Genomic diversity and population structure of four South African indigenous Sheep

K.S. Nxumalo[1, 2], J.P. Grobler[1], K. Ehlers[1], C.B. Banga and N.O. Mapholi[3]

[1]Department of Genetics, University of the Free State, P. O Box 339, Bloemfontein, 9300, South Africa

[2]Agricultural Research Council - Animal Production Institute, Private Bag X2 Irene 0062, South Africa

[3]Department of Life and Consumer Sciences, University of South Africa, Private Bag X6, Florida 1710, South Africa

Corresponding author:.maphon@unisa.ac.za

Summary

The indigenous sheep breeds of South Africa are of importance than any other livestock within the country and they are subdivided into different breed's types which differ based on their geographical origins of origin. These types are threatened with extinction due to terminal crossbreeding effect. An understanding of the genetic diversity of types is essential prior to conservation and implementation for proper breeding programmes. 96 South African Ovine samples, comprising of 25 Pedi, 25 Namaqua, 25 Swazi and 21 Damara were genotyped using the Ovine50KSNP BeadChip assay. Genome Studio (Illumina) software was used to make the genotypes from the raw genotypic data. Genotypic data were subjected to quality control (QC) measures using PLINK software. After QC of 54 241SNPs, 48 429 SNPs remained for analysis (MAF >0.05). The outcomes of this study delivered valuable understanding regarding genetic variation within and among four indigenous sheep breeds.

Keywords: Nguni sheep, breeding programmes, crossbreeding.

Introduction

Sheep (*Ovis Aries*) is the first grazing animal to be domesticated by man for meat, milk and wool (Chessa *et al.,* 2009). South Africa was estimated to have approximately 24.5 million by National Department of Agriculture (2010). This

statistics in South Africa is made by a pool of indigenous and locally developed sheep breeds. The sheep breeds of South Africa are of importance than any other livestock within the country, this is due to their production and performance includes meat, fat and milk for food, bone and horn for implements, skins and wool for clothing and gut for containers (Almeida, 2011). However, the landrace breeds are not seen as a value – added advantage due to their low production but often used to develop hardy composites and help to conserve pure nucleus (Ramsay *et al.,* 2001). The indigenous breeds are mainly kept by South African poor resource farmers whereby there is no proper breeding management and high level of crossbreeding. Their extremely hardy nature makes them adaptable to all the varied climatic regions of South Africa from the harsh Karoo to the tropical aspects of the country and they are naturally tolerance of ticks diseases and intestinal nematodes (Nyamushamba *et al*., 2017). The aim of this study was to identify genomic population structure of four South African indigenous sheep breeds using OvineSNP50 chip.

Materials and methods

Sample collection and DNA extraction

A total of 96 indigenous sheep of South Africa were sampled based up on farm pedigree records to maximize the level of Genetic Diversity within the populations. Two populations were representing South African Nguni sheep n = 25 (Swazi and Pedi) and 2 Indigenous South African indigenous sheep were also represented Namaqua Afrikaner (n = 25) and Damara Sheep (n = 21). Genomic DNA (gDNA) was isolated from whole blood samples using the Qiagen DNeasy extraction kit (www.Qiagen.com). Genomic DNA for all samples was quantified by measuring both concentration and DNA purity using Qubit® 2.0 Fluorometer..

DNA genotyping and quality control

SNP Genotyping for the Illumina *OvineSNP50 BeadChip* was commercially completed at Agricultural Research Council-Biotechnology Platform (ARC-BTP), Onderstepoort in South Africa. This Inifinium whole genome genotyping test is designed to cross examine a large amount of SNPs at boundless levels for loci multiplexing (www.illumina.com). All genotype calls were done from the raw data using Genome Studio. Genotypic data were subjected to quality control (QC) measures using PLINK software. Individuals with a low genotyping call rate (≤90%) were removed from further analysis, while SNP positions were removed based on average call rate per SNP (Geno ≤95%), minor allele frequency (MAF≤0.05) and lack of Hardy-Weinberg Equilibrium (HWE p-value≤0.001). LD-based pruning was performed to remove SNPs that were in linkage disequilibrium with one another using PLINK's

Statistical analysis

PLINK 1.9 software was used to perform analysis including estimation of relatedness, mean expected (H_E) and observed (H_O) heterozygosity and average individual inbreeding coefficients (F_{IS}), which were calculated for LD-filtered

mapped, autosomal SNPs within and between the different population. Relatedness was calculated according to PLINK's PI_HAT value between individual pairs as the proportion identity-by-decent (IBD). The QC was performed again after four population's datasets were merged. Individuals with a sample call rate of less than 80% and SNPs that had a call rate below 95%, MAF below 5% or violated HWE ($P<0.001$) were removed from further analysis. After QC of 54 241SNPs, 48 429 SNPs remained for analysis. PLINK's simple pairwise threshold model (command: --indep-pairwise 50 5 0.2) were used to calculate heterozygosity and individual inbreeding, SNPs in linkage disequilibrium. Eigenvalues and eigenvectors were estimated using the principal component (PCA) (command:--pca) and evec file which was used to construct PCA on Genesis software, were estimated using plink2evec. PGD Spider version 2.0.8.2 was used to convert PLINK MAP and PED files to Arlequin format. Genetic differentiation for South African Nguni sheep populations was evaluated by means of Arlequin and the genetic distances between pairs of South African Nguni sheep was measured using pairwise F_{ST} obtained using Arlequin (Excoffier *et al*., 2005).

Results and Discussions

Distribution of single nucleotide polymorphisms

High levels of SNPs polymorphism were displayed when 48 429 from 54 241 SNPs remained for analysis after quality control. Expected heterozygosity (*He*) ranged between 0.28 (Namaqua Afrikaner) and 0.34 (Swazi) while observed heterozygosity (*Ho*) varied between 0.21 (Damara) and 0.38 (Swazi) in table 1. The majority individuals (96.9%) had successfully passed the animal call rate >0.1%, 2 (Namaqua Afrikaner) and 1 Damara had a call rate of <0.1% and they were excluded from further analysis.

Minor Allele Frequency and SNP polymorphism within sheep Breeds

The average MAF ranged from 0.175 (Pedi) to 0.347 (Damara) and 0.260 (merged) (Table 1). These results was comparable with the previous findings on South African Merino (Sandenbergh, 2015). Higher proportions of polymorphic SNPs were observed from indigenous Nguni types Pedi (42 078) and Swazi (43 546) as compared with the Namaqua Afrikaner (37 473), but there similar range of polymorphic SNPs was observed in Damara sheep (44 709). Sandenbergh; 2015, observed least amount of polymorphic SNPs 34 448 (Namaqua Afrikaner). The use of Ovine50KSNP BeadChip will be informative in South African indigenous sheep and they will provide useful insight for the genotyping of the indigenous Nguni sheep breeds.

Pairwise genetic distances and Phylogenetic structure SA indigenous sheep breeds

The Namaqua Afrikaner was clearly discriminated from other populations

based on the PCA (principal component analysis). PCA also clearly revealed a distinction between Pedi sheep and other Nguni types (Figure 1). Low levels of admixture were demonstrated between populations in STRUCTURE analysis (Figure 2).

It ca be concluded that the findings of this study demonstrated that the indigenous sheep populations of South Africa has a moderate level of genetic difference. The results of the Nguni sheep were comparable with the results from other indigenous sheep. Further investigation using ovine SNP chip for characterization of breed diversity is recommended to differentiate within the breed types.

Acknowledgements

The authors would like to express gratitude to the ARC- Animal Production Institute and Biotechnology Platform for samples processing. Financial support from the SASCAL and DAFF is conspicuously valued along with NRF for student support grant.

Table: 1 Measures of genetic diversity amongst populations studied

	SNP call rate<95%	Ho	He	MAF	F
Pedi	42 078	0.28	0.251	0.175	-0.118
Swazi	43 546	0.027	0.038	0.204	-0.063
Damara	44 709	0.418	0.402	0.347	-0.043
Namaqua Afrikaner	37 473	0.259	0.251	0.178	-0.04
Merged	48 429	0.315	0.332	0.26	-0.047

[1]observed (Ho) and expected heterozygosity (He), Fixation indices (F).
[2]Minor allele frequency (MAF) >0.05

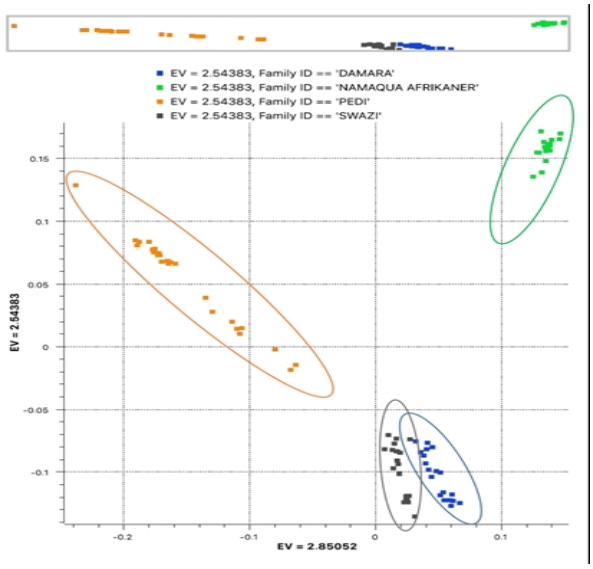

Figure 1. The genetic relationships between four populations measured using principal component analysis (PCA1 vs PCA2).

Figure 2. The genetic population structure K= 4 (Blue: Namaqua Afrikaner, Green: Pedi, Yellow: Damara & Red: Swazi).

References

Annual report. 2009/10 agriculture, forestry & fisheries. Department: Agriculture, Forestry and Fisheries. Republic of South Africa.

Chessa B, Pereira F, Arnaud F, Amorim A, Goyache F, Mainland I, Kao RR, Pemberton JM, Beraldi D, Stear MJ, Alberti A, Pittau M, Iannuzzi L, Banabazi MH, Kazwala RR, Zhang YP, Arranz JJ, Ali BA, Wang Z, Uzun M, Dione MM, Olsaker I, Holm LE, Saarma U, Ahmad S, Marzanov N, Eythorsdottir E, Holland MJ, Ajmone-Marsan P, Bruford MW, Kantanen J, Spencer, Palmarini M (2009) Revealing the history of sheep domestication using retrovirus integrations. Science 324: 532-536

Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, *1*, 47–50.

Nyamushamba, G. B., Mapiye, C., Tada, O., Halimani, T. E., & Muchenje, V. (2017). Conservation of indigenous cattle genetic resources in Southern Africa's smallholder areas: turning threats into opportunities — A review. *Asian-Australasian Journal of Animal Sciences*, *30*(5), 603–621. http://doi.org/10.5713/ajas.16.0024

Ramsay, K., Harris, L. & Kotzé, A., (2001). Landrace breeds: South Africa's indigenous and locally developed farm animals. Farm animal conservation trust, Pretoria

Sandenbergh, L. (2015). Identification of SNPs associated with robustness and greater reproductive success in the South African Merino sheep using SNP chip technology. Ph.D (Genetics). Thesis. Department of Animal Sciences, Faculty of Agricultural and Forestry Sciences, University of Stellenbosch. Stellenbosch.