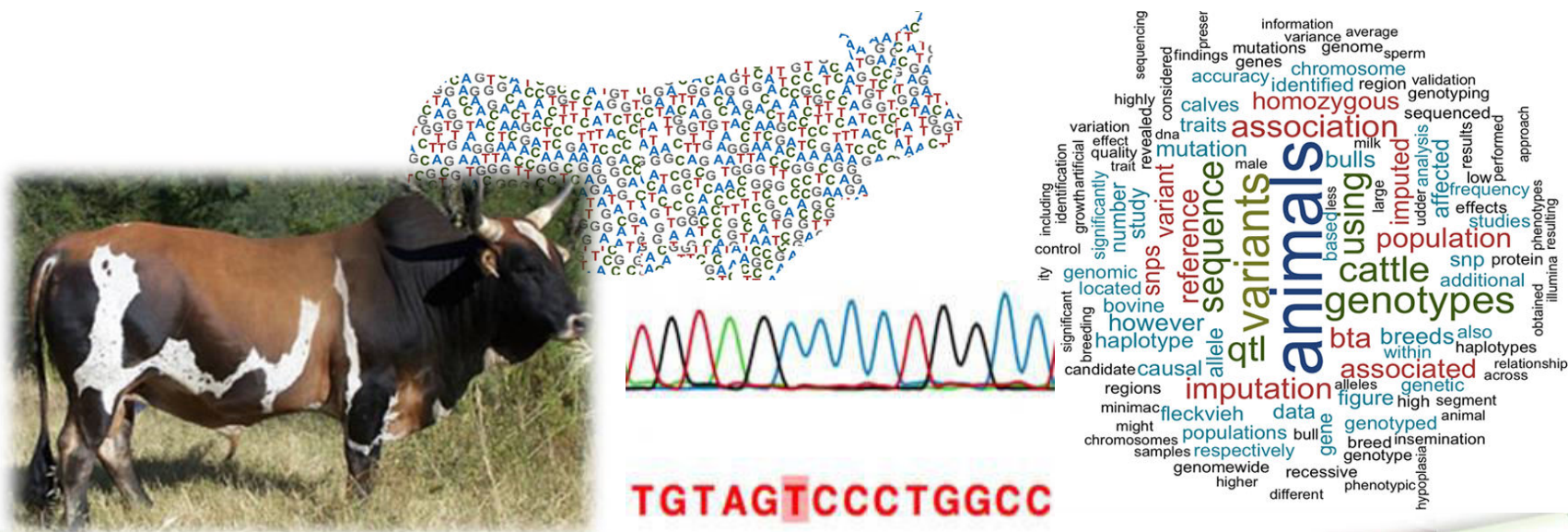


Developments in SNP genotyping and next generation sequencing in South African indigenous cattle

Dr AA Zwane: Researcher, Animal Breeding and Genetics, Agricultural Research Council-Animal Production (ARC-AP), Irene, Pretoria, South Africa

zwanea@arc.agric.za



Presentation Outline

Introduction

Genomics application

Genetic characterization of SA cattle

Limitations

Whole genome sequencing studies

Biobank Overview

Acknowledgements



Introduction

- o South Africa (SA) is a home of five indigenous cattle breeds namely, Afrikaner, Bonsmara, Drakensberger, Nguni and Tuli
- o They played important role in the history and development of the country and also in social-cultural roles in many African societies (Ramsey et al., 2000; Mwai et al., 2015).
- o They represent valuable resources vital for food security in the era of climate change
- o However, crossbreeding and inbreeding has negatively affected production and fitness traits
 - o and have contributed to loss of diversity in most cattle populations, especially indigenous breeds (Pinaar et al., 2014).
- o Quantitative breeding methods such as artificial insemination
 - o has resulted in more intense selection pressure on economic important traits
 - o Therefore, maintaining within-breed genetic diversity is essential for selection



Nguni



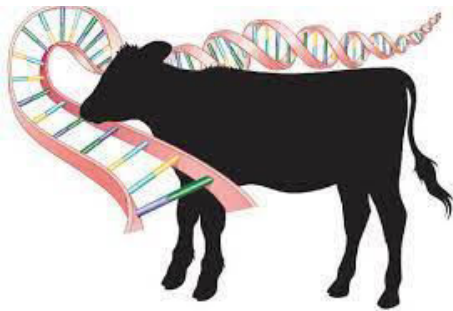
Afrikaner



Drakensberger



Introduction

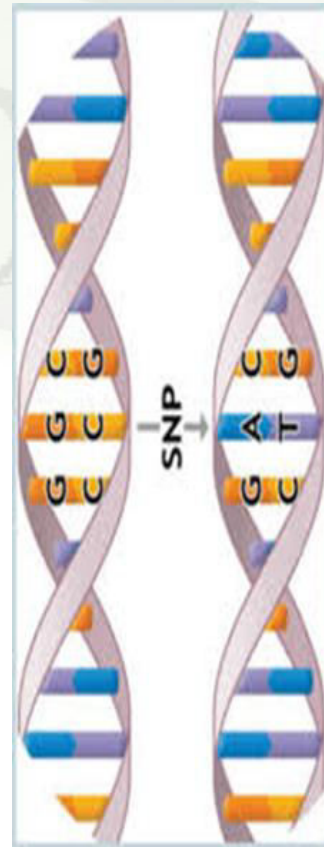


- In few years, indigenous breeds would be expected to have high 'functional' genetic diversity,
 - while being adapted to their usual local environments
- Diversity of many of African taurine and zebu are reaching marginal diversity of endangerment
 - and future diversity is expected to decrease in few years.
- Therefore, breeding objectives must be improved to avoid loss of diversity in indigenous cattle.
- Thus, extensive genetic studies are needed to improve cattle breeds against future climate changes
 - and also, to preserve current indigenous cattle as valuable genetic resources.



Introduction

- o Development in next generation sequencing (NGS) technologies has been witnessed over the past three decades,
 - o providing new prospects for development of genomic tools to enhance genetic progress in livestock production (Anderson & Schrijver, 2010)
- o These technologies have allowed the discovery of thousands of SNPs in livestock species including cattle (Nishimura et al., 2013).
- o Due to growing human population, and the global demand for livestock products
 - o These tools and developments play a central role in current and future studies in livestock genetics and improvement
- o Currently, low and high-density SNP genotyping assays, and whole genome sequencing technologies are available to perform genome-wide analyses in livestock



Genomics Application

- The analysis of NGS data
 - Allows for the discovery **new genes** and **regulatory sequences** their positions
 - It makes available large collections of **molecular markers** for genetic **diversity studies**, **GWAS** and **genomic selection**.
 - Provide breeders with an understanding of the **molecular basis of complex traits** (adaptation, production, disease traits).
 - Allows for screening mutant and germplasm collections for allelic variants in target genes.
- Thus, advances in genomics are providing breeders with new tools and methodologies that allow a great leap forward in animal breeding
 - i.e. breeding for complex traits.



Genetic characterization in South African indigenous cattle

- In 2014, Makina et al, conducted the first study to genetically characterize six South African cattle populations using Bovine SNP50K chip

frontiers in
GENETICS

ORIGINAL RESEARCH ARTICLE

published: 22 September 2014
doi: 10.3389/fgene.2014.00333



Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel

Sithembile O. Makina^{1,2*}, Farai C. Muchadeyi³, Este van Marle-Köster², Michael D. MacNeil^{1,4,5} and Azwihangwisi Maiwashe^{1,4}

¹ Agricultural Research Council-Animal Production Institute, Irene, South Africa

² Department of Animal and Wildlife Sciences, University of Pretoria, Hatfield, South Africa

³ Agricultural Research Council-Biotechnology Platform, Onderstepoort, South Africa

⁴ Department of Animal, Wildlife and Grassland Sciences, University of Free State, Bloemfontein, South Africa

⁵ Delta G, Miles City, MT, USA

Edited by:

Johann Sölkner, BOKU -University of Natural Resources and Life Sciences Vienna, Austria

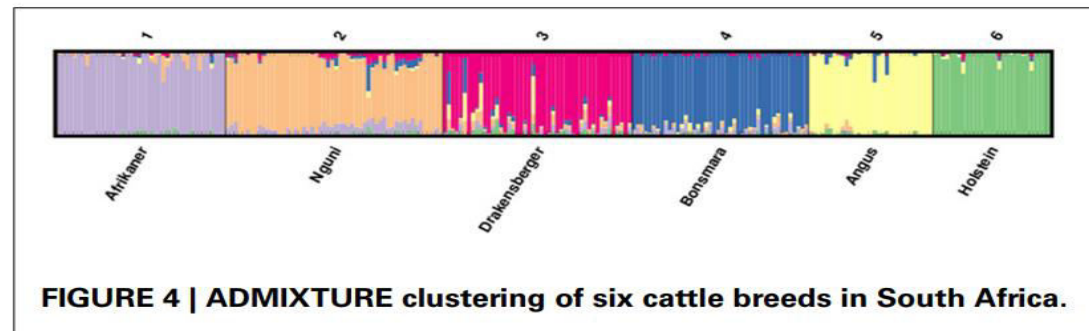
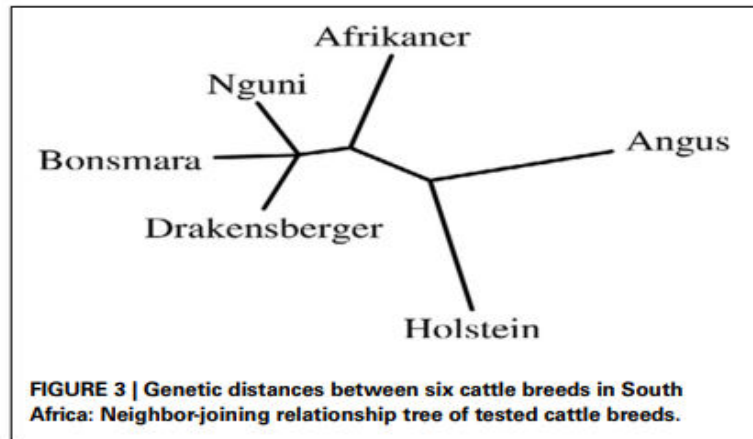
Reviewed by:

Rodolfo Juan Carlos Cantet, Universidad de Buenos Aires, Argentina

Kwan-Suk Kim, Chungbuk National University, South Korea

Information about genetic diversity and population structure among cattle breeds is essential for genetic improvement, understanding of environmental adaptation as well as utilization and conservation of cattle breeds. This study investigated genetic diversity and the population structure among six cattle breeds in South African (SA) including Afrikaner ($n = 44$), Nguni ($n = 54$), Drakensberger ($n = 47$), Bonsmara ($n = 44$), Angus ($n = 31$), and Holstein ($n = 29$). Genetic diversity within cattle breeds was analyzed using three measures of genetic diversity namely allelic richness (A_R), expected heterozygosity (H_e) and inbreeding coefficient (f). Genetic distances between breed pairs were evaluated

- They were able to differentiate between the breeds, and also to check the level of admixture among populations



- In 2016 they compared South African breeds with the other 20 breeds worldwide to determine the pattern of ancestry.

Makina et al. *Genet Sel Evol* (2016) 48:88
DOI 10.1186/s12711-016-0266-1



RESEARCH ARTICLE

Open Access



Insight into the genetic composition of South African Sanga cattle using SNP data from cattle breeds worldwide

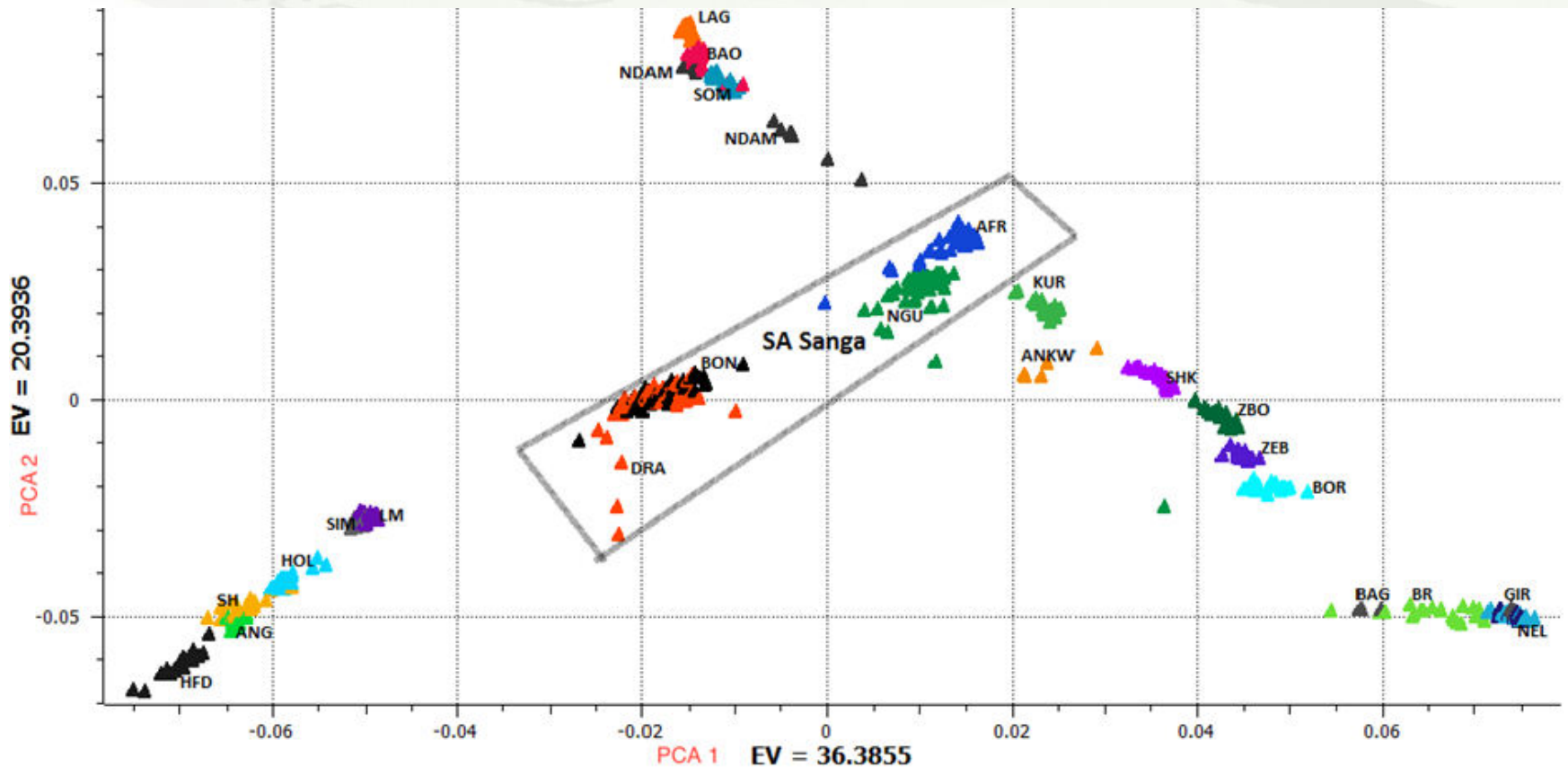
Sithembile O. Makina^{1*}, Lindsey K. Whitacre², Jared E. Decker², Jeremy F. Taylor², Michael D. MacNeil^{1,3,4}, Michiel M. Scholtz^{1,3}, Este van Marle-Köster⁵, Farai C. Muchadeyi⁶, Mahlako L. Makgahlela¹ and Azwihangwisi Maiwashe^{1,3}

Abstract

Background: Understanding the history of cattle breeds is important because it provides the basis for developing appropriate selection and breed improvement programs. In this study, patterns of ancestry and admixture in Afrikaner, Nguni, Drakensberger and Bonsmara cattle of South Africa were investigated. We used 50 K single nucleotide polymorphism genotypes that were previously generated for the Afrikaner (n = 36), Nguni (n = 50), Drakensberger (n = 43) and Bonsmara (n = 40) breeds and 6,384 European cattle breeds from 5 continents for African cattle.



- The PCA analysis revealed the distribution and the relationship between SA cattle and other cattle breeds worldwide



Makina et al., 2016. Insight into the genetic composition of South African Sanga cattle using SNP data from cattle breeds worldwide

- Another study was done to conduct a genome-wide scan for selection signatures using the bovine assay and identified 47 regions of selection and their associated genes (Makina et al., 2015), while another study was done to identify CNVs (Pierce et al., 2015), and QTLs associated with tick resistance in Nguni cattle (Mapholi et al., 2015)

RESEARCH ARTICLE

Open Access



Genome-wide scan for selection signatures in six cattle breeds in South Africa

Sithembile O. Makina^{1,2*}, Farai C. Muchadeyi³, Este van Marle-Köster², Jerry F. Taylor⁴, Mahlako L. Makgahlela¹ and Azwihangwisi Maiwashe^{3,5}

Abstract

Background: The detection of selection signatures in breeds of livestock species can contribute to the identification of regions of the genome that are, or have been, functionally important and, as a consequence, have been targeted by selection.

Methods: This study used two approaches to detect signatures of selection within and between six cattle breeds in South Africa, including Afrikaner (n = 44), Nguni (n = 54), Drakensberger (n = 47), Bonsmara (n = 44), Angus (n = 31) and Holstein (n = 29). The first approach was based on the detection of genomic regions in which haplotypes have been driven towards complete fixation within breeds. The second approach identified regions of the genome that had very different allele frequencies between populations (F_{ST}).

Results and discussion: Forty-seven candidate genomic regions were identified as harbouring putative signatures of selection using both methods. Twelve of these candidate selected regions were shared among the breeds and ten were validated by previous studies. Thirty-three of these regions were successfully annotated and candidate genes were identified. Among these genes the keratin genes (KRT222, KRT24, KRT25, KRT26, and KRT27) and one heat shock protein gene (HSP89) on chromosome 19 between 42,896,570 and 42,897,840 bp were detected for the Nguni breed. These genes were previously associated with adaptation to tropical environments in Zebu cattle. In addition, a

Limitations

- The Bovine SNP genotyping assays have shown to be biased on the analysis of the local breeds - due to less informativeness – chip design
 - shown to be adequate for studies of European taurine breeds, and are when applied to indicine or indigenous livestock breeds
- They have shown lower levels of LD and lower MAF compared to the European taurine breeds, Makina et al. (2015)
- This was also confirmed by the study of Zwane et al. (2016),
 - when identifying informative markers to discriminate between three South African (SA) indigenous cattle breeds, using the BovineSNP50 and GGP80K BeadChip



- o No substantial numbers of informative SNPs ($F_{st} \geq 0.6$) were identified among indigenous breeds.
- o Only **eleven** SNPs were validated as discriminating between SA indigenous breeds from African breeds and
- o did not possess breed specific alleles

South African Journal of Animal Science 2016, 46 (No. 3)

Genome-wide identification of breed-informative single-nucleotide polymorphisms in three South African indigenous cattle breeds

A.A. Zwane^{1,2*}, A. Maiwashe^{1,5}, M.L. Makgahlela¹, A. Choudhury³,
J.F. Taylor⁴ & E. van Marle-Köster²

¹Department of Animal Breeding and Genetics, ARC-API, P/Bag X2, Irene, 0062, ²Department of Animal and Wildlife Sciences, University of Pretoria, P/Bag X20, Hatfield, Pretoria, 0028, ³Sydney Brenner Institute of Molecular Bioscience, University of the Witwatersrand, P/Bag 3, Wits, Gauteng, 2050, ⁴Division of Animal Sciences, University of Missouri, 920 East Campus Drive, Columbia, MO 65211-5300, USA, ⁵Department of Animal, Wildlife and Grassland Sciences, University of the Free State, Bloemfontein 9300, South Africa

(Received 1 June 2016; Accepted 31 July 2016; First published online 20 September 2016)

Copyright resides with the authors in terms of the Creative Commons Attribution 2.5 South African Licence.

See: <http://creativecommons.org/licenses/by/2.5/za>

Condition of use: The user may copy, distribute, transmit and adapt the work, but must recognise the authors and the South African Journal of Animal Science.

Abstract

Access to genotyping assays enables the identification of informative markers that discriminate between cattle breeds. Identification of these markers can assist in breed assignment, improvement and conservation. The objective of this study was to identify breed informative markers to discriminate between three South African indigenous cattle breeds. Data from BovineSNP50 and GeneSeek Genomic Profiler (GGP-80K) assays were generated for Afrikaner, Drakensberger and Nguni, and were analysed for their genetic differentiation. Hereford and Angus were included as outgroups. Breeds were differentiated using principal component analysis (PCA). Single-nucleotide polymorphisms (SNPs) within the breeds were

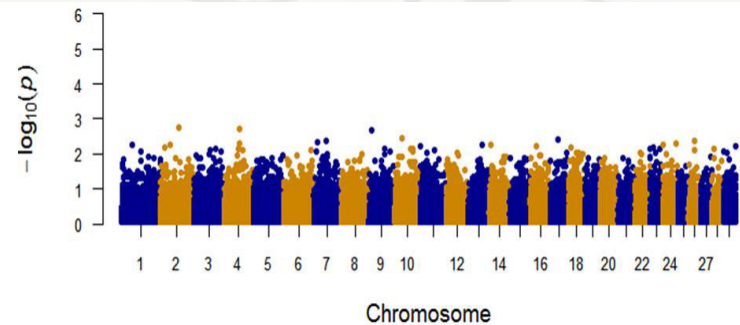


Fig 1: Manhattan plot of hapFLK statistic significance values for the analysis including indigenous and taurine ANG and HFD breeds.

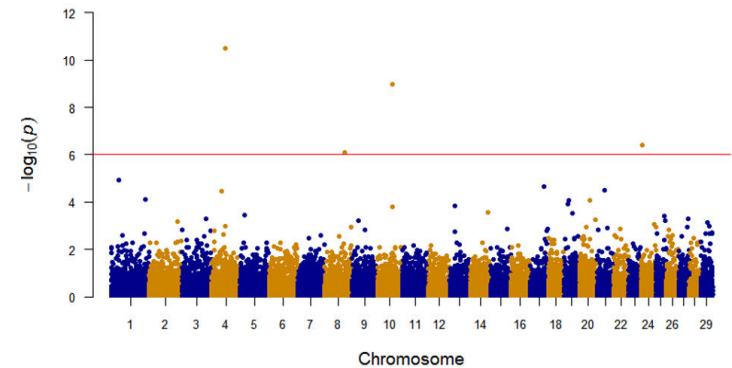


Fig 2: Manhattan plot of hapFLK statistic significance values for the analysis including the three SA breeds and three African breeds.

Whole genome sequencing studies

1. To discover novel SNPs in indigenous Afrikaner, Drakensberger and Nguni cattle breeds, and also identify selection signatures (Zwane et al., 2019)



ORIGINAL RESEARCH
published: 29 March 2019
doi: 10.3389/fgene.2019.00273



Genome-Wide SNP Discovery in Indigenous Cattle Breeds of South Africa

**Avhashoni A. Zwane^{1,2*}, Robert D. Schnabel^{3,4}, Jesse Hoff³, Ananyo Choudhury⁵,
Mahlako Linah Makgahlela^{1,6}, Azwihangwisi Maiwashe^{1,6}, Este Van Marle-Koster² and
Jeremy F. Taylor³**

¹ Department of Animal Breeding and Genetics, Agricultural Research Council-Animal Production, Irene, South Africa, ² Department of Animal and Wildlife Sciences, University of Pretoria, Pretoria, South Africa, ³ Division of Animal Sciences, University of Missouri, Columbia, MO, United States, ⁴ Informatics Institute, University of Missouri, Columbia, MO, United States, ⁵ Sydney Brenner Institute of Molecular Bioscience, University of the Witwatersrand, Johannesburg, South Africa, ⁶ Department of Animal, Wildlife and Grassland Sciences, University of the Free State, Bloemfontein, South Africa

Single nucleotide polymorphism arrays have created new possibilities for performing genome-wide studies to detect genomic regions harboring sequence variants that affect complex traits. However, the majority of validated SNPs for which allele frequencies have been estimated are limited primarily to European breeds. The objective of this study was to perform SNP discovery in three South African indigenous breeds (Afrikaner, Drakensberger, and Nguni) using whole genome sequencing. DNA was extracted

OPEN ACCESS

Edited by:



In this study:

- DNA pools were sequenced at 30X coverage
- Selections signatures were identified using the Z-transformed pooled heterozygosity (ZHp) scores

		SNPs		Indels	
Breed	No. Variants	No. SNPs	Proportion SNPs	No. Indels	Proportion indels
AFR	11,165,172	9,950,384	0.89	1,212,231	0.11
DRA	7,049,789	6,327,515	0.90	721,628	0.10
NGI	12,514,952	11,164,415	0.89	1,347,215	0.11
Total	17,243,304	15,442,314	0.89	1,908,137	0.11

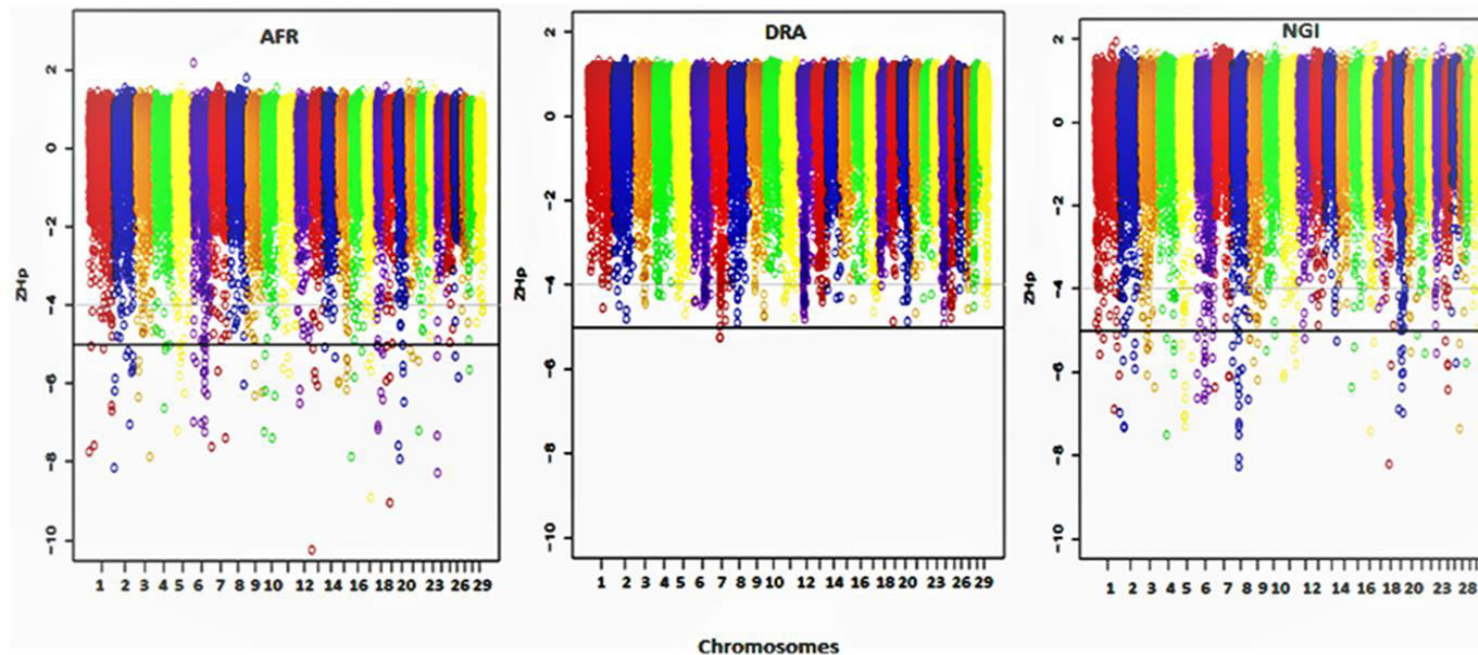
Table 1: Summary of SNPs and Indels identified in Afrikaner (AFR), Drakensberger (DRA) and Nguni (NGI)

	All Variants				SNPs			
Breed	Known variants	Novel variants	Total	Proportion Novel variants	Known SNPs	Novel SNPs	Total	Proportion Novel SNPs
AFR	9,381,545	614,536	9,996,081	0.07	8,576,732	617,296	9,194,028	0.07
DRA	6,307,154	381,743	6,688,897	0.06	5,764,627	413,795	6,178,422	0.06
NGI	10,693,999	631,412	11,325,411	0.07	9,793,635	647,269	10,440,904	0.07
Total	26,382,698	1,627,691	28,010,389	0.07 (Av)	24,134,994	1,678,360	25,813,354	0.07 (Av)

Table 2: Novel variants identified in the three breeds through comparison to 1000 Bull Genomes Project Run 6 data



Identification of selection sweeps regions



465 candidate selective sweeps (ZHp Z-score ≤ -4) were identified in AFR, DRA and NGI, and **223** regions were putative selective sweeps (ZHp Z-score ≤ -5)

Figure 3: Distribution of ZHp Z-scores across all 29 autosomes for Afrikaner (AFR), Drakensberger (DRA), and Nguni (NGI).

Functional analysis of selection signatures



South African Journal of Animal Science 2021, 51 (No. 6)



Gene-set enrichment analysis of selective sweeps reveals phenotypic traits in Nguni cattle

A.A. Zwane^{1#}, K.S. Nxumalo¹, M.L. Makgahlela^{1,3}, E. Van Marle-Koster² & N. Maiwashe^{1,3}

¹Animal Breeding and Genetics, Agricultural Research Council-Animal Production, Private Bag X2, Irene, 0062, South Africa

²Department of Animal and Wildlife Sciences, University of Pretoria, Pretoria, South Africa

³Department of Animal, Wildlife and Grassland Sciences, University of the Free State, Bloemfontein, South Africa

(Submitted 28 July 2021; Accepted 24 October 2021; Published 26 December 2021)

Copyright resides with the authors in terms of the Creative Commons Attribution 4.0 South African Licence.

See: <http://creativecommons.org/licenses/by/4.0/za>

Condition of use: The user may copy, distribute, transmit and adapt the work, but must recognise the authors and the South African Journal of Animal Science.

Abstract

Adaptation of animals to different environments is typically associated with structural and functional genomic variations. High throughput SNP genotyping and next-generation sequencing (NGS) have made it possible to study positive selection footprints and adaptation traits. Nguni is a small frame-size breed, mostly horned, and well known for being adapted to diverse South African environmental conditions. This study used previously identified selective sweeps to perform functional analysis of genes related to phenotypic characteristics in Nguni. Two hundred and sixty-four candidate selective sweeps were used for gene-set enrichment analysis in molecular functional categories (KEGG pathways) using the database for annotation, visualization, and integrated discovery (DAVID). In total, 107 genes were identified across all the chromosomes with 74 genes associated with eight phenotype queries, including fat content, milk production,

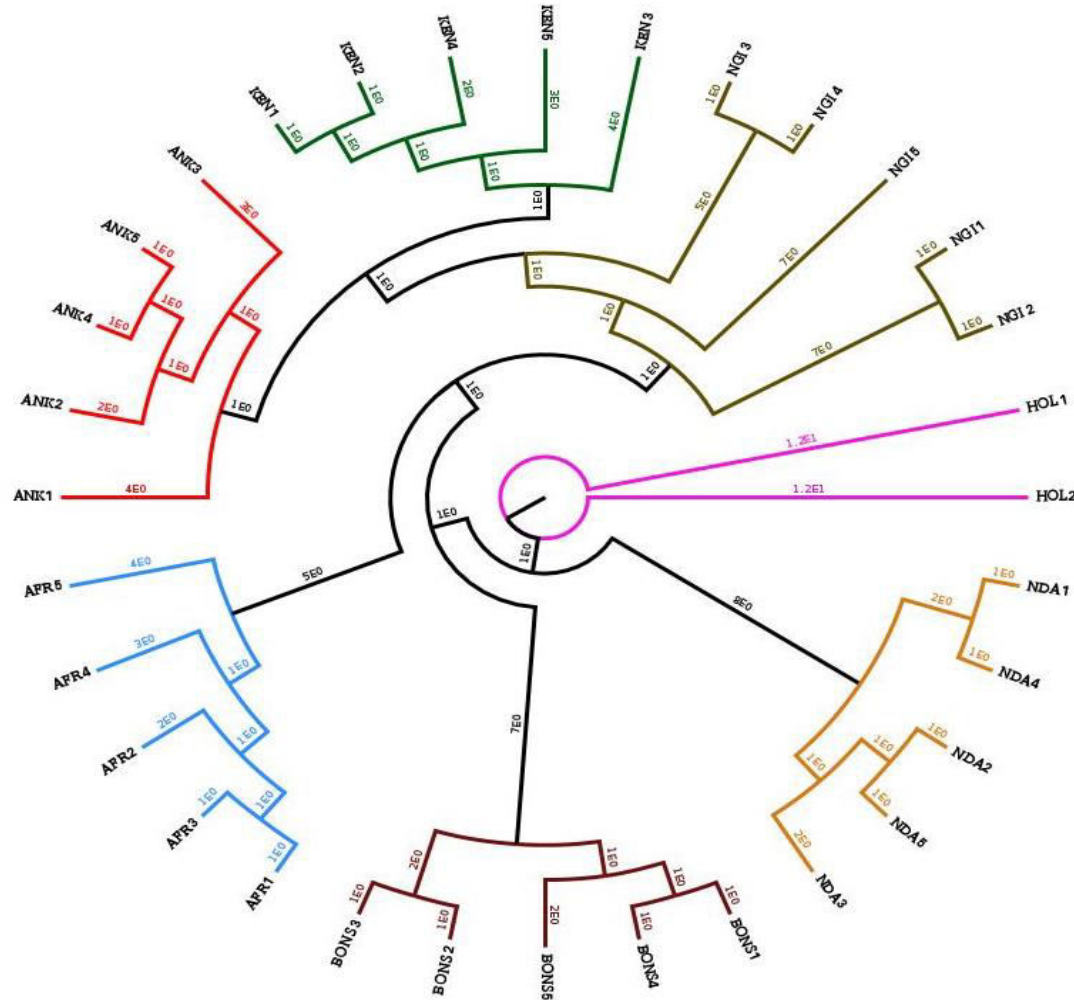


2. Current studies have focused on leveraging the WGS data to further characterize the SA indigenous breeds :

- To determine the variations among SA and African breeds, do phylogeny and introgression analysis (Mr Malima, MSc)
- And also to characterize SA breeds with other breeds in different continents (Africa and Europe, OPTIBOV)
- Identify selection signatures, genes and variants associated with adaptation, production and reproduction (Mr Nxumalo, PhD)
- Development of a customized low density SNP Chip.



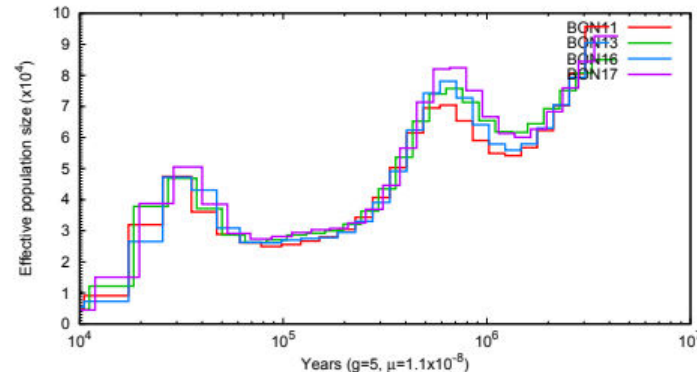
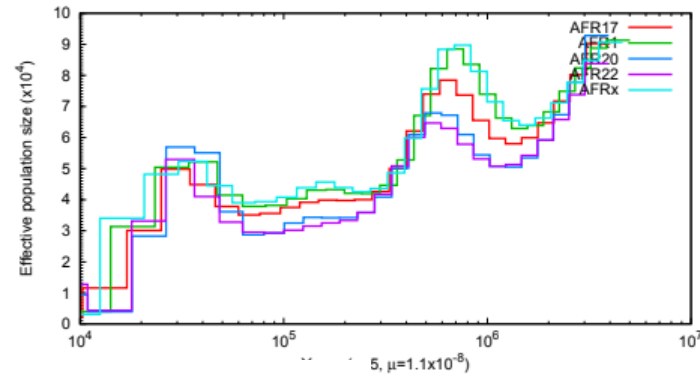
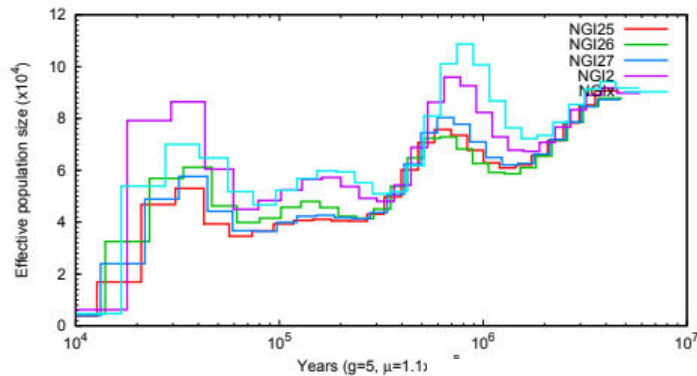
Relationships



Relationship between SA and African indigenous cattle breeds, ancestry (ANGSD)

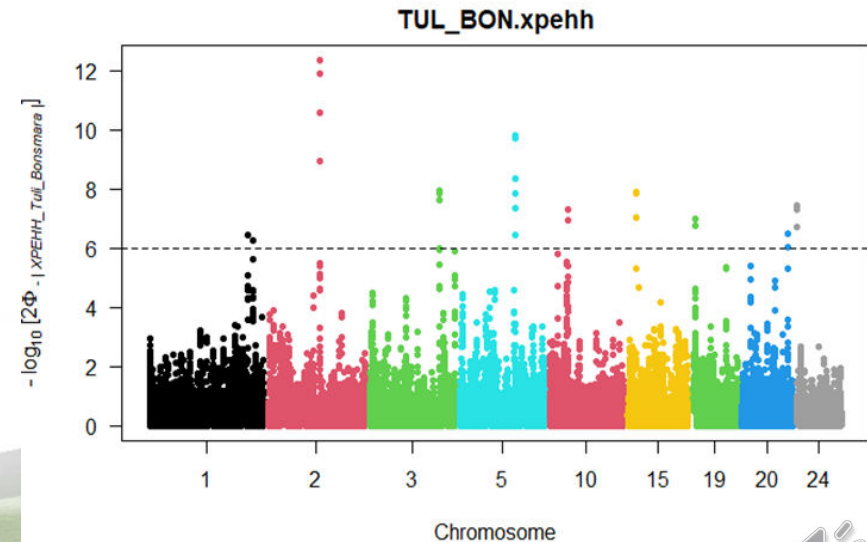
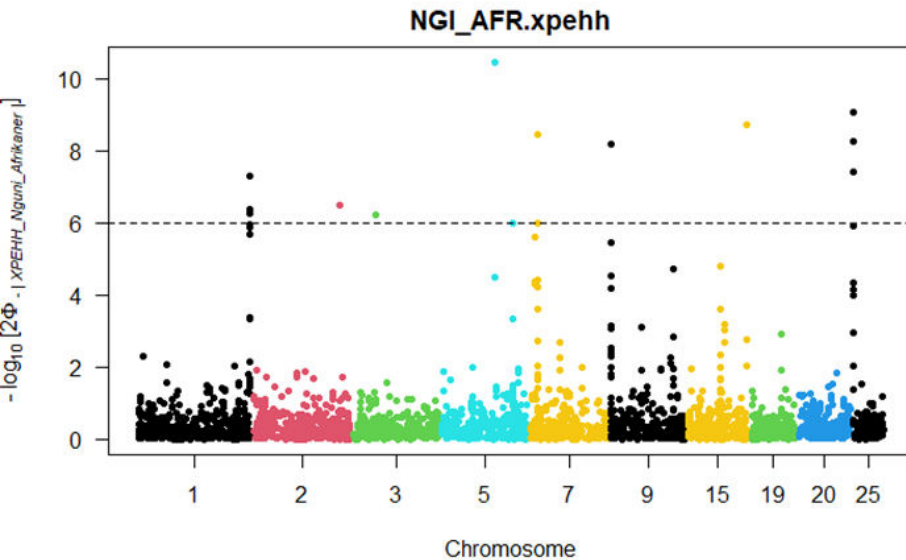
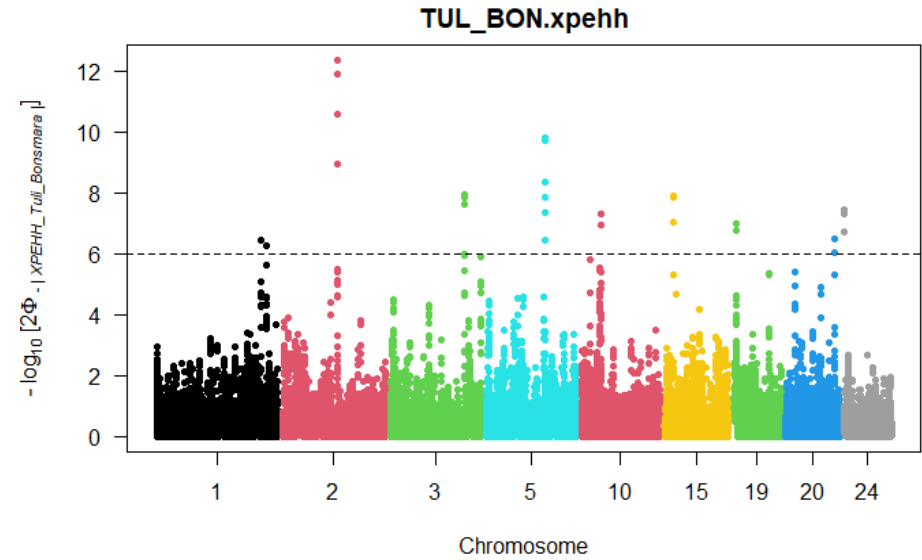
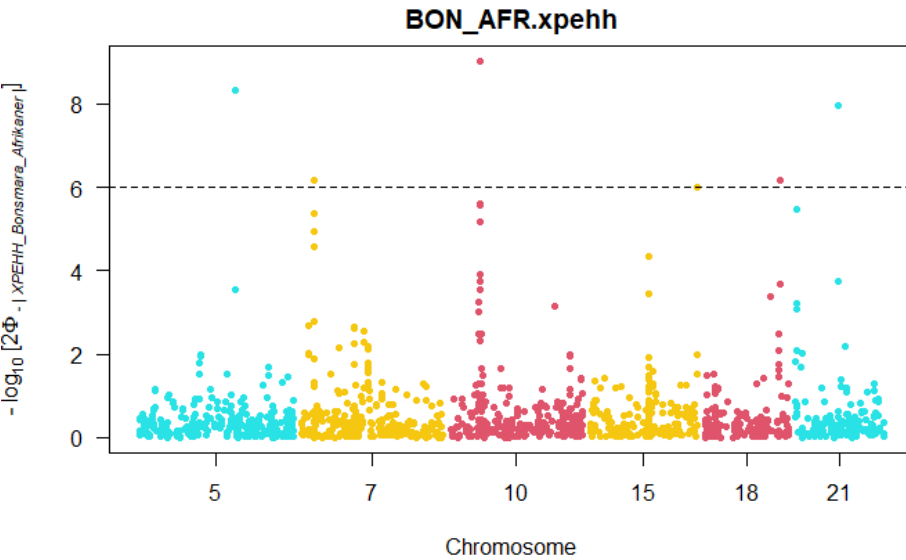
Demographic analysis

Showing how the effective population size has changed over time (Unpublished data)



Effective population size is determined by the diversity that is existing within the population (PSMC).

OPTIBOV data

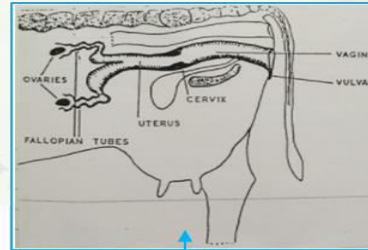


Regions of selection signatures in Afrikaner, Bonsmara, Nguni and Tuli, using XP-EHH method across breeds

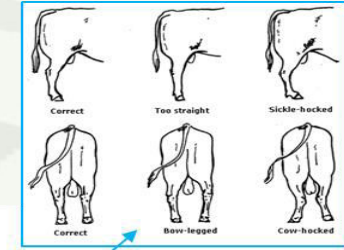
Gene associations



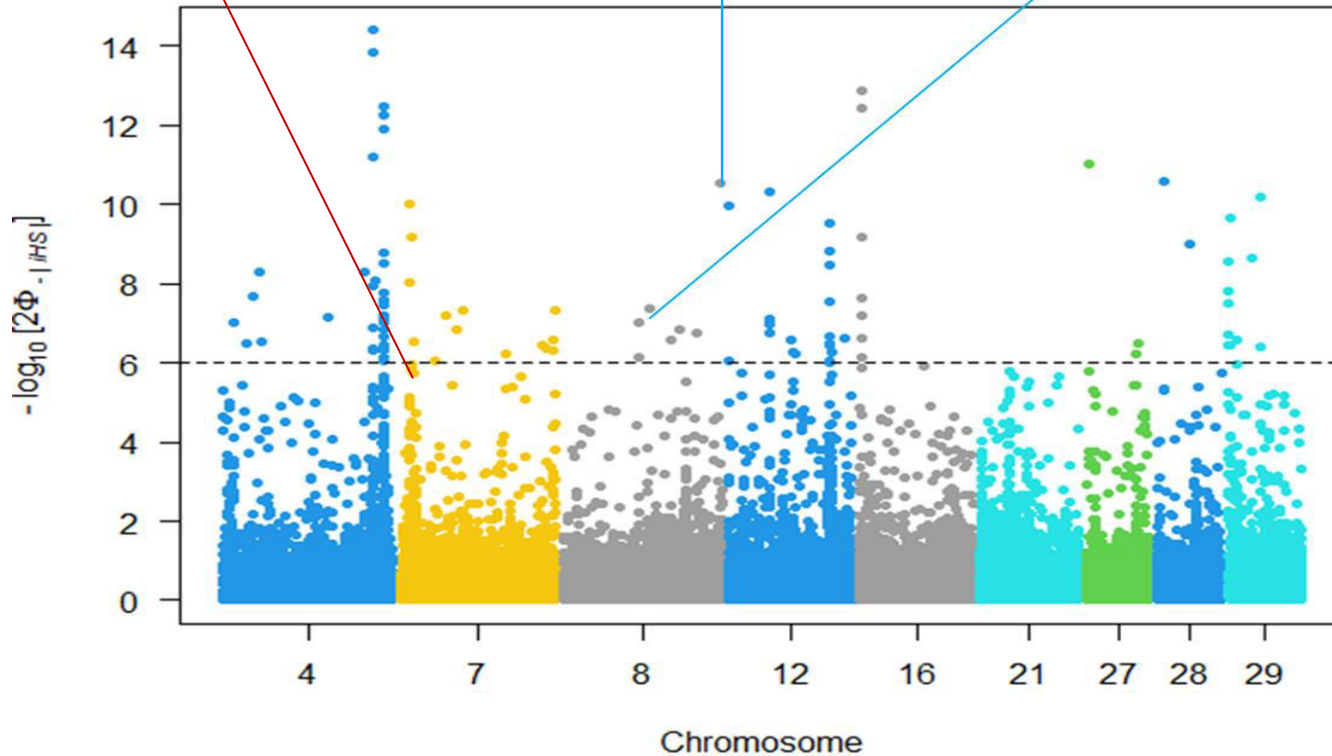
CHR4: regulates the production of progesterone by the placenta



CHR8: reproductive system development and regulation



CHR8: body conformation or Structural soundness



Genes associated with reproduction in SA indigenous cattle (unpublished data)

Conclusion

- There is notable difference in using SNP data and whole genome sequence data in characterizing SA indigenous cattle breeds
- WGS data holds the potential for the improvement of indigenous livestock breeds due to the in-depth analysis genome-wide
- Genetic diversity studies gives the understanding of population structure, demographic history, introgression and heterozygosity levels in SA indigenous cattle
- Identification of selective signatures provides insight into selection events that have shaped the genomes of indigenous cattle breeds, and allows the identification of important genes
- This will allow genomic selection, sooner, in our indigenous breeds, and more studies are needed for other indigenous livestock species



Acknowledgements

