Animal Prod. Sci.* 61 (1), 84–91. doi:10.1071/AN20030

Mapholi, N. O., Marufu, M. C., Maiwashe, A., Banga, C. B., Muchenje, V., MacNeil, M. D., et al. (2014). Towards a genomics approach to tick (Acari: ixodidae) control in cattle: a review. *Ticks Tick-Borne Dis.* 5 (5), 475–483. doi:10.1016/j.ttbdis.2014.04.006

Mapiye, C., Chikwanha, O. C., Chimonyo, M., and Dzama, K. (2019). Strategies for sustainable use of indigenous cattle genetic resources in Southern Africa. *Diversity* 11, 214. doi:10.3390/d11110214

Mapiye, C., Chimonyo, M., Dzama, K., Raats, J. G., and Mapekula, M. (2009). Opportunities for improving Nguni cattle production in the smallholder farming systems of South Africa. *Livest. Sci.* 124, 196–204. doi:10.1016/j.livsci.2009.01.013

Marshall, K., Gibson, J. P., Mwai, O., Mwacharo, J. M., Haile, A., Getachew, T., et al. (2019). Livestock genomics for developing countries: African examples in practice. *Front. Genet.* 10, 297. doi:10.3389/fgene.2019.00297

Mdladla, K. (2016). *Landscape genomic approach to investigate genetic adaptation in South African indigenous goat populations*. PhD Thesis (Durban, South Africa: University of KwaZulu-Natal).

Morrison, J. W. (2007). A guide to the identification of the natural indigenous goats of southern Africa tics. Available at: http://www.landbou.com/wp-content/uploads/03/f2297405-a93f-4399-bdb7-6f3de538d75d.pdf (Accessed July 22, 2024).

Mujibi, F. D., Rao, J., Agaba, M., Nyambo, D., Cheruiyot, E. K., Kihara, A., et al. (2019). Performance evaluation of highly admixed Tanzanian smallholder dairy cattle using SNP-derived kinship matrix. *Front. Genet.* 10, 375. doi:10.3389/fgene.2019.00375

Nyamushamba, G. B., Mapiye, C., Tada, O., Halimani, T. E., and Muchenje, V. (2017). Conservation of indigenous cattle genetic resources in Southern Africa's smallholder areas: turning threats into opportunities—a review. *Asian-Australasian J. Animal Sci.* 30, 603–621. doi:10.5713/ajas.16.0024

Ojango, J. M., Marete, A., Mujibi, F. D. N., Rao, E. J. O., Poole, E. J., Rege, J. E. O., et al. (2014). *A novel use of high-density SNP assays to optimize choice of different crossbred dairy cattle genotypes in smallholder systems in East Africa*. Vancouver, British Columbia, Canada: Proceedings, 10th World Congress of Genetics Applied to Livestock Production.

Palmer, T., and Ainslie, A. (2006). *Country pasture/forage resource profiles: South Africa*. Rome, Italy: FAO Publishing Policy and Support Branch.

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

Popescu, A.-A., Huber, K. T., and Paradis, E. (2012). Ape 3.0: new tools for distance-based phylogenetics and evolutionary analysis in R. *Bioinformatics* 28, 1536–1537. doi:10.1093/bioinformatics/bts184

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi:10.1093/genetics/155.2.945

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81 (3), 559–575. doi:10.1086/519795

R Core Team (2019). *R: a language and environment for statistical computing*. New York: Chapman and Hall/CRC: R Foundation for Statistical Computing.

Radhika, G., Aravindakshan, T. V., Jinty, S., and 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. doi:10.1080/10495398.2017.1297719

Rege, J. E. O., Ayalew, W., and Getahun, E. (2007). *Domestic animal genetic resources information system (DAGRIS)' domestic animal genetic resources information system (DAGRIS)*. Addis Ababa, Ethiopia: International Livestock Research Institute.

Rewe, T. O., Herold, P., Kahi, A. K., and Zárate, A. V. (2009). Breeding indigenous cattle genetic resources for beef production in Sub-Saharan Africa. *Outlook Agric.* 38, 317–326. doi:10.5367/000000009790422205

Scholtz, M. M. (2010). *Beef breeding in South Africa*. 2nd ed. Pretoria, South Africa.

Scholtz, M. M., Bester, J., Mamabolo, J. M., and Ramsay, K. A. (2008). Results of the national cattle survey undertaken in South Africa, with emphasis on beef. *Appl. Anim. Husb. and Rural Dev.* 1, 1–9.

Sharma, R., Kumar, B., Arora, R., Ahlawat, S., Mishra, A. K., and Tantia, M. S. (2016). Genetic diversity estimates point to immediate efforts for conserving the endangered Tibetan sheep of India. *Meta Gene* 8, 14–20. doi:10.1016/j.mgene.2016.01.002

Shirasawa, K., Ishii, K., Kim, C., Ban, T., Suzuki, M., Ito, T., et al. (2012). Development of Capsicum EST–SSR markers for species identification and *in silico* mapping onto the tomato genome sequence. *Mol. Breed.* 31, 101–110. doi:10.1007/s11032-012-9774-z

Söderquist, L., Broberg, A., Rosenberg, V., and Sletvold, N. (2020). Predicting heterosis and inbreeding depression from population size and density to inform management efforts. *J. Appl. Ecol.* 57 (8), 1459–1468. doi:10.1111/1365-2664.13643

Steyl, L. (2024). Bonsmara SA celebrates 60 years in style. *Dié Rooi Ras.* 8 (2), 7–11.

Teneva, A., Dimitrova, I., Georgiev, G., Polihronova, L., and Ivanova, K. (2009). *Molecular characterization of Bulgarian livestock genetic resources and their optimized utilization for animal production*. Vienna, Austria: FAO/IAEA International Symposium on Sustainable Improvement of animal Production and Health, 8-11 June 2009. Synopses -126-127.

Thornton, P. K. (2010). Livestock production: recent trends, future prospects. *Philosophical Trans. R. Soc. B Biol. Sci.* 365, 2853–2867. doi:10.1098/rstb.2010.0134

Tittensor, D. P., Walpole, M., Hill, S. L. L., Boyce, D. G., Britten, G. L., Burgess, N. D. A., et al. (2014). A mid-term analysis of progress toward international biodiversity targets. *Science* 346, 241–244. doi:10.1126/science.1257484

Van Marle-Köster, E., Lashmar, S. F., Retief, A., and Visser, C. (2021). Whole-genome SNP characterisation provides insight for sustainable use of local South African livestock populations. *Front. Genet.* 12, 714194. doi:10.3389/fgene.2021.714194

Xu, L., Zhao, G., Yang, L., Zhu, B., Chen, Y., Zhang, L., et al. (2019). Genomic patterns of homozygosity in Chinese local cattle. *Sci. Rep.* 9 (1), 16977. doi:10.1038/s41598-019-53274-3

Zeitler, L., and Gilbert, K. J. (2024). *Using runs of homozygosity and machine learning to disentangle sources of inbreeding and infer self-fertilization rate*. doi:10.1101/204.02.20.581206