Short Communication: A Within- and Across-country Assessment of the Genomic Diversity and Autozygosity of South African and Eswatini Nguni Cattle

Simon Frederick Lashmar (simon.lashmar@up.ac.za)
University of Pretoria https://orcid.org/0000-0003-1798-6673
Carina Visser
UP: University of Pretoria
Moses Okpeku
University of Kwazulu-Natal
Farai Catherine Muchadeyi
Agricultural Research Council
Ntanganedzeni Olivia Mapholi
University of South Africa
Esté van Marle-Köster
University of Pretoria

Research Article

Keywords: animal genetic resources, diversity, cattle, inbreeding, indigenous

Posted Date: November 10th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1060515/v1

Version of Record: A version of this preprint was published at Tropical Animal Health and Production on November 1st, 2022. See the published version at https://doi.org/10.1007/s11250-022-03373-7.

Short communication: A within- and across-country assessment of the genomic diversity and autozygosity of South African and eSwatini Nguni cattle

S.F. Lashmar[1], C. Visser[1], M. Okpeku[2], F.C. Muchadeyi[3], N.O. Mapholi[4] & E. van Marle-Köster[1]

[1] Department of Animal Science, University of Pretoria, South Africa [2] Discipline of Genetics, School of Life Sciences, University of KwaZulu Natal, South Africa [3] Biotechnology Platform, Agricultural Research Council, South Africa [4] Department of Agriculture and Animal Health, University of South Africa, South Africa

Abstract

In southern Africa, the Nguni cattle breed is classified as an indigenous and transboundary animal genetic resource that manifests unique adaptation abilities across distinct agroecological zones. The genetic integrity of various ecotypes is under potential threat due to both indiscriminate crossbreeding and uncontrolled inbreeding. The aim of this study was to assess the genetic diversity and autozygosity that exists both across countries (ES: eSwatini; SA: South Africa) and within-country (SA), between purebred stud animals (SA-S) and research herds (SA-R). Subsets of 96 ES, 96 SA-S and 96 SA-R genotyped for 40 930 common SNPs were used to study inbreeding, runs of homozygosity (ROH) and heterozygosity (ROHet) profiles as well as population structure. The highest proportion (0.513) of the 3 595 ROH was <4Mb in length, while the majority (0.560) of the 4 409 ROHet segments fell within the 0.5-1Mb length category. Inbreeding coefficients indicated low inbreeding (F_{ROH} range: 0.025 for SA-S to 0.029 for SA-R). Principal component (PCA) and population structure (K=5) analyses illustrated genomic distinctiveness between SA and ES populations, greater admixture for SA-R (mean±standard deviation proportion shared=0.631±0.353) compared to SA-S (mean±standard deviation proportion shared=0.741±0.123), and three subpopulations for ES. Overall, results illustrated that genetic distinctiveness in the Nguni resulted from both geographic isolation and exposure to different production strategies. Although no impending threat to genetic diversity was observed, further loss should be monitored to prevent endangerment of unique and beneficial indigenous resources.

Keywords: animal genetic resources, diversity, cattle, inbreeding, indigenous

Introduction

The Nguni cattle breed is one of more than 150 recognized breeds indigenous to the African continent (Mwai et al., 2015). Classified as a Sanga breed (*Bos taurus africanus*), these small to medium-framed cattle have an admixed genetic composition that is intermediate between *Bos taurus* and *Bos indicus* subspecies (Hanotte et al., 2002). Nguni cattle are known for relatively low maintenance requirements (Musemwa et al., 2010), longevity and high calving rates (Matjuda et al., 2014). Their heat tolerance (Katiyatiya et al., 2017), as well as resistance to ticks and tick-borne diseases (Muchenje et al., 2008; Mapholi et al., 2014) deem them adaptable to the diverse production environments characteristic of southern Africa.

Accompanying the settlement of several ethnic groups in different geographic and climatic regions of South Africa, distinct Sanga cattle ecotypes, phenotypically distinguishable by coat colour variations, developed over time (Bester et al., 2003) which include the Shangaan, Pedi and Nkone (van Marle-Köster et al., 2021). In the early 1930's, Nguni cattle populations were classified as non-descript and an official breed society was only established in 1986. The SA Nguni as a breed has a relatively short history of animal recording and objective selection (van Marle-Köster et al., 2021).

Currently, the breed is farmed throughout South Africa, but is most popular in the regions with higher temperature and humidity in both commercial and smallholder production systems. In rural communities Nguni cattle are milked for household consumption and play an integral role in cultural and religious ceremonies and are often only marketed when cash is required (Mapiye et al., 2019). Due to their smaller frame sizes, and relatively slower growth, Nguni weaners are undesirable and fetch lower prices on the commercial level (i.e. in feedlots; Leeuw & Jiyana, 2020). There are approximately 240 stud farmers taking part in animal recording and genetic evaluations (SA Studbook Annual Report, 2016). Genetic characterisation of SA Nguni ecotypes is limited to one study based

on microsatellite markers (Sanarana et al., 2016) and another using the 50 000-SNP genotyping panel but with limited sample size (Makina et al., 2014).

In eSwatini, Nguni are found in all six ecological regions of the country where they play a significant role in the economy of rural communities, contributing more than 70% to the livelihoods of communities dependent on livestock (Vilakati, 1994). Indigenous eSwatini Nguni cattle are noted as an important heritage of the eMaswati (people of the kingdom of eSwatini) and has been previously listed as an endangered breed (Scherf, 2000), but no genetic characterisation has been performed to date. More than two decades ago, four conservation and breeding stations were established for conservation of the eSwatini Nguni. The mandate of these farm units in 1975 has changed from conservation alone to a combined focus on conservation and improvement for beef cattle production. The improvement program consists of inter-crossing of the six different indigenous lines for multiplication purposes and up-grading of selected populations with exotic breeds that include the Angus and Brahman (FAO, 2004). Uncontrolled breeding and the approved influx of exotic genetic material for crossbreeding and up-grading of indigenous populations is considered a threat to the adapted local breeds such as the Nguni (Keller and Waller, 2002; Taberlet et al., 2008).

Advancements in SNP-based methodologies have allowed the identification of homozygosity and heterozygosity-rich regions to study both genetic uniformity and diversity. Runs of homozygosity (ROH) may be used to describe the inbreeding status (in terms of degree and age) of populations or to identify genomic regions under selection (Biscarini et al., 2020). Runs of heterozygosity (ROHet) are genomic regions where diversity might be beneficial (e.g. adaptive traits) and may indicate balancing selection events (Biscarini et al., 2020). These and other inbreeding and diversity parameters may assist in evaluating current genetic characteristics and for endangerment risk assessment of southern African Nguni populations. This study aimed to assess the genetic diversity and autozygosity that exists both across countries (ES: eSwatini; SA: South Africa) and within-country (SA), between purebred stud animals (SA-S) and research herds (SA-R).

Materials and Methods

Subsets of 96 SA purebred stud animals (SA-S), 96 SA animals from research stations (SA-R) and 96 eSwatini cattle from three governmental breeding stations (ES) were studied. DNA was extracted from hair (ES, and SA-S) and/or blood samples (SA-R). The ES and SA-R populations were genotyped with the Illumina® Bovine SNP50 version 2 genotyping panel (54 609 SNPs), while the SA-S population was genotyped with the GeneSeek® Genomic Profiler™ uHD bovine genotyping panel (141 716 SNPs). A common set of 40 930 autosomal SNPs were retained for each population after the exclusion of SNPs on non-autosomal chromosomes, unmapped SNPs and SNPs that had duplicated genomic positions. Using PLINK software (Purcell et al., 2007) exclusion of samples and SNPs was based on sample call rate (<90%), SNP call rate (<95%), low minor allele frequency (MAF<1%) and Hardy-Weinberg Equilibrium (P<0.001). Liberal linkage disequilibrium (LD) based filtering was additionally employed (r^2>0.5) before genetic relatedness and population structure analyses. For the estimation of extended homozygous and heterozygous fragments (discussed hereafter), separate data sets with no MAF nor LD filtering applied were used in concordance with guidelines provided by Meyermans et al. (2020).

Runs of homozygosity (ROH) and heterozygosity (ROHet) were identified using the R package *detectRUNS* (Biscarini et al., 2018) by executing the consecutive-SNP-based detection method (Marras et al., 2015) using data sets For ROH and ROHet, a minimum of 45 and 20 SNPs, respectively, as well as 1Mb was required to constitute a segment. A maximum number of one opposite (heterozygous for ROH and homozygous for ROHet) and two missing genotypes were allowed with a maximum inter-segment gap of 1Mb. Two coefficients of inbreeding were calculated: 1) F_{SNP}, which was a SNP-by-SNP based estimation of excess in homozygosity as implemented in PLINK and 2) F_{ROH}, which was a ROH based coefficient. The F_{ROH} coefficient was calculated as S_{ROH}/L_{GEN}, where S_{ROH} represented the summed length of ROH per animals and L_{GEN} represented the base pair length of the genome covered by SNPs.

To assess the genetic relatedness between individuals, a genomic relationship matrix (GRM) was constructed using GCTA (Genome-wide Complex Trait Analysis; Yang et al., 2011) and used in principal component analysis (PCA) to estimate eigenvectors per individual. The cross-validation (CV) procedure was used to identify the ideal number of ancestral populations (K) in ADMIXTURE software (Alexander et al., 2009); the K-value producing

the lowest CV error was considered ideal. For visualization, population structure bar plots were produced with GENESIS software (Buchmann & Hazelhurst, 2014).

Results and discussion

A total of 3 595 ROH segments were identified, ranging from 1 105 segments for ES to 1 325 segments for SA-R. The sub-totals of ROHet segments were 1 629, 1 467 and 1 313 for the ES, SA-R and SA-S populations, respectively, which was similar to per-population numbers observed in other studies (e.g. 1 702 for Montana Tropical composite; Mulim et al., 2021). The mean ROH and ROHet segment lengths were 6.47Mb versus 0.61Mb across the three populations. Albeit complex to compare (due to differing criteria for defining ROH and ROHet segments), mean lengths were similar to those reported for ROH (8.55Mb) and ROHet (0.70Mb) by Biscarini et al. (2020) for semi-feral Maremanna cattle. As illustrated in Fig. 1, the highest proportion of ROH segments (range: 0.500-0.525) were <4Mb in size, which translates to the majority of inbreeding effects occurring up to 12.5 generations ago (Howrigan et al., 2011). Similar to Biscarini et al. (2020), the 0.5-1Mb length category ranked the highest in terms of mean proportion of ROHet segments (0.560 across populations).

Fig. 1 The proportion of runs of homozygosity (ROH; a) and runs of heterozygosity (ROHet; b) within different segment length categories

The highest occurring consensus ROH and ROHet runs were identified only in the SA-R population, and are summarized in Table 1. The most prevalent consensus ROH region encompasses 24 *Ensembl* protein-coding genes assembled to the ARS-UCD 1.2 cattle assembly, whereas the ROHet consensus region contained nine. The 1.24Mb ROH region includes, among others, genes involved in inflammatory response (e.g. *HDAC3*), collagen binding (e.g. *RELL2*) and integrated stress response (e.g. *DELE1*), which causes an upregulation of genes in response to internal or environmental stressors. The 0.48Mb ROHet region includes genes that may play a role in myogenesis (e.g. *ZNF609* and *TRIP4*) and endocytosis (e.g. *CSNK1G1*).

Table 1 Description and frequency of most prevalent consensus runs of homozygosity and heterozygosity across populations

	BTA	Start (bp)	End (bp)	N SNPs	Frequency
ROH	7	52 224 595	53 468 463	11	~20%
ROHet	10	45 351 906	45 834 171	8	~19%

The inbreeding coefficients indicated low but positive levels of inbreeding. The SA-S population ranked the highest for the F_{IS} coefficient (mean±standard deviation=0.017±0.046) and the ES population the lowest (0.001±0.037), which was in concordance with higher F_{IS} estimates previously observed for populations under high selections pressures (e.g. 0.086 for Hereford versus -0.017 for Montana Tropical composite; Mulim et al., 2021). Although also low positive, the F_{ROH} coefficients ranked marginally different with the SA-R population the highest (F_{ROH}=0.029) and SA-S the lowest (F_{ROH}=0.025) and this may be due to the frequency distribution (Fig. 2). For ES and SA-S, there were few highly inbred animals with F_{ROH}>0.1 (three and five, respectively) and this could be a sampling effect or the result of increased over usage of high-impact animals (especially for SA-S). The distinct and more uniform distribution of F_{ROH} for SA-R, and lowest proportion of >32Mb segments, conveyed a lesser risk of more recent increases in inbreeding.

Fig. 2 Violin plots illustrating the distribution of F_{ROH} values per Nguni population

The PCA analysis (Fig. 3a) indicated a clear separation of ES and SA populations, with more dispersion observed within the ES population than the overall SA population, and is consistent with a longer history of methodic within-breed selection for SA overall (since 1986) compared to a longer history of indiscriminate crossbreeding with exotic breeds for ES (since 1975). The tighter cluster observed for SA-S than SA-R supports more intense exposure to artificial selection and the utilization of high-impact animals across herds. Population structure results (Fig. 3b) supported the clustering patterns depicted in the PCA. The ideal number of ancestral populations (K)

estimated, with the lowest cross-validation error (CV=0.499), for the merged Nguni data set was five. The genomic composition of the SA Nguni population was largely derived from two ancestral populations, with SA-S being more uniform (mean proportion shared=0.741) and SA-R displaying a higher degree of admixture (mean±standard deviation proportion shared=0.631±0.353). The three distinct ancestral contributors observed for ES (shared genetics=0.404 for yellow, 0.249 for pink and 0.286 for green) supported the existence of three separate breeding lines (consistent with sampling), however, the degree of admixture illustrated increased crossbreeding across lines and individual ecotypes could be at risk of genetic erosion.

Fig. 3 Genetic structure of South African cattle breeds, according to principal component analysis (a), and model-based clustering (b)

In conclusion, the genomic information proved useful to provide insight into the genomic diversity and inbreeding among the three Nguni populations. The SA and ES populations could be distinguished as separate clusters and require further investigation for potential ecotyopes. Breeding strategies for ES populations should be monitored to prevent genetic erosion of subpopulations. Higher resolution genomic profiling (e.g. with whole-genome sequencing information) of Nguni populations, will provide a more comprehensive picture of the genes included in consensus ROH and ROHet regions.

Acknowledgements

Authors wish to acknowledge the permission/consent provided by the respective breed societies to use the genotypes for this study, and funding provided by the respective funding bodies.

Author contributions

Conceptualization: E. van Marle-Köster & M. Okpeku; Methodology: S.F. Lashmar; Formal analysis and investigation: S.F. Lashmar; Writing – original draft preparation: E. van Marle-Köster, C. Visser, S.F. Lashmar; Writing – review and editing: S.F. Lashmar, C. Visser, M. Okpeku, F.C. Muchadeyi, N.O. Mapholi & E. van Marle-Köster. All authors read and approved the final manuscript.

Data availability

The datasets generated during and/or analysed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Declarations

Conflict of interest

The authors declare that they have no conflict of interest.

Ethics approval

Ethical approval was granted by the Ethics Committee of the Faculty of Natural and Agricultural Sciences, University of Pretoria, for the use of external data (EC-180000097) and written consent was provided by the University of eSwatini's UNESWA Research Board.

Funding

The genotyping of ES cattle was co-funded by the ARC-Biotechnology Platform. Genotyping of SA-R was funded by the ARC-Animal Improvement Institute and The University of Stellenboch (NRF student bursary). Genotyping of SA-S was funded through the Beef Genomic Program (BGP).

References

Alexander, D., Novembre, J. and Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Research 19, 1655–1664.

Bester, J., Matjuda, L.E., Rust, J M. and Fourie, H.J., 2003. Nguni: a case study. In: Vilakati et al. (eds) 2003. Community-based management of animal genetic resources. Rome: FAO

Biscarini, F., Cozzi, P., Gaspa, G. and Marras, G., 2018. detectRUNS: Detect runs of homozygosity and runs of heterozygosity in diploid genomes.

Biscarini, F., Mastrangelo, S., Catillo, G., Senczuk, G. and Ciampolini, R., 2020. Insights into genetic diversity, runs of homozygosity and heterozygosity-rich regions in Maremmana semi-feral cattle using pedigree and genomic data. Animals, 10(12), 2285.

Buchmann, R. and Hazelhurst, S., 2014. Genesis manual. Johannesburg: University of the Witwatersrand.

Food and Agriculture Organization (FAO) 2004. Country report on state of animal genetic resources in Swaziland: A contribution to the first report on the state of the world's animal genetic resources. http://www.fao.org/3/a1250e/annexes/CountryReports/Swaziland.pdf.

Hanotte, O., Bradley, D.G., Ochieng, J.W., Verjee, Y., Hill, E.W. and Rege, J.E.O., 2002. African pastoralism: genetic imprints of origins and migrations. Science, 296(5566), 336-339.

Katiyatiya, C.L.F., Bradley, G. and Muchenje, V., 2017. Thermotolerance, health profile and cellular expression of HSP90AB1 in Nguni and Boran cows raised on natural pastures under tropical conditions. Journal of Thermal Biology, 69, 85-94.

Keller, L.F. and Waller, D.M., 2002. Inbreeding effects in wild populations. Trends in Ecology and Evolution 17, 230-241.

Leeuw, K.J. and Jiyana, S., 2020. Feedlotting with Nguni cattle. Proceedings of the 3rd Animal Husbandry Research Symposium, North West Department of Agriculture, South Africa, 2020.

Makina, S.O., Muchedeyi, F.C., Van Marle-Koster, E., MacNeil, M.D. and Maiwashe, A., 2014. Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel. Frontiers in Genetics 5, 333. Doi:10.3389/fgene.2014.00333.

Mapholi, N.O., Marufu, M.C., Maiwashe, A., Banga, C.B., Muchenje, V., MacNeil, M.D., Chimonyo, M. and Dzama, K., 2014. Towards a genomics approach to tick (Acari: Ixodidae) control in cattle: A review. Ticks and Tick-Borne Diseases 5, 475-483.

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(11), 214.

Marras, G., Gaspa, G., Sorbolini, S., Dimauro, C., Ajmone-Marsan, P., Valentini, A., Williams, J.L. and Macciotta, N.P., 2015. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. Animal genetics, 46(2), 110-121.

Matjuda, L.E., MacNeil, M.D., Maiwashe, A., Leesburg, V.R. and Malatje, M., 2014. Index-in-retrospect and breeding objectives characterizing genetic improvement programs for South African Nguni cattle. South African Journal of Animal Science 44, 161-172.

Meyermans, R., Gorssen, W., Buys, N. and Janssens, S., 2020. How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. BMC genomics, 21(1), 1-14.

Muchenje, V., Dzama, K., Chimonyo, M., Raats, J.G. and Strydom, P.E., 2008. Tick susceptibility and its effects on growth performance and carcass characteristics of Nguni, Bonsmara and Angus steers raised on natural pasture. Animal, 2(2), 298-304.

Mulim, H.A., Brito, L.F., Pinto, L.F.B., Ferraz, J.B.S., Grigoletto, L., Silva, M.R. and Pedrosa, V.B., 2021. Characterization of Runs of Homozygosity, Heterozygosity-Enriched Regions, and Population Structure in Cattle Populations Selected for Different Breeding Goals. Research Square. DOI: 10.21203/rs.3.rs-785546/v1.

Musemwa, L., Mushunje, A., Chimonyo, M. and Mapiye, C., 2010. Low cattle market off-take rates in communal production systems of South Africa: Causes and mitigation strategies. Journal of sustainable development in Africa, 12(5), 209-226.

Mwai, O., Hanotte, O., Kwon, Y.J. and Cho, S., 2015. African indigenous cattle: unique genetic resources in a rapidly changing world. Asian Australasian Journal of Animal Science 28, 911-921.

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 analysis. American Journal of Human Genetics 81, 559-575.

SA Stud Book Annual Report, 2016. SA Stud Book Annual Report. Bloemfontein. Available online at: www.sastudbook.co.za

255 Sanarana, Y., Visser, C., Bosman, L., Nephawe, K., Maiwashe, A. and van Marle-Köster, E., 2016. Genetic
256 diversity in South African Nguni cattle ecotypes based on microsatellite markers. Tropical Animal Health
257 and Production 48, 379-385.
258 Scherf, B.D., 2000. World watch list for domestic animal diversity. Rome: FAO.
259 Taberlet, P., Valentini, A., Rezaei, H.R., Naderi, S., Pompanon, F., Negrini, R. and Ajmone-Marsan, P., 2008.
260 Are cattle, sheep and goats endangered species? Molecular Ecology 17, 275-284.
261 Marle-Köster, V., Visser, C., Sealy, J. and Frantz, L., 2021. Capitalizing on the Potential of South African
262 Indigenous Beef Cattle Breeds: A Review. Sustainability, 13(8), 4388.
263 Vilakati, D.D. 1994. National Beef Cattle Breeding Programme. In Livestock Sub-sector Review and Range
264 Survey. Swaziland. Vol.III, working paper 6. FAO, Rome, Italy.
265 Yang, J., Lee, S.H., Goddard, M.E. and Visscher, P.M., 2011. GCTA: a tool for genome-wide complex trait
266 analysis. American Journal of Human Genetics 88, 76-82.

Figures

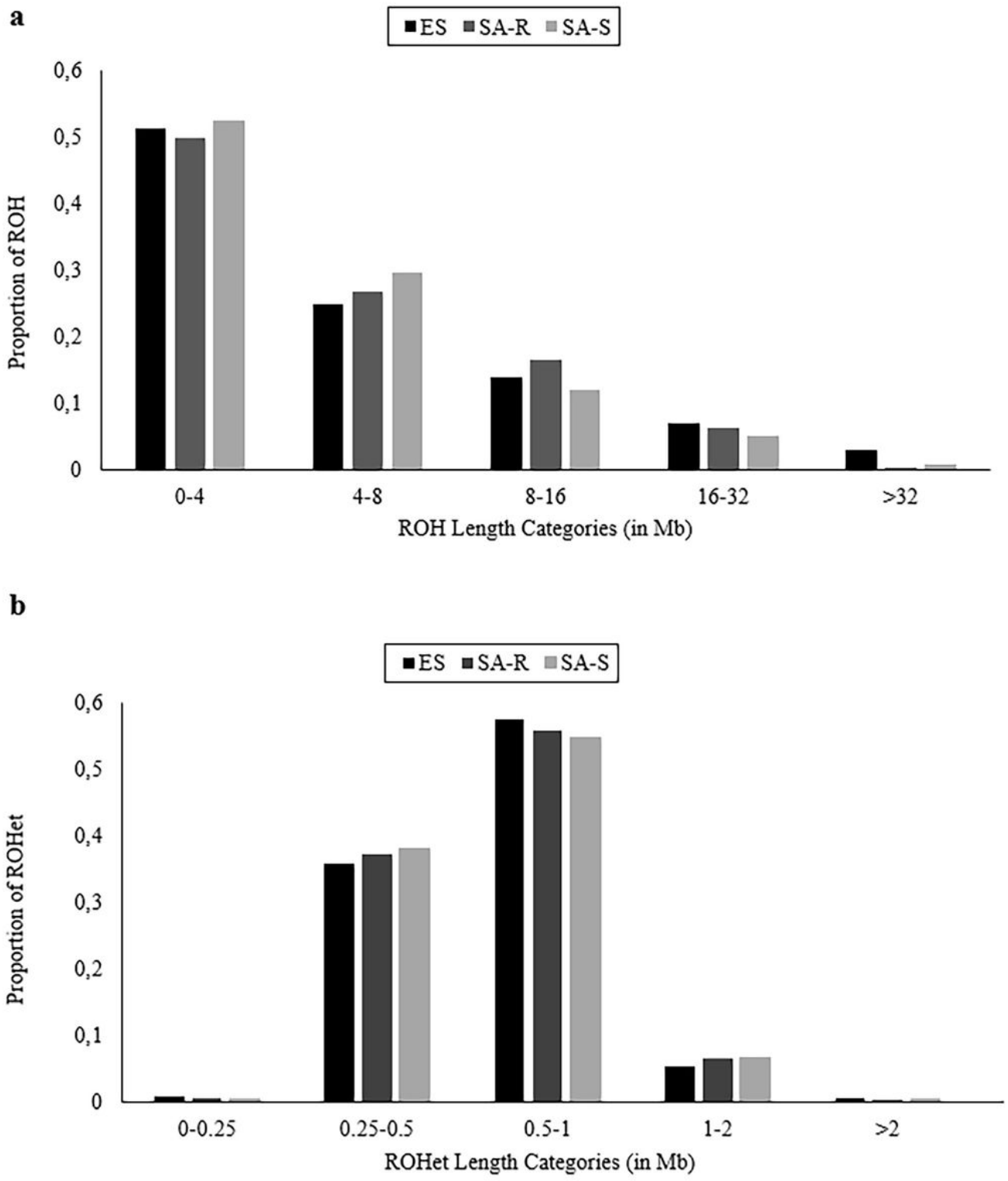

Figure 1

The proportion of runs of homozygosity (ROH; a) and runs of heterozygosity (ROHet; b) within different segment length categories

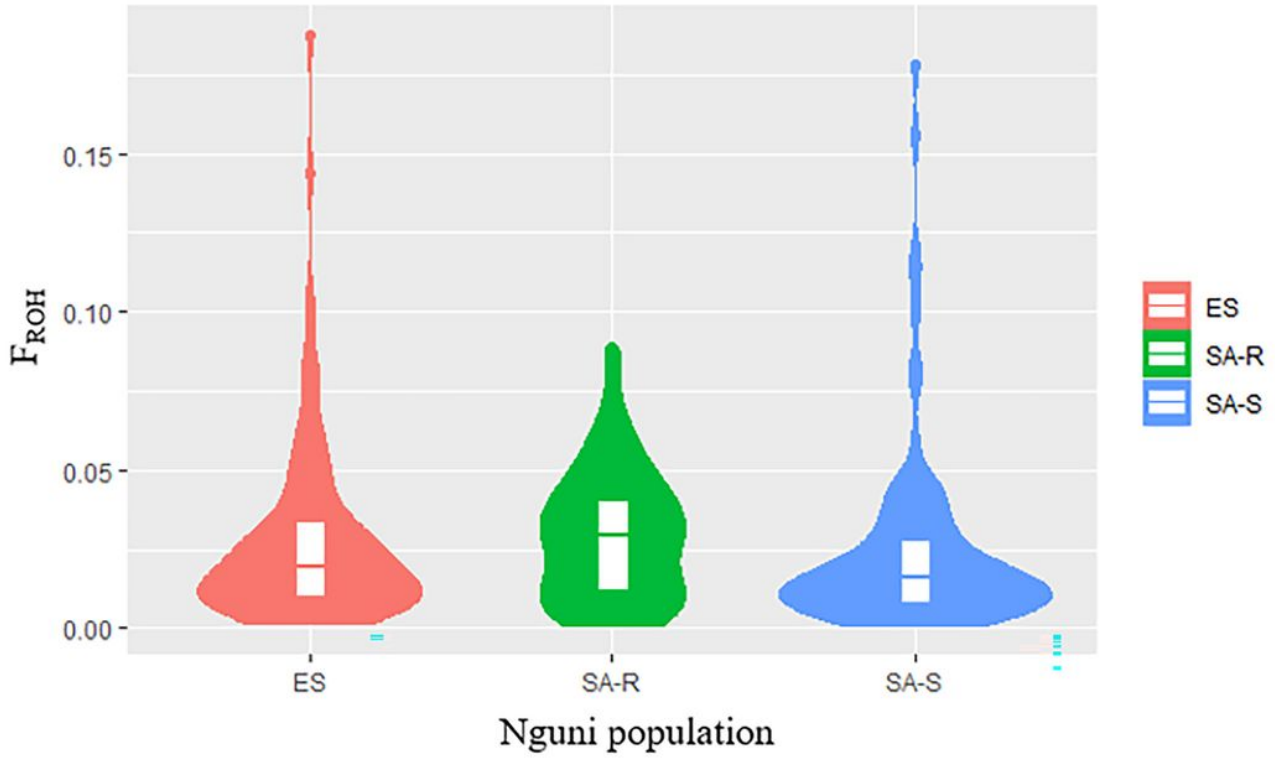

Figure 2

Violin plots illustrating the distribution of FROH values per Nguni population

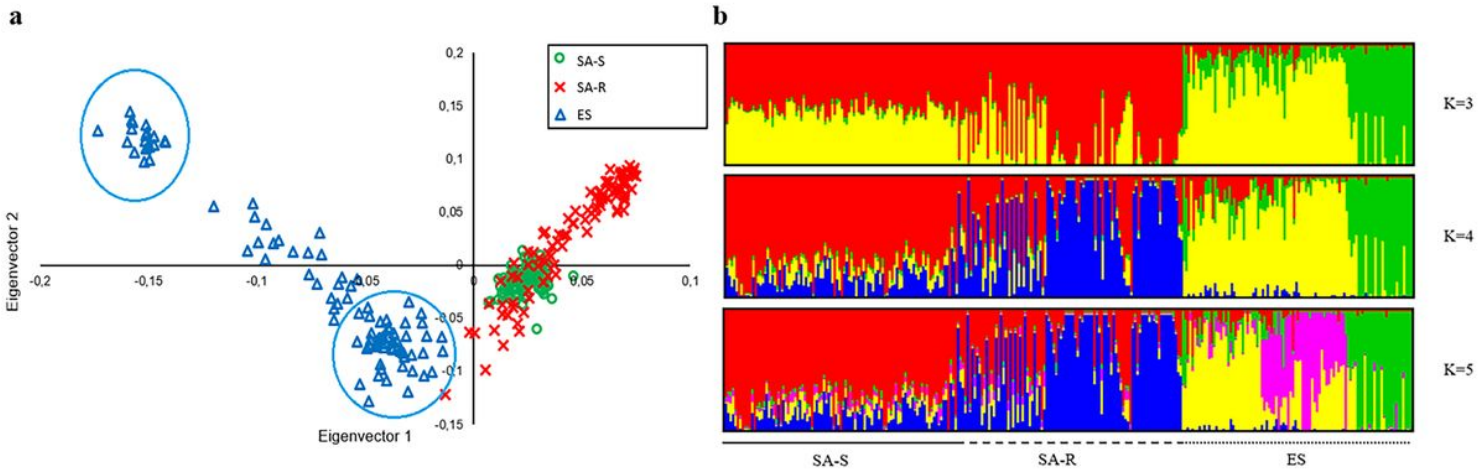

Figure 3

Genetic structure of South African cattle breeds, according to principal component analysis (a), and model-based clustering (b)