Exploring the Genetic Diversity of South African Indigenous and Rural Horse Breeds

By

Erin Frost

Submitted in fulfillment of the academic requirements for the degree of

Master of Science in Research

In the Discipline of Genetics

School of Life Science

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg Campus

2025

Supervisor: Dr Edgar Farai Dzomba

Co-Supervisor: Dr Nompilo Lucia Hlongwane

Co-Supervisor: Dr Magrieta Aletta van der Nest

Co-Supervisor: Dr Farai Catherine Muchadeyi

Conference outputs

Conferences I attended and presented research.

1. 54th Annual South African Society for Animal Science (SASAS) (03 – 05 July 2024)
   Unraveling Genetic Variability in South African Indigenous and Rural Horse Breeds through SNP Genotypic Markers
   E. Frost, N.L. Hlongwane, M.A. van der Nest, F.C. Muchadeyi, E.F Dzomba
   Published in "Book of abstracts for the 54th SASAS Congress" (https://www.sasas.co.za/wp-content/uploads/2024/10/Final-SASAS-ABSTRACTS-2024-Oral-Poster.pdf)

Papers submitted for publication

Publication 1 (Chapter 3) – Paper under review for publication in BMC Genomics

Frost, E., Hlongwane, N.L., van der Nest, M.A., Muchadeyi, F.C., Nini, S., Dzomba, E.F. (Under review) Unraveling genetic variability in South African indigenous and rural horse breeds through SNP genotypic markers. BMC Genomics.

Author contributions:

EF collected samples, analysed the data and wrote the draft manuscript of this paper. FM and ED designed the statistical experiments and sourced funding for this project. NH, MVDN and ED coordinated the project's execution and contributed to manuscript preparations and revisions. SN assisted in facilitating sample collection for this research. All authors read and approved the manuscript.

Progress of peer review:

The manuscript is currently under review by the BMC Genomics Journal.

COLLEGE OF AGRICULTURE, ENGINEERING AND SCIENCE

DECLARATION 1 - PLAGIARISM

The data described in this thesis were collected in the Biotechnology Platform laboratory at the Agricultural Research Council (ARC) Onderstepoort Campus, Pretoria, under the supervision of Dr Edgar Farai Dzomba, Dr Nompilo Lucia Hlongwane, Dr Magrieta Aletta van der Nest and Dr Farai Catherine Muchadeyi.

I, ……**Erin Frost**……., declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   Their words have been re-written but the general information attributed to them has been referenced
   Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

[redacted]

Erin Frost

March 2025

I certify that the above statement is correct and as the candidate's supervisors, I have approved this thesis for submission.

[redacted]

Dr Edgar Farai Dzomba
Supervisor
March 2025

[redacted]

Dr Nompilo Lucia Hlongwane
Co-Supervisor
March 2025

[redacted]

Dr Magrieta van der Nest
Co-Supervisor
March 2025

[redacted]

Dr Farai Catherine Muchadeyi
Co-Supervisor
March 2025

Acknowledgements

I would like to extend my sincere gratitude to the following people and organizations for their assistance and input within this valuable research study that was very close to my heart.

Firstly, I would like to thank all of my supervisors – for without you none of this would have been possible. Dr Edgar Farai Dzomba, you were my lecturer in undergrad and followed suit as my supervisor for both my postgraduate degrees – always encouraging me to study further. Thank you for that. Dr Nompilo Lucia Hlongwane, thank you for all the sampling trips, imparting knowledge of genomics lab techniques and your constant support throughout this project – most importantly, all the laughs. Dr Magrieta Aletta van der Nest, thank you very much for your constant support and contributions towards this study, your guidance as a researcher has profoundly influenced me. To Dr Farai Catherine Muchadeyi, thank you for giving me the opportunities involved within this study, you paved the way for me – thank you.

I would like to share my thanks to the rural horse farmers who allowed their horses to be used in this study, the assistance of the Department of Agriculture, Rural Development and Land Reform in KwaZulu-Natal and Eastern Cape – including the research technicians from Agricultural Research Council (ARC) Kaonafatso ya Dikgomo (KYD) scheme – Ncera in East London. I share my sincere gratitude to the commercial breeders/trainers who allowed their horses to be a part of this study from the Kotzen Racing Stables and the Saddlebred Society of South Africa. Your contribution towards our South African horse breed conservation and utilization is admirable.

Thank you to the Agricultural Research Council (ARC) for accepting me to be a part of the Professional Development Program (PDP) – the knowledge I have gained within the genomics field is greatly appreciated. Thank you to the National Research Fund (NRF) for the funding opportunity to conduct this important research, allowing our indigenous and rural horse breeds to be honoured.

To my girls, Kendell Frohlich, Cayley Shum, Courtney Noton, Courtney Odell and Shannen Venter - you all inspire me. To Bongekile Tabu, we got through our Masters' together – thank you. To my family, most people say that they cannot help the family that is given to them, but I thank the universe every day for giving me all of you. To my brother, Kyle Frost, thank you for being my best friend. Lastly, but most importantly, thank you to my parents Bridget and Warren Frost. Thank you for giving me such a beautiful life, thank you for raising me to believe that I can achieve anything I set my mind to and thank you for inspiring the love of horses within me. I promise to always make you proud.

For the love of horses.

Table of Contents

Conference outputs ii
Papers submitted for publication iii
DECLARATION 1 - PLAGIARISM iv
Acknowledgements v
List of Figures ix
List of Tables xi
Abstract xiii
List of Abbreviations xv
Chapter 1 1
1.1 General Introduction 1
1.2 Justification 4
1.3 Aims and Objectives of MSc Research Study 6
1.4 Outline of thesis 6
1.5 References 8
Chapter 2 15
Literature Review 15
2.1 Introduction 15
2.2 Modern-day horse breed origin determination and breed classification 16
2.3 The establishment of horse breeds in South Africa 18
2.4 Historic and modern-day economic, socio-economic and cultural societal roles of horse breeds within South Africa 22
2.4.1 Basotho pony 23
2.4.2 Nooitgedacht 25
2.4.3 South African Boerperd 27
2.4.4 English Thoroughbred 28
2.5 Evaluation of previously known and future potential studies conducted on the genetic diversity of indigenous horse populations within South Africa 32
2.5.1 Previous studies conducted that assess the genetic diversity of the indigenous horse populations 32
2.5.2 Future potential genomic studies that can be conducted to examine the genetic diversity of the indigenous horse breed populations 33
2.6 Conclusion 33
2.7 References 34
Chapter 3 41
Understanding the diverse roles of rural, indigenous and commercial horse breeds within South African Society 41
3.1 Abstract 41
3.2 Introduction 42
3.3 Materials and Methods 44

3.3.1 Detailed Surveys and Phenotypic Horse Data 44
3.3.2 Determination of Functional Roles between Communities 45
3.3.3 Phenotypic Analysis of Horse Functional Roles 46
3.3.4 Assessment of Socio-economic Impact in Relation to Functional Roles 47
3.4 Results 48
3.4.1 Determination of Functional Roles between Communities 48
3.4.2 Phenotypic Analysis of Horse Functional Roles 56
3.4.3 Assessment of Socio-economic Impact in Relation to Functional Roles 58
3.5 Discussion 61
3.6 Conclusion 64
3.7 References 65
Chapter 4 71
Unravelling Genetic Variability in South African Indigenous and Rural Horse Breeds through SNP Genotypic Markers 71
4.1 Abstract 71
4.3 Materials and Methods 73
4.3.1 Breeds/Populations Sampled 73
4.3.2 Genotyping and Quality Control 75
4.3.3 Population Structure and Differentiation (PCA, sNMF, Fst Pairwise, AMOVA) 75
4.3.4 Population Genetic Diversity (Runs of Homozygosity (ROH and ROH Islands), Inbreeding and FROH) 76
4.4 Results 77
4.4.1 Population Structure and Differentiation 77
4.4.2 Population Genetic Diversity 83
4.5 Discussion 88
4.6 Conclusion 97
4.7 References 98
Chapter 5 110
Genome-wide Scan for Selection Signatures to Provide Insight into the Adaptive Potential of Indigenous and Rural South African Horse Populations 110
5.1 Abstract 110
5.2 Introduction 111
5.3 Materials and Methods 112
5.3.1 Sample Collection 112
5.3.2 Genotyping and Quality Control 113
5.3.3 Within population scan for selection signatures (iHS-based Scores) 114
5.3.4 Cross-population scan for selection signatures (XP-EHH) 114
5.3.5 Candidate gene identification, functional annotation and enrichment analysis 115
5.4 Results 115

5.4.1 Within population scan for selection signatures 115
5.4.2 Cross-population scan for selection signatures 117
5.4.3 Candidate gene identification, functional annotation and enrichment analysis 120
5.5 Discussion 158
5.6 Conclusion 160
5.7 References 161
Chapter 6 167
6.1 General Discussion 167
6.1.1 Evaluation of economic, socio-economic and cultural roles of rural, indigenous and commercial horse populations 167
6.1.2 Determination of the genetic diversity of South African horse breed populations 168
6.1.3 Identification of selection signatures found within and between South African horse breed populations 169
6.2 References 170
Chapter 7 175
7.1 Conclusions and Recommendations 175

List of Figures

Figure 2.1 The “Groot Trek” migration pattern throughout South Africa that determined the establishment and geographical evolution of South African indigenous horse breeds 20

Figure 2.2 The phenotypic appearance and associated utilization of the indigenous Basotho pony within the region of Lesotho..... 24

Figure 2.3 The Nooitgedacht breed that was the first indigenous South African horse breed to be recognized and recorded within the South African Studbook Association (SA Studbook)..... 26

Figure 2.4 The historical and modern-day phenotypic appearance of the South African Boerperd horse breed that was genetically shaped throughout South African history 27

Figure 2.5 The historical and modern-day phenotypic form of the English Thoroughbred selected for specific athletic ability within the equine sporting world..... 29

Figure 2.6 The historical and modern-day phenotypical form of the Saddlebred that associated with their breed utilization throughout the centuries 31

Figure 3.1 GIS map of sample regions within Eastern Cape, KwaZulu-Natal and Gauteng Provinces..... 44

Figure 3.2 Functional role options within detailed survey 45

Figure 3.3 Monthly income per household/stud farm/training facility in relation to direct or indirect functional roles of horses belonging to each community 48

Figure 3.4 Population clustering of functional roles 54

Figure 3.5 Agglomeration of clusters based on hierarchical clustering method 54

Figure 3.6 Photographic identification demonstrating categorical trait-specific phenotypes for each horse breed population..... 56

Figure 3.7 Categorical Principal Components Analysis (CATPCA) exhibiting population clustering based on phenotypic traits 57

Figure 4.1 GIS map of sample regions within Eastern Cape, KwaZulu-Natal and Gauteng Provinces..... 74

Figure 4.2 Principal Component Analysis (PCA) exhibiting population clustering 78

Figure 4.3 Principal Component Analysis (PCA) exhibiting population clustering including Petersen et al. (2013) breed populations (99,530 SNPs) 79

Figure 4.4 sNMF based clustering of ancestral proportions..... 81

Figure 4.5 Pairwise Fst values arranged within a heatmap plot 82

Figure 4.6 Total number of Runs of Homozygosity (ROH) stretches and associated run length (Mbps)..... 84

Figure 4.7 Mean length (Mbps) Runs of Homozygosity (ROH) per chromosome 85

Figure 4.8 Genome-wide population inbreeding levels coefficients (FROH) based on ROH stretches 87

Figure 4.9 Chromosomal-level population inbreeding coefficients (FROH) based on ROH stretches 88

Figure 5.1 GIS map of sample regions within Eastern Cape, KwaZulu-Natal and Gauteng Provinces..... 113

Figure 5.2 Genome-wide scan for distribution of iHS scores across six horse breed populations..... 116

Figure 5.3 Genome-wide scan for distribution of XP-EHH scores across indigenous BSP breed and rural breed populations..... 118

Figure 5.4 Genome-wide scan for distribution of XP-EHH scores across rural breed populations 119

List of Tables

Table 2.1 Five recognized horse breed types and their associated breed characteristics 17

Table 3.1 Phenotypic trait assessment for each horse breed population in relation to functional ability..... 47

Table 3.2 Cross-comparison of a functional role in agriculture for each breed population 49

Table 3.3 Cross-comparison of a functional role in transport for each breed population 50

Table 3.4 Cross-comparison of a functional role in meat production for each breed population..... 51

Table 3.5 Cross-comparison of a functional role in rural racing for each breed population 52

Table 3.6 Cross-comparison of a functional role in commercial showing and racing for each breed population 53

Table 3.7 Agglomeration schedule of hierarchical clustering of the six breed populations..... 55

Table 3.8 Comparison of household monthly income (ZAR/Month) and functional roles of Basotho-type (BSP) breed population 58

Table 3.9 Comparison of household monthly income (ZAR/Month) and functional roles of East London Rural (ELR) breed population 59

Table 3.10 Comparison of household monthly income (ZAR/Month) and functional roles of KwaZulu-Natal (KZN) rural breed population..... 60

Table 3.11 Comparison of household monthly income (ZAR/Month) and functional roles of Eastern Cape Rural (ECR) breed population..... 60

Table 4.1 Population based *Fst* pairwise analysis 82

Table 4.2 *F*-statistical genetic differentiation within and between representative breed populations..... 83

Table 4.3 ROH islands/peaks detected among the six representative breed populations..... 86

Table 4.4 QTL analysis of the top 30 percentile ROH 92

Table 5.1 Gene identification, functional annotation and enrichment of iHS selection signatures in BSP breed population..... 121

Table 5.2 Gene identification, functional annotation and enrichment of iHS selection signatures in ECR breed population..... 124

Table 5.3 Gene identification, functional annotation and enrichment of iHS selection signatures in ELR breed population..... 126

Table 5.4 Gene identification, functional annotation and enrichment of iHS selection signatures in KZN breed population..... 128

Table 5.5 Gene identification, functional annotation and enrichment of iHS selection signatures in SAD breed population ..... **Error! Bookmark not defined.**

Table 5.6 Gene identification, functional annotation and enrichment of iHS selection signatures in THR breed population ..... **Error! Bookmark not defined.**

Table 5.5 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across indigenous BSP and rural ECR breed populations 134

Table 5.6 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across indigenous BSP and rural ELR breed populations 140

Table 5.7 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across indigenous BSP and rural KZN breed populations 143

Table 5.8 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across rural ECR and rural ELR breed populations 146

Table 5.9 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across rural ECR and rural KZN breed populations 149

Table 5.10 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across rural ELR and rural KZN breed populations 152

Table 5.11 Intersected selection signatures detected within population-specific iHS analyses and cross-population XP-EHH analyses 156

Abstract

South Africa boasts a diverse range of indigenous and rural horse breeds that thrive in both commercial breeding operations and rural farming communities. The infusion of Javanese war-horses by Dutch settlers marked the genesis of the Cape Horse, involving subsequent crossbreeding, with various equine stocks – some of which included the English Thoroughbred, the American Saddlebred and the Persian Arabian. The emergence of modern, South African horse populations stands as a testament to human ingenuity, natural selection, and adaptability. To date, there is a paucity of genomic research that has been conducted on indigenous Basotho pony and rural, non-descript South African horse breeds focusing on their multi-functional roles, genetic diversity and adaptive traits. This study aimed to characterize the genetic and functional diversity of the rural and indigenous horse populations in South Africa and providing essential data to counter potential threats of extinction, uncontrolled crossbreeding and exotic breed replacements.

Data was collected from 110 surveys and 282 phenotypic records across indigenous (Basotho Pony-type (BSP)), rural (Eastern Cape Rural (ECR); East London Rural (ELR) and KwaZulu-Natal (KZN) rural) and commercial (SA Saddlebred (SAD); Thoroughbred (THR)) horse breeds - collected in KwaZulu-Natal, Eastern Cape and Gauteng provinces. Data was collected across different production systems and horse breed representatives within defined provisional locations, based on literature and consultations with officers belonging to the Department of Agriculture and Rural Development, ensuring selections were relative to local breed availability. Farmers located within rural areas were interviwed based upon voluntary participation, whereby some farmers owned either a minimum of one horse, to some owning multiple horses. Horses were selected based on distant relatedness to minimize genetic bias and therefore the number of horse samples per farmer was not proportionate. This approach was adopted to ensure an unbiased representation of the breed populations, increasing accuracy of genetic diversity estimations. Similarly, the commercial breeder and trainer also participated voluntarily, with horse selection based on both distant relatedness and the number of available horses within the breeding stud and training facility. Data categorized and structured from the detailed surveys and phenotypic records were analysed using IBM SPSS Statistics version 29.0.2.0. Cross-tabulation analyses demonstrated that all four rural horse breeds (BSP, ELR, ECR and KZN), were significantly utilized for transportation and rural racing ($p = 0.001$). The BSP population also showed a unique functional role in meat production within Sotho and Tshwana cultural communities ($p = 0.001$), as supported by hierarchical cluster analysis (HCA), which placed BSP as the most distinct cluster. HCA revealed the closest similarity between commercial breeds SAD and THR, with an agglomeration coefficient of 0.000, indicating identical functional roles. CATPCA of six phenotypic traits identified five clusters, with rural breeds overlapping due to shared traits, while SAD and some THR individuals distributed separately, reflecting breed-specific phenotypic traits. Ordinal regression analysis found no significant socio-economic impact of these functional roles ($p > 0.050$), although their contributions improved livelihoods in rural communities through indirect means.

The Illumina EquineSNP70k and EquineSNP80k BeadChips were used to analyze genetic diversity of the horse breed populations. Principal Component Analysis (PCA) analysis performed using Golden Helix SNP Variation Suite (SVS) version 8.5.0, identified four distinct clusters, with the commercial breeds (SAD and THR) and the ELR separating from BSP, ECR, and KZN. Global PCA showed all South African breeds clustering together within continental representative clusters. Ancestral proportions analysis (sNMF) conducted within R/Rstudio version 4.3.1, suggested shared ancestry among the clusters, linked to SAD and THR precursor breeds. Minimal differentiation between indigenous and rural breeds was confirmed by Pairwise Fixation Index (Golden Helix SNP Variation Suite (SVS) version 8.5.0) and AMOVA (ARLEQUIN version 3.1) results (low inter-population variation, (F_{st} < 0.060; high intra-population variation, F = 94.390%). Runs of Homozygosity (ROH) analysis conducted within R/Rstudio version 4.3.1, revealed shorter accumulative runs (<300 Mbps) in BSP, ECR, and THR populations, while ELR showed variable lengths (300–800 Mbps), indicating recent inbreeding or selective events. SAD and KZN populations exhibited higher total runs. ROH peaks (>0.70) highlighted adaptive selection in KZN and ELR for traits like fertility (*NLRP13*; *NLRP5*), immunity (*FIZ1*; *SSC5D*), and growth (*EPN1*; *SNAI1*). Commercial breeds showed selection for sporting traits but also contained deleterious genes (*MAGI1*; *ADGRL3*) in SAD. Inbreeding levels (F_{ROH} = 0.032–0.326) were highest in SAD, ELR, and KZN, and lowest in BSP, ECR, and THR populations.

SNP data extracted from the 282 horse specimens was used to identify and evaluate genomic regions under positive selection within and between the six breed populations in order to provide insight into the adaptive genetic traits preserved. These findings thus contribute to conservation efforts and support the sustainable management of horse populations that hold cultural and economic significance. Strong selective sweeps from population-specific iHS analyses were identified on chromosomes 1, 2, 4, and 7, linked to traits such as fertility and reproduction (*COBL*, *CFAP58*, *SPACA3*), nervous system development (*UNC13*, *ARHGAP35*, *DTNA*, *HTR1A*), growth (*FBN1*), and feed efficiency (*GLB1L2*, *B3GAT1*, *PPARA*). Traits associated with nervous system development were found to be significant within these horse populations, highlighting their control system in shaping the behaviour, cognitive function and response to environmental stimuli, which collectively facilitates their rapid adaption to diverse working roles. Cross-population XP-EHH analyses revealed similar selection patterns, reinforcing the fixation of these genomic regions. Intersection of XP-EHH and iHS analyses highlighted unique selection signatures: hair follicle development (*SOS1*) in ECR, locomotory behaviour (*ADAM22*) in ELR, and fear response (*SLC1A1*) in ELR vs. KZN, reflecting adaptations to ecological niches. Additional traits related to blood vessel morphogenesis (*SEMA3E*, *TSPAN12*, *JAM3*) and olfactory responses (*OR7D20*, *OR1M1*) further illustrated environmental adaptation. These results underscore the functional and genetic importance of indigenous and rural horse breeds, advocating for conservation policies that sustain their role, in South African societies.

Keywords: South African horse populations, functional characterization, CATPCA, genetic diversity, global PCA, ROH islands, selection signatures (iHS and XP-EHH)

List of Abbreviations

AHSV – African Horse Sickness Virus

AMOVA – Analysis of Molecular Variance

ARC-BTP – Agricultural Research Council-Biotechnology Platform

BCE – Before Common Era

BSP – Basotho Pony-type

CATPCA – Categorical Population Cluster Analysis

CV – Cross-validation error

df – Degrees of Freedom

DAVID – Database for Annotation, Visualization and Integrated Discovery

DNA – Deoxyribonucleic Acid

EAT – Equine-assisted Therapy

ECR – Eastern Cape Rural

ELR – East London Rural

EMS – Equine Metabolic Syndrome

EU – European Union

FAO – Food and Agricultural Organization

FDR – False Discovery Rate

F*ROH* – Inbreeding coefficient based on Runs of Homozygosity lengths

F*CT* – F-statistical among populations differentiation

F*IS* – F-statistical intra-population differentiation

F*st* – Fixation Index

HCA – Hierarchical Cluster Analysis

HWE – Hardy-Weinberg Equilibrium

IBD – Identity By Descent

IBH – Insect Bite Hypersensitivity

IBM – International Business Machines

iHS – Integrated Haplotype Score

Kb – Kilo base

K-value – Assumed number of ancestral populations

KZN – KwaZulu-Natal rural

LD – Linkage Disequilibrium

L*genome* – Total length of genome

L*ROH* – Total length of Runs of Homozygosity

MAF – Minor Allele Frequency

Mbps – Mega base Pairs

PC1 – Principal Component 1

PC2 – Principal Component 2

PCA – Population Cluster Analysis

PLUM – Polytomous Universal Model

QC – Quality Control

ROH – Runs of Homozygosity

SAD – South African Saddlebred

SLL – Supporting Limb Laminitis

SMMEs – Small, Medium and Micro Enterprises

sNMF – Sparse Non-negative Matrix Factorization

SNP – Single Nucleotide Polymorphism

SPSS – Statistical Package for Social Sciences

SVS – SNP Variation Suite

THR – South African Thoroughbred

VOC – Vereenigde Oost-Indisch Compagnie (Dutch East India Company)

XP-EHH – Cross-population Extended Haplotype Homozygosity

ZAR – South African Rand

Chapter 1

1.1 General Introduction

Horses have been integral to human history, evolving from wild prey animals to domesticated partners (Ahmad et al., 2020; Hole et al., 2023). Initially existing as prey within their environments, wild horses developed physiological adaptations for rapid flight and unpredictable manoeuvres to evade predators, including humans (Hole et al., 2023). Archaeological evidence indicate the use of horses as a food source in early human settlements, with specialized tools crafted for hunting and horse remains used to create weapons (Hutson et al., 2024; Hoffecker et al., 2018). Over time, humans recognized the potential of horses in advancing civilization, leading to their domestication. This process, as noted by Van Dierendonck and Sprijt (2012), reduced rapid prey responses and facilitated trust while maintaining herd social structures. Horses' ability to adapt to human-controlled environments enabled them to fulfil various roles, including agricultural labour, warfare, and transportation (Brubaker and Udell, 2016; Koski and Spannari, 2024; Lenders, 2024).

The introduction of horses to South Africa in 1652 by Dutch settlers marked a transformative moment in the region's history. Despite the lack of archelogical evidence, Swart (2003) advocates for the evidence of horses introduced into South Africa through VOC (Vereenigde Oost-Indisch Compagnie) shipping logs and journal entries by Ducth settlers, as the indigenous equid species that were found within the country could not be tamed and workhorse power was needed to initiate colonization infrastructure efforts. Initially used in mounted warfare and infrastructure development, these animals influenced both the biophysical and socio-political landscapes of the colonies (Swart, 2003; Swart, 2010). For instance, Basotho warriors utilized resilient Basotho ponies to defend their landlocked country, which became a symbol of cultural identity and economic resilience (Millin, 2007; Mofuoa, 2015; Phafoli and Zulu, 2012). These horses continue to play vital roles in rural South African communities, providing transportation, agricultural power, and aiding towards food security (Wadham et al., 2022; Danso-Abbeam et al., 2024). Simultaneously, the importation of English Thoroughbred stallions to improve local stock laid the foundation for South Africa's Thoroughbred racing industry, a sector that has grown into a significant economic contributor and cultural phenomenon (McShane and Swart, 2011; Bailey et al., 2022; Surujlal et al., 2011).

The genetic history of South African horse breeds reflects a rich tapestry of anthropogenic and natural selective pressures, from the importation of Javanese ponies by Dutch settlers to the crossbreeding with various global bloodstocks to consistently improve the initial horse population; the resulting Cape Horse exemplifies the interplay of genetic diversity and functional adaptability (Swart, 2003). The global infusion toward genetic improvement included the ancient Arabian breed, which promoted a lighter frame and high-arched conformation to the small-statured Javanese stock, enhancing athletic ability across varied terrains (Machmoum et al., 2020). This refinement was further advanced through the genetic contribution of the Persian Arabian breed, distinguished by a slightly taller and leaner stature relative the standard Arabian (Sadeghi et al., 2018). The incorporation of English Thoroughbred bloodstock enchanced the stamina, speed and overall

athletisim within the initial South African horse population, thereby increasing their suitability for a wider range of working roles (Thiruvenkadan et al., 2009). The British invasion of the Cape Colony in 1795 further diversified these populations, with horses becoming integral to military and agricultural endeavours during the Great Trek and the Mfecane wars (Potgieter and Grundlingh, 2007; Knapp, 2021). The Mfecane wars involved widespread conflicts involving various ethnic groups from Lesotho and surrounding areas, which triggered a series of raids on Voortrekker settlements (Knapp, 2020). During these events, the voortrekker horses were taken and brought into the Kingdom of Lesotho, where they were presented to King Moshoeshoe as a gift – contributing to the emergence of the Basotho people's mounted culture (Thabane, 2023). The isolation of Cape Horses in Lesotho's harsh environmental conditions gave rise to the Basotho pony, renowned for its endurance and adaptability (Millin, 2007; McCormack, 2018). This isolation led to the admixture of this indigenous breed into crossbred rural populations, conserving beneficial breed-specific traits that increase survivability to face environmental conditions. These contributing historical events contributed towards the diversification of locally adapted indigenous and rural horse breed populations that possess key selection signatures, desired phenotypic traits and multi-functional abilities that need to be genetically and functioning characterized.

The most recent report of the horse population statistics within South Africa, documents a total of 328 thousand horses as of 2021 (Cowling, 2023) that were maintained within the various equine sectors that comprise the South African equine industry of today. The distribution of the horse population is based on their designated roles within society, including involvement within the competitive equine sport industry, breeding of registered horse breeds that contribute towards the import and export trade as well as their use for transport and agricultural production within rural communities (Mdlulwa et al., 2021). The Thoroughbred horse racing industry established within South Africa, dating back to records of the early 1800s within the British Cape colony, that later expanded to settlements within the Eastern Cape and KwaZulu-Natal provinces (Herringer. 2004). The spread of the competitive horse racing industry and the associated economic growth that the sport provided throughout the years, escalated to a great significance of contribution towards the country's GDP of over R2 billion annually (Ngalonkulu, 2018). Ryan (2019) reports that the horse racing industry provides over 177 thousand employment opportunities across associated sectors, including Thoroughbred breeding studs and saervices. Not only has the commercial horse racing industry provided economic growth for the country but so has rural horse racing, that has gained traction in terms of significance for cultural tourism within South Africa, due to major rural racing events that in turn supports the socio-economic status of local rural communities (Nkwanyana et al., 2016). Major rural racing events such as the Dundee July held in the uMzinyathi district of KwaZulu-Natal (Mhlauli and Ezeuduji, 2022), as well as the Annual Berlin November race held in Buffalo City district of Eastern Cape (Ngumbela, 2023), highlight the cultural and economic significance of indigenous and crossbred horse breeds.

The import and export trade of horses within and out of South Africa has throughout the years, significantly contributed to the economic GDP, gaining R250 million per annum which now has the potential of growing to around R1 billion (Sanne, 2019). This growth is attributed to a risk model developed by Sergeant et al. (2016),

which estimates a low probability of exporting undetected African Horse Sickness (AHS). AHS is a highly infectious, often fatal viral disease that is endemic to sub-Saharan Africa, spread through the bite of an infected *Culicoides* midge spieces in which amongst all equid species found within South Africa – horses are the most susceptible (Dennis et al., 2019; Rodriguez et al., 2020; Venter et al., 2014). Marlow (2010) documents the first record of a major outbreak of AHS in South Africa in 1719, occurring shortly after the Dutch initiated continuous breeding improvement efforts of the orginal Javenese imports, which marked the beginning the spread of this highly infectious disease that continues to affect horse populations in South Africa today. Over the years, extensive research efforts have been undertaken in order to aid in control of the spread of this disease (de Klerk et al., 2023; Grewer et al., 2012; Zientara et al., 2015). Key efforts include the ongoing development of live AHS vaccines covering all identified viral strains (von Teichman et al., 2010) and the establishment of an AHS-free zone in the Western Cape, which facilitates the export of horses (Buhrmann, 2016). While international export of horses from South Africa has been hindered as to control AHS spread within Sub-Saharan Africa, low-risk zones and extensive quarantine protocol have facilitated trade with the European Union (EU) (Buhrmann, 2016; Hofmeyr, 2011). This direct export approval reduces costs for South African sport horses participating in international events and increases profits for breeders supplying internal buyers (Pillay, 2024).

These versatile breeds are in high demand for multidisciplinary equine sports, enabling profitable outcomes for breeders (Ryan, 2023). In contrast, the Basotho pony, another recognized indigenous breed, lacks a dedicated breeding society. Past preservation efforts by Lesotho's Department of Livestock Services and the National Stud conserved the Basotho pony population after its decline post-1996 Anglo-Boer War (Lekota, 2014). However, uncontrolled crossbreeding in rural areas bordering Lesotho threatens the breed's genetic preservation (Martin, 2007). Despite this, crossbreeding promotes economic gains and preserves genetic resources in locally adapted populations (de Klerk, 2020).

Locally adapted horse breed populations play crucial roles in supporting South Africa's rural economy, specifically in regions where small-scale farmers face significant challenges, due to limited resources to utilize mechanized farming equipment (Zantsi and Bester, 2019). Small-scale farmers that are densely concentrated within the provinces of KwaZulu-Natal (Lewu and Assefa, 2009) and Eastern Cape (Zantsi and Bester, 2019), that significantly contribute to the economic growth of the rural farming sector within South Africa, advocate for the utilization of horse power to pull basic agricultural tools as a low-maintenance and more cost-effective alternative. These rural horse breeds provide another functional role in terms of transport to people within rural areas that lack resources to purchase motorized vehicles and specifically local farmers that require transport across farmlands that comprise of rocky, mountainous terrain that motorized vehicles would not be able to travel through (de Klerk et al., 2020). Not only do these rural horse breeds fulfil practical roles within rural communities, they provide a sense of cultural identity to the people within these regions that have long been intertwined within their cultural history (Wadham et al., 2022).

Despite the significance of locally adapted rural horse breeds, genomic research on them remains limited. A pioneering study by Cothran and Van Dyk (1998) analysed genetic diversity in three indigenous breeds (Basotho Pony, Nooitgedacht, and SA Boerperd) using blood groups and biochemical loci. Their findings revealed significant genetic similarity to the SA Thoroughbred and North American Saddle breeds, indicating shared ancestry. Breytenbach et al. (2020) conducted a performance trait analysis study focusing on gait- and height-associated mutations within the SA Boerperd breed using seventeen microsatellite markers that were genotyped for the sampled breed population – which are essential indicators for the breed specifics of the indigenous breed. This microsatellite data was also used to determine the genetic diversity of the sampled SA Boerperd breed population, whereby upon comparison to nine European breeds, the indigenous population exhibited increased heterozygosity and a high occurrence of unique alleles.

With the advances of modern high-throughput SNP genotyping technologies, researchers now have unprecedented opportunities to delve deeper into the genetic architecture of these populations. These advanced tools enable the exploration of genetic diversity, demographic histories and ancestral lineages with greater precision (Balague-Dobon et al., 2022). Moreover, such genomic studies can identify selection signatures that reveal how these breeds have adapted to their environments and the anthropogenic pressures they have faced. To this date, no studies have previously been published that assess the genetic diversity of rural and indigenous horse breed populations using SNP genotypic markers, imploring the use of SNP BeadChip genomic tool endorsed by Illumina (Illumina, USA). The development of the Equine SNP70K BeadChip, shortly after the release of the Equine SNP50K BeadChip (Schaefer and McCue, 2020), was designed using a higher-density array of mapped SNP markers that were sequenced from seven representative breeds that include Akhal-Teke, Andalusian, Arabian, Icelandic, Quarter horse, Standardbred and Thoroughbred breeds (McCue et al., 2012). This high-throughput genomic tool has gained traction in the utilization of genomic studies that focus on locally adapted horse breeds on a global scale, such as indigenous Brazilian horse breeds (Nogueira et al., 2022), Iranian indigenous horse breeds (Mousavi et al., 2023) and many more. By bridging the gap between traditional knowledge and cutting-edge genomic techniques (Raudsepp et al., 2019), future research can not only enhance our understanding of our locally adapted South African horse breeds but also contribute to their preservation and sustainable utilization in South African society. Therefore, the findings of this study will enchance our understanding of the genetic diversity of these indigenous and rural horse populations and potential admixture with other exotic breeds. This research study will provide essential information on the importance of these indigenous horse breeds within South Africa, as to hinder the potential threat of extinction of these breed populations, explosive crossbreeding as well as exotic breed replacements. Thus, allowing for the propagation of increased management, utilization and conservation of South African horse breeds.

1.2 Justification

Despite extensive research focused on horse domestication (Librado et al., 2021; Warmuth et al., 2012) and the development of global horse breeds in Europe (Sole et al., 2019), Asia (Radovic et al., 2024), North America (Khanshour et al., 2014) and South America (Jimenez et al., 2012) – limited research studies focusing

on characterizing the genetic and functional diversity of South African indigenous and rural horse breeds have been conducted. Although a few genetic studies that focus on South African indigenous horse breeds have been published (Cothran and Van Dyk, 1998; van der Merwe and Martin, 2002) that provided a basis on population genetic diversity, the incorporation of high-throughput genomic tools to provide extensive insight into the genome-wide population dynamics of both indigenous and rural horse breeds has yet to be conducted. Therefore, the incorporation of genome-wide SNP genotyping technology, that has significantly benefited research within livestock breeding and genomic programs (Gurgul et al., 2014), within this research study enabled comprehensive analyses of genetic variation within and between populations that facilitated linkage of functional characterization and diversity of each population. This linkage emphasizes the importance of locally adapted horse breeds that contain advantageous, environmentally adaptive traits (Hoffmann, 2013), for their multi-purpose use within both rural farming and commercial breeding sectors - advocating for their preservation within modern society and avoidance of exotic breed replacement.

Phenotypic characterization of locally adapted horse breeds provides the crucial initial step in genetic resource conservation programs, as phenotypic data enables the classification of individuals within a population (Ribeiro et al., 2024). Most indigenous and rural horse populations are the product of various precursor breeds that incorporates trait integration, characteristic of crossbred individuals (Monau et al., 2018), which results in high phenotypic variation across populations and increases difficulty in determining breed-specific traits within each population. However, the association of particular phenotypic traits that includes body shape and conformation that were highlighted within this study (Brooks et al., 2010), allows for the functional characterization of these breed populations utilized within different communities and emphasizes their importance in their functional performance. The inclusion of precursor horse breeds and their known breed-specific traits that contributed to breed improvements of South African horse stocks (Marlow, 2010), provides source identification within these populations as well as the degree of variation in which individuals have adapted to specific environments. Functional characterization and association of phenotypic performance of these indigenous and rural horse breeds within this community-driven research study, drives the economic, socio-economic and cultural validation to promote continuous conservation and management of locally adapted breeds.

The genetic characterization of these horse populations that incorporates population structure and genetic diversity provides key links of the demographic histories of indigenous and rural breeds (Bazvand et al., 2024) to draw conclusions with comparisons of purebred precursor breeds. Genomic analyses utilized within this study measured the degree of genetic differentiation between the indigenous, rural and purebred commercial breeds and the extent of admixture, revealing the level of gene flow between populations (Pritchard et al., 2000) – which further validates the demographic histories and breeding patterns of locally adapted populations. Further genetic characterization through the identification of regions within the genome under positive selection, provides insight into the selection events undertaken within each breed population (de Simoni Gouveia et al., 2014), that have shaped the modern state of their adaptive potential. Associated selection

signatures detected within these regions facilitate the link of advantageous phenotypic traits that ensure the adaptability of these populations subjected to both human influence and harsh environmental conditions (Tiwari et al., 2024) – further providing their validation of essential South African genetic resources.

Therefore, the findings of this research study will provide support for the formation of necessary policies and management practices for the conservation of indigenous and rural horse breeds, thereby ensuring their continued contribution to rural livelihoods, cultural identity and diversity in South Africa.

1.3 Aims and Objectives of MSc Research Study

This genomics assisted, community-driven research study aims to characterize the genetic and functional diversity of the rural and indigenous horse populations within South Africa. This achieved through the following objectives:

i. To conduct detailed surveys with rural farmers, breed societies and horserace training facilities in order to evaluate the economic, socio-economic and cultural roles that these breed populations play in various communities.
ii. To genotype sampled horse populations (using the Illumina EquineSNP70k and EquineSNP80k BeadChip genotype array) in order to determine the genetic diversity of the rural, indigenous and commercial South African horse breed populations.
iii. To use the genotyped data to identify and study the signatures of selection found within and between the rural, indigenous and commercial South African horse breed populations.

1.4 Outline of thesis

The dissertation opens with a comprehensive introduction and relevant literature that comprehensively explores the development of indigenous and rural horse breeds, intertwined within South African history – paying tribute to the ground work of previous genetic diversity research studies conducted on indigenous horse breeds but imploring the use of high-throughput genomic tools to functionally and genetically characterize both indigenous and rural horse breed populations. The chapters contained within the dissertation are organized as follows:

Chapter 1: General introduction, overall aim and objectives of this research study.

Chapter 2: Literature review.

Chapter 3: Understanding the diverse roles of rural, indigenous and commercial horse breeds within South African Society.

Chapter 4: Unravelling genetic variability in South African indigenous and rural horse breeds through SNP genotypic markers.

Chapter 5: Genome-wide scan for selection signatures to provide insight into the adaptive potential of indigenous and rural South African horse populations.

Chapter 6: General discussion.

Chapter 7: Conclusions and recommendations.

1.5 References

Ahmad, H.I., Ahmad, M.J., Jabbir, F., Ahmar, S., Ahmad, N., Elokil, A.A., Chen, J. (2020). The domestication makeup: Evolution, survival, and challenges. Frontiers in Ecology and Evolution 8. Available: https://doi.org/10.3389/fevo.2020.00103 [accessed November 20, 2024].

Bailey, E., Petersen, J.L., Kalbfleisch, T.S. (2022). Genetics of Thoroughbred racehorse performance. Annual Review of Animal Biosciences 10, 131-150.

Balague-Dobon, L., Caceres, A., Gonzalez, J.R. (2022). Fully exploiting SNP arrays: a systematic review on the tools to extract underlying genomic structure. Briefings in Bioinformatics 23(2). Available: https://doi.org/10.1093/bib/bbac043 [accessed October 05, 2023].

Bazvand, B., Rashidi, A., Zandi, M.B., Moradi, M.H., Rosamzadeh, J. (2024). Genome-wide analysis of population structure, effective population size and inbreeding in Iranian and exotic horses. PLOS One 19(3). Available: https://doi.org/10.1371/journal.pone.0299109 [accessed December 17, 2024].

Brooks, S.A., Makvandi-Nejad, S., Chu, E., Allen, J.J., Streeter, C., Gu, E., McCleery, B., Murphy, B.A., Bellone, R., Sutter, N.B. (2010). Morphological variation in the horse: defining complex traits of body size and shape. Animal Genetics 41(2), 159-165.

Brubaker, L., Udell, M.A.R. (2016). Cognition and learning in horses (*Equus caballus*): What we know and why we should ask more. Behavioural Processes 126, 121-131.

Buhrmann, G. (2016). AHS vaccinations – what you need to know. Sabinet African Journals 13 (1), 14-15.

Burgess, M. (2014). Contributing to the Nooitgedachter legacy. Farmer's Weekly. Available: https://www.farmersweekly.co.za/rural-insight/contributing-to-the-nooitgedachter-legacy/ [accessed September 13, 2022].

Coleman, A. (2008). The sturdy bloodline of the SA Boerperd. Farmer's Weekly. Available: https://www.farmersweekly.co.za/archive/the-sturdy-bloodline-of-the-sa-boerperd/ [accessed September 13, 2022].

Cothran, E. G., Van Dyk, E. (1998). Genetic analysis of three South African horse breeds. Journal of the South African Veterinary Association 69(4): 120–125.

Cowling, N. (2023). Number of horse population in South Africa 2010-2021. Statista. Available: https://www.statista.com/statistics/1306731/stock-of-horses-in-south-africa/#:~:text=The%20number%20of%20horses%20in,increased%20from%20nearly%20325%20thousand [accessed February 01, 2025].

Danso-Abbeam, G., Ogundeji, A.A., Asale, M.A., Baiyegunhi, L.J.S. (2024). Effects of livestock ownership typology on household food security in rural Lesotho. GeoJournal 89. Available: https://doi.org/10.1007/s10708-024-11049-y [accessed November 22, 2024].

De Klerk, J.N., Quan, M., Grewar, J.D. (2020). Socio-economic impacts of working horses in urban and peri-urban areas of the Cape Flats, South Africa. Journal of the South African Veterinary Association 91. Available: https://doi.org/10.4102/ jsava.v91i0.2009 [accessed November 25, 2024].

De Klerk, J.N., Gorsich, E.E., Grewar, J.D., Atkins, B.D., Tennant, W.S.D., Labuschagne, K., Tildesley, M.J. (2023). Modelling African horse sickness emergence and transmission in the South African control area using a deterministic metapopulation approach. PLOS Computational Biology 19(9). Available: https://doi.org/10.1371/journal.pcbi.1011448 [accessed May 15, 2025].

De Simoni Gouveia, J.J., Barbosa da Silva, M.V.G., Paiva, S.R., Pinheiro de Oliveria, S.M. (2014). Identification of selection signatures in livestock species. Genetics and Molecular Biology 37(2), 330-342.

Dennis, S.J., Meyers, A.E., Hitzeroth, I.I., Rybicki, E.P. (2019). African horse sickness: A review of current understanding and vaccine development. Viruses 11(9). Available: https://doi.org/10.3390/v11090844 [accessed May 15, 2025].

Grewar, J.D., Weyer, C.T., Guthrie, A.J., Koen, P., Davey, S., Quan, M., Visser, D., Russouw, E., Buhrmann, G. (2012). The 2011 outbreak of African horse sickness in the African horse sickness controlled area in South Africa. Journal of the South African Veterinary Association 84(1). Available: http://dx.doi.org/10.4102/jsava.v84i1.973 [accessed May 15, 2025].

Gurgul, A., Semik, E., Pawlina, K., Szmatola, T., Jasielczuk, I., Bugno-Poniewierska, M. (2014). The application of genome-wide SNP genotyping methods in studies on livestock genomes. Journal of Applied Genetics 55, 197-208.

Herringer, P. (2004). Thoroughbred horse racing and breeding in South Africa. Thoroughbred Heritage. Available: https://www.tbheritage.com/Breeders/SABreeding/SABreeders.html [accessed September 13, 2022].

Hoffecker, J.F., Holliday, V.T., Stepanchuk, V.N., Lisitsyn, S.N. (2018). The hunting of horse and the problem of the Aurignacian on the central plain of Eastern Europe. Quaternary International 492, 53-63.

Hoffmann, I. (2013). Adaption to climate change – exploring the potential of locally adapted breeds. Animal 7, 346-362.

Hofmeyr, I. (2011). Get your globe-trotter from A to B. SA Horseman, 10-11.

Hole, C., Murray, R., Marlin, D., Freeman, P. (2023). Equine behavioural and physiological responses to auditory stimuli in the presence sand absence of noise-dumping ear covers. Animals (Basel) 13(9). Available: https://doi.org/10.3390/ani13091574 [accessed November 20, 2024].

Hutson, J.M., Villaluenga, A., Garcia-Moreno, A., Turner, E., Gaudzinski-Windheuser, S. (2024). Persistent predators: Zooarchaeological evidence for specialized horse hunting at Schoningen 13II-4. Journal of Human Evolution 196. Available: https://doi.org/10.1016/j.jhevol.2024.103590 [accessed November 20, 2024].

Jimenez, L.M., Mendez, S., Dunner, S., Canon, J., Cortes, O. (2012). Colombian creole horse breeds: same origin but different diversity. Genetics and Molecular Biology 35(4), 790-796.

Khanshour, A., Juras, R., Blackburn, R., Cothran, E.G. (2014). The legend of the Canadian horse: genetic diversity and breed origin. Journal of Hereditary 106(1), 37-44.

Knapp, A. (2020). Horse breeds indigenous to South Africa – a noble heritage. The Design Train. Available: https://www.thedesigntrain.com/post/horse-breeds-indigenous-to-south-africa-a-noble-heritage [accessed September 05, 2022].

Knapp, A. (2021). Lesotho's national transport – a brief history of the Basotho pony. The Design Train. Available: https://www.thedesigntrain.com/post/the-basotho-pony-lesotho-s-national-transport-a-brief-look-at-this-fascinating-history [accessed September 20, 2023].

Koski, S.E., Spannari, J. (2024). Experiences of intersubjectivity and agency in human-horse relationships differ between box- and open-shed-stable environments. Anthrozoös 37(4), 687-704.

Lekota, T.A. (2014). The state of the Basotho pony in Lesotho. FAO. Available: https://www.fao.org/fileadmin/user_upload/nr/sustainability_pathways/docs/The%20State%20of%20the%20Basotho%20Pony%20in%20Lesotho.pdf [accessed September 12, 2022].

Lewu, F.B., Assefa, Y. (2009). Farmers' knowledge in the cropping systems of northern KwaZulu-Natal, South Africa: Current challenges and solution for sustainable future food production. African Journal of Agricultural Research 4(11), 1148-1153.

Librado P., Khan N., Fages A., Kusliy M. A., Suchan T., Tonasso-Calvière L., Schiavinato S., Alioglu D., Fromentier A., Perdereau A., Aury J.-M., Gaunitz C., Chauvey L., Seguin-Orlando A., Der Sarkissian C., Southon J., Shapiro B., Tishkin A. A., Kovalev A. A., Alquraishi S., Alfarhan A. H., Al-Rasheid K. A. S., Seregély T., Klassen L., Iversen R., Bignon-Lau O., Bodu P., Olive M., Castel J.-C., Boudadi-Maligne M., Alvarez N., Germonpré M., Moskal-del Hoyo M., Wilczyński J., Pospuła S., Lasota-Kuś A., Tunia K., Nowak M., Rannamäe E., Saarma U., Boeskorov G., Lõugas L., Kyselý R., Peške L., Bălăşescu A., Dumitraşcu V., Dobrescu R., Gerber D., Kiss V., Szécsényi-Nagy A., Mende B. G., Gallina Z., Somogyi K., Kulcsár G., Gál E., Bendrey R., Allentoft M. E., Sirbu G., Dergachev V., Shephard H., Tomadini N., Grouard S., Kasparov A., Basilyan A. E., Anisimov M. A., Nikolskiy P. A., Pavlova E. Y., Pitulko V., Brem G., Wallner B., Schwall C., Keller M., Kitagawa K., Bessudnov A. N., Bessudnov A., Taylor W., Magail J., Gantulga J.-O., Bayarsaikhan J., Erdenebaatar D., Tabaldiev K., Mijiddorj E., Boldgiv B., Tsagaan T., Pruvost M., Olsen S., Makarewicz C. A., Valenzuela Lamas S., Albizuri Canadell S., Nieto Espinet A., Iborra M. P., Lira Garrido J., Rodríguez González E., Celestino S., Olària C., Arsuaga J. L., Kotova N., Pryor A., Crabtree P., Zhumatayev R., Toleubaev A., Morgunova N. L., Kuznetsova T., Lordkipanize D., Marzullo M., Prato O., Bagnasco Gianni G., Tecchiati U., Clavel B., Lepetz S., Davoudi H., Mashkour M., Berezina N. Ya., Stockhammer P. W., Krause J., Haak W., Morales-Muñiz A., Benecke N., Hofreiter M., Ludwig A., Graphodatsky A. S., Peters J., Kiryushin K. Yu., Iderkhangai T.-O., Bokovenko N. A., Vasiliev S. K., Seregin N. N., Chugunov K. V., Plasteeva N. A., Baryshnikov G. F., Petrova E., Sablin M., Ananyevskaya E., Logvin A., Shevnina I., Logvin V., Kalieva S., Loman V., Kukushkin I., Merz I., Merz V., Sakenov S., Varfolomeyev V., Usmanova E., Zaibert V., Arbuckle B., Belinskiy A. B., Kalmykov A., Reinhold S., Hansen S., Yudin A. I., Vybornov A. A., Epimakhov A., Berezina N. S., Roslyakova N., Kosintsev P. A., Kuznetsov P. F., Anthony D., Kroonen G. J., Kristiansen K., Wincker P., Outram A., Orlando L. (2021). The origins and spread of domestic horses from the Western Eurasian steppes. Nature 598, 634–640.

Lenders, R. (2024). 'Wild' horses in medieval and early modern landscapes of Europe. Landscape History 44(2), 15-35.

Machmoumz, M., Boujenane, S., Azelhak, R., Badaoui, B., Petit, D., Piro, M. (2020). Genetic diversity and population structure of Arabian horse populations using microsatellite markers. Journal of Equine Veterinary Science 93. Available: https://doi.org/10.1016/j.jevs.2020.103200 [accessed May 14, 2025].

Martin, J. (2007). Revitalization of the Basotho Pony mare camps: a step towards increasing populations of the Basotho pony in Lesotho. People and Animals. Available: https://openknowledge.fao.org/server/api/core/bitstreams/9a5e5d11-8d58-44ad-940b-9a677cb3525c/content#page=32 [accessed February 02, 2025].

McCormack, J. (2018). PSXVI-14 the Basotho pony. A breed improvement program in the highlands of Lesotho. Journal of Animal Science 96. Available: http://dx.doi.org/10.1093/jas/sky404.551 [accessed September 21, 2023].

McCue, M.E., Bannasch, D.L., Petersen, J.L., Gurr, J., Bailey, E., Binns, M.M., Distl, O., Guerin, G., Hasegawa, T., Hill, E.W., Leeb, T., Lindgren, G., Penedo, M.C.T., Roed, K.H., Ryder, O.A., Swinburne, J.E., Tozaki, T., Valberg, S.J., Vaudin, M., Lindblad-Toh, K., Wade, C.N., Mickelson, J.R. (2012). A high density SNP array for the domestic horse and extant Perissodactlya: Utility for association mapping, genetic diversity, and phylogeny. PLOS Genetics 8(1). Available: https://doi.org/10.1371/journal.pgen.1002451 [accessed November 16, 2023].

McShane, C., Swart, S. (2011). Designing equids in South Africa and North America. Safundi 12(2), 203-227.

Mdlulwa, Z., Masemola, M., Lubisi, B.A., Chaminuka, P. (2021). The financial burden of African Horse Sickness: a case of the European Union trade ban on South Africa's horse industry. Agricultural Economics Research, Policy and Practice in Southern Africa 60(4), 353-369.

Mhlauli, O., Ezeuduji, I.O. (2022). When a tourism event goes rustic: Marketing a rural horse racing event in KwaZulu-Natal. Journal of Hospitality, Tourism and Leisure 11(1), 156-168.

Millin, P. (2007). The renowned Basotho pony. SA Horseman. Available: https://journals.co.za/doi/pdf/10.10520/EJC14081 [accessed September 08, 2022].

Mofuoa, K.V. (2015). The exemplary ethical leadership of the King Moshoeshoe of Basotho of Lesotho in the nineteenth century Southern Africa. Journal of Public Administration and Governance 5(3), 21-35.

Monau, P.I., Visser, C., Nsoso, S.J., van Marle-Koster, E. (2018). Phenotypic and genetic characterization of indigenous Tswana goats. South African Journal of Animal Science 48(5), 925-934.

Mousavi, S.F., Razmkabir, M., Rostamzadeh, J., Seyedabadi, H.R., Naboulsi, R., Petersen, J.L., Lindgren, G. (2023). Genetic diversity and signatures of selection in four indigenous horse breeds of Iran. Hereditary 131(2), 96-108.

Ngalonkulu, M. (2018). The South African horse industry: a potential billion-rand market. Forbes Africa. Available: https://www.forbesafrica.com/economy/2018/04/10/south-african-horse-industry-potential-billion-rand-market/ [accessed February 01, 2024].

Ngumbela, X.G. (2023). Empowering rural women through tourism entrepreneurship: The case of the Annual Berlin November horse racing in the Eastern Cape Province. International Journal of Social Science Research and Review 6(11), 195-207.

Nkwanyana, M.S., Ezeuduji I.O., Nzama, A.T. (2016). Cultural heritage tourism in South Africa: perceived a panacea for rural development. Acta Universitatis Danubius Oeconomia 12(6), 160-167.

Nogueira, M.B., de Faria, D.A., Ianella, P., Paiva, S.R., McManus, C. (2022). Genetic diversity and population structure of locally adapted Brazilian horse breeds assessed using genome-wide single nucleotide polymorphisms. Livestock Science 264. Available: https://doi.org/10.1016/j.livsci.2022.105071 [accessed February 02, 2025].

Phafoli, L.S., Zulu, N.S. (2012). The complexity of the cultural identity of Basotho in Lesotho. LWATI: A Journal of Contemporary Research 9(1), 238-253.

Pillay, T. (2024). Lifting the ban of horse export from SA brings relief. Available: https://www.iol.co.za/ios/news/lifting-on-ban-of-horse-export-from-sa-brings-relief-ccb09d73-a6ca-4b0e-9bbd-71d95206be02 [accessed February 02, 2025].

Potgieter, T., Grundlingh, A. (2007). Admiral Elphinstone and the Conquest and the Defence of the Cape of Good Hope. South African Journal of Military Studies 35(2), 39-67.

Pritchard, J.K., Stephens, M., Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics 155(2), 945-959.

Radovic L., Remer V., Rigler D., Bozlak E., Allen L., Brem G., Reissman M., Brockmann G. A., Ropka-Molik K., Stefaniuk-Szmukier M., Kalinkov L., Kalashnikov V. V., Zaitev A. M., Raudsepp T., Castaneda C., von Butler-Wemken I., Patterson Rosa L., Brooks S. A., Novoa-Bravo M., Kostaras N., Abdurasulov A., Antczak D. F., Miller D. C., Lopes M. S., da Câmara Machado A., Lindgren G., Juras R., Cothran G., Wallner B. (2024). The global spread of oriental horses in the past 1,500 years through the lens of the Y chromosome. PNAS 121. Available: https://doi.org/10.1073/pnas.2414408121 [accessed December 17, 2024].

Raudsepp, T., Finno, C.J., Bellone, R.R., Petersen, J.L. (2019). Ten years of the horse reference genome: insights into equine biology, domestication and population dynamics in the post-genome era. Animal Genetics 50(6), 569-597.

Ribeiro, N.L., Rodrigues de Medeiros, G., Vieira da Silva, N., Soares, K.O., Arandas, J.K.G., Vieira do Nascimento, G., Ribeiro, M.N. (2024). Phenotypic characterisation and production system of the Nordestino horse form a multivariate perspective. Scientific Reports 14. Available: https://doi.org/10.1038/s41598-023-51018-y [accessed November 26, 2024].

Rodriguez, M., Joseph, S., Pfeffer, M., Raghavan, R., Wernery, U. (2020). Immune response of horses to inactivated African horse sickness vaccines. BMC Veterinary Research 16(322). Available: https://doi.org/10.1186/s12917-020-02540-y [accessed May 15, 202].

Ryan, J. (2019). SA equine industry pushing for direct exports. Farmer's Weekly (South Africa). Available: https://www.pressreader.com/south-africa/farmers-weekly-south-africa/20191122/281891595112568 [accessed February 01, 2025].

Ryan, J. (2023). Why horses and donkeys still have a place in agriculture. Farmer's Weekly. Available: https://www.farmersweekly.co.za/animals/horses/why-horses-and-donkeys-still-have-a-place-in-agriculture/ [accessed February 02, 2025].

Sadeghi, R., Moradi-Shahrbabak, M., Ashtiani, S.R.M., Schlamp, F., Cosgrove, E.J., Antczak, D.F. (2019). Genetic diversity of Persian Arabian horses and their relationship to other native Iranian horse breeds. Journal of Heredity 110(2), 173-182.

Sanne, I. (2019). Published scientific study confirms that the risk of exporting African horse sickness is negligible. Equine Health Fund. Available: https://www.equinehealthfund.co.za/Portals/0/Risk%20Assessment.pdf [accessed February 02, 2025].

Schaefer, R.J., McCue, M.E. (2020). Equine genotyping arrays. Veterinary Clinics of North America: Equine Practice 36(2), 183-193.

Sergeant, E.S., Grewar, J.D., Weyer, C.T., Guthrie, A.J. (2016). Quantitative risk assessment for African Horse Sickness in live horses exported from South Africa. PLOS ONE 11(3). Available: https://doi.org/10.1371/journal.pone.0151757 [accessed February 02, 2025].

Sole, M., Ablondi, M., Binzer-Panchal, A., Velie, B., Hollfelder, N., Buys, N., Ducro, B.J., Francois, L., Janssens, S., Schurink, A., Viklund, A., Eriksson, S., Isaksson, A., Kultima, H., Mikko, S., Lindgren, G. (2019). Inter- and intra-breed genome-wide copy number diversity in a large cohort of European equine breeds. BMC Genomics 20 (759). Available: http://dx.doi.org/10.21203/rs.2.10580/v3 [accessed December 17, 2024].

Surujlal, J., Mosai, W., Dhurup, M. (2011). 'Sport of kings': analysis of motives for gambling on horseracing. African Journal for Physical Health Education, Recreation and Dance 17(2). Available: https://hdl.handle.net/10520/EJC121735 [accessed November 25, 2024].

Swart, S. (2003). Riding high – horses, power and settler society, c. 1654-1840. Kronos, Environmental History 29. Available: https://doi.org/10.2307/41056494 [accessed September 05, 2022].

Swart, S. (2010). "The world the horses made": A South African case study of writing animals into social history. International Review of Social History 55(2), 241-263.

Thabane, M. (2023). History of Lesotho. Oxford Research Encyclopedia of African History. Available: https://doi.org/10.1093/acrefore/9780190277734.013.1328 [accessed May 15, 2025].

Thiruvenkadan, A.K., Kandasamy, N., Panneerselvam, S. (2009). Inheritance of racing performance of Thoroughbred horses. Livestock Science 121 (2), 308-326.

Tiwari, M., Gujar, G., Shashank, C.G., Ponsuksili, S. (2024). Selection signatures for high altitude adaption in livestock: a review. Gene 927. Available: https://doi.org/10.1016/j.gene.2024.148757 [accessed December 18, 2024].

Van der Merwe, F.J., Martin, J. (2002). Four South African horse breeds. Animal Genetic Resources 32. Available: http://dx.doi.org/10.1017/S1014233900001565 [accessed December 17, 2024].

Van Dierendonck, M., Spruijt, B.M. (2012). Coping in groups of domestic horses – Review from a social and neurobiological perspective. Applied Animal Behaviour Science 138(3), 194-202.

Venter, G.J., Labuschagne, K., Majatladi, D., Boikanyo, S.N.B., Lourens, C., Ebersohn, K., Venter, E.H. (2014). *Culicoides* species abundance and potential over-wintering of African horse sickness virus in the Onderstepoort area, gauten, South Africa. Journal of the South African Veterinary Association 85(1). Available: http://dx.doi.org/10.4102/jsava.v85i1.1102 [accessed May 15, 2025].

Von Teichman, B.F., Dungu, B., Smit, T.K. (2010). *In Vivo* cross-protection to African horse sickness Serotypes 5 and 9 after vaccination with Serotypes 8 and 6. Vaccine 28(39), 6505-6517.

Wadham, H., Wallace, C., Furtado, T. (2022). Agents of sustainability: How horse and people co-create, enact and embed the good life in rural places. Sociologica 63(3), 390-414.

Warmuth, V., Eriksson, A., Bower, M.A., Barker, G., Barrett, E., Hanks, B.K., Li, S., Lomitashvili, D., Ochir-Goryaeva, M., Sizonov, G.V., Soyonov, V., Manica, A. (2012). Reconstructing the origin and spread of horse domestication in the Eurasian steppe. PNAS 109(21), 8202-8206.

Zantsi, S., Bester, B. (2019). Revisiting the benefits of animal traction to subsistence smallholder famers: a case study of Ndabakazi villages in Butterworth, Eastern Cape province of South Africa. South African Journal of Agricultural Extension.

Zientara, S., Weyer, C.T., Lecollinet, S. (2015). African horse sickness. Revue Scientifique et Technique Office International des Epizooties 34(2), 315-327.

Chapter 2

Literature Review

2.1 Introduction

The domesticated horse (*Equus ferus caballus*) is described as a single, uneven-toed, ungulate mammal that was proposed to be the evolutionary result of climate change from a humid forest environment to a drier environment (Solounias et al., 2018). This transition accelerated a diverse range of morphological changes due to adaption to new food resources that occupied throughout the ancestral lineage of the Equidae family. According to MacFadden et al. (2012), the fossilized evidence of the earliest ancestor of the Equidae family (which includes domesticated horses (*Equus caballus*), wild horses (*Equus ferus*) and horse-like mammals such as zebra (*Equus Quagga*)) was scientifically referred to as "Eohippus" or "Hyracotherium" that originated within the Eocene epoch geological age (56.5 to 35.4 million years ago). This polydactyl ancestor was characterized to be small in size, a jaw structure that was designed for browsing broad-leafed plants, in a marshy forest habitat that provided protection from predators. As the Eocene geological age began to change and become drier, this accelerated the established grazing feeding practices of the mammal and underwent extreme digit reduction which resulted in the monodactyl anatomy of the *Equus* genus (Solounias et al., 2018). Research by Solounias et al. (2018) indicated that the polydactyl condition of ancestral equids is preserved in the skeletal and superficial anatomy of modern horses, such as the chestnut – a callus-like growth on the inner side of their legs. This monodactyl adaption increased the length of equine limbs, resulting in greater height and stride length compared to their ancestors. These traits proved advantageous for high-speed locomotion in open grasslands, where evading predators was critical for survival (Janis and Bernor, 2019). The development of these adaptions not only facilitated the survival of wild equine populations, but may also have attracted human attention, leading to the eventual domestication of horses as a locomotive resource (Librado et al., 2016).

The domestication of the horse species represents one of the most influential events in ancient human history, fundamentally altering local transport, migration, trade and warfare (Guimaraes et al., 2020). Several theories on the timing and location of horse domestication have emerged over the decades. Researchers have debated on whether the domestication of horses occurred through a single event involving one ancestral group, or through multiple events involving several ancestral populations (Jansen et al., 2002). One theory posits that domestication began during the Bonze period (about 2,000 BCE to 700 BCE), with the Sintashta culture in the Ural Mountains of Western Russia and Northwestern Kazakhstan, based on archaeological evidence such as chariot burials and horse bridles (Allentoft et al., 2015). However, this theory suggests only the use of horses, not necessarily their domestication, prompting researchers to investigate earlier evidence. Researchers, such as Librado et al. (2021), provided genetic evidence that supports domestication of the horse in the Western Eurasian Steppes about 3,500 BCE. Through the analysis of 273 ancient horse genomes from suspected domestication centers that include Anatolia, Iberia and Central Asia, the researchers concluded that modern domestic horse populations replaced earlier local populations due to their expansion across the Western Eurasian steppes. Librado et al. (2021) analyzed ancient horse genomes using DNA extracted from fossil

remains representing proposed regions of horse domestication, that were dated using radiocarbon dating to narrow down the era of domestication, and sequenced using shotgun sequencing. Although sequencing of archeological remains proposes to be challenging in DNA quality, due to preservation lifespan (Orlando et al., 2021), shotgun sequencing has proven to be successful in providing genomic data with the assiastance of statistical imputation efforts (Todd et al., 2023). Two prominent genes, namely the *GSDMC* gene that was associated with a docile nature and the *ZFPM1* gene associated with weight-bearing ability through the facilitation of a strong spine structure, were identified as key factors enabling domestication (Cassella, 2021). These findings indicate that modern horses likely originated from a domesticated population in what is now Ukraine and Western Kazakhstan, spreading globally through human intervention.

Warmuth et al. (2012) expanded on these findings through analyzing mitochondrial DNA and whole-genome sequences from over 300 modern horse populations across 12 Eurasian regions. Using demographic modeling, the researchers concluded that the wild ancestors of domesticated horses (*Equus ferus*) originated in Eastern Eurasia around 160,000 years ago and were domesticated in the Western Eurasian Steppes nearly 6,000 years ago (Williams, 2012). This study not only presented one of the most plausible and accurate genetic evidence of the domestication of the modern-day horse but also proposed the fact that our modern-day, domesticated and feral horse populations should accurately be classified as *Equus ferus caballus* – due to the interbreeding of ancient wild horse populations with the domesticated ancestral horse populations. The work of Warmuth et al. (2012), provided insights into the genetic diversity, migration and spread of horse populations worldwide.

Therefore, the aim of this review paper was to delineate the evolutionary history of horse breeds over time, with a particular focus on South Africa and the key historical events, which resulted in the development of the three unique indigenous Basotho Pony, Nooitgedacht and Boerperd breeds of South Africa.

2.2 Modern-day horse breed origin determination and breed classification

Building upon the domestication theory, determining the origins of modern horse breeds has become essential for understanding selective breeding strategies, as it provides critical inisghts into their genetic foundations, historical development and adaptive traits (Eser and Erat, 2022; Todd et al., 2023). Wallner et al. (2017) combined historical studbooks, which documented horse pedigrees dating back to the 1700s, with Y chromosome DNA sequencing data from 363 male horses across 57 modern breeds. Their analysis of evolutionary descent according to sire breeding lines identified two prominent ancient lineages as precursors of modern breeds: the Arabian horse lineage from the Arabian Peninsula and the extinct Turkoman horse lineage from the Western Eurasian Steppe (Price, 2017). These lineages were introduced into Europe approximately 700 years ago, where diverse roles led to selective breeding for specific traits that included performance, conformation and gait (Hanot et al., 2018). This ultimately resulted in the explosion of diverse horse breeds that maintained standards of breed-specific characteristics, often with influence of morphological adaptions to regional environmental factors.

Throughout the centuries of horse breed refinement, equine researchers and health specialists classified these unique and diverse breeds into five prominent breed types that were based on standardized weight, height, body conformation associated with specific phenotypic features and societal roles – namely Draft horses, Light horses, Warmbloods, Ponies and Miniature horses (Griffin, 2020). Horse breed types cannot be defined based on coat colour due to the diverse range of coat colours and patterns that horses phenotypically manifest, as according to Thiruvenkadan et al. (2008), coat colour in horses is controlled by genes located at 12 different loci and is therefore not specific to horse breeds or horse breed types. The specific weight, height, body conformation and breed examples of each recognized breed type is shown in **Table 2.1** below.

Table 2.1 Five recognized horse breed types and their associated breed characteristics

[The five recognized horse breed types are differentiated based on their weight, height and body conformation associated with specific phenotypic features. Examples of the horse breeds categorized within the five breed types are given.) [Adapted from "Breeds of Equine", by Chappalwar, A., Sakhare, M., Katkuri, S. (2022). Pashudhan Praharee. Available: https://www.pashudhanpraharee.com/breeds-of-equine/ [accessed August 30, 2022] and "Horse", by Cothran, E.G., Podhajsky, A.W. (2022). *Encyclopaedia Britannica*. Available: https://www.britannica.com/animal/horse [accessed August 30, 2022]]

[*1 Hand (hh) is equal to 10.160cm]

Breed Type	Weight (Kg)	Height (hh) (Hands)*	Body Conformation	Societal Roles	Breed Examples
Draft Horses	635Kg and above	14.2hh – 17.2hh	- Muscular - Large limbs (including head, foreleg and hindleg, hoof size and body size) - Feathery feet	- Agricultural production (pulling carts, farming machines) - Parade cart towers	- Clydesdale - Shire - Percheron - Belgian
Light Horses	400Kg – 635Kg	12hh – 17.2hh	- Light-framed, athletic build - Slender legs - Fine, smaller heads	- Sporting events (Show jumping, horse racing, eventing, dressage and showing) -Leisure riding and trail riding	- Thoroughbred - American Quarter horse - Arabian breeds - American Standardbred - Saddlebred
Warmbloods	545Kg – 700Kg	16hh – 18hh	- Medium-framed to large-framed, athletic build - Muscular - Long, thicker-built forelegs and hindlegs	- Sporting events particularly show jumping - Leisure riding	- Irish Sport horse - Dutch warmblood - Hanoverian - Holsteiner
Ponies	226Kg – 400Kg	14.2hh or less	- Small, stocky body frame - Proportionally smaller legs	- Sporting events (Show jumping, dressage and showing) for children	- Shetland pony - Welsh - Connemara

			- Small head size - Heavy-set coat	- Leisure riding for children	- Pony of Americas
Miniature Horses	68Kg – 160Kg	9.5hh or less	- Extremely small-framed - Limbs extremely small but proportionate to body	- Historic use in carting materials in mines - Companion animals/pets	- Falabella - Dutch Miniature - South African Miniature - American Miniature horse

Through the classification of horse breed types (**Table 2.1**), associated breed specific traits were assigned to each breed, leading to the formal registration of many developed horse breeds worldwide. However, there has been major controversy over the number of classified and registered horse breeds throughout the past few decades. Aldridge (2000) addresses the research study conducted by Khadka (2010), which received extensive criticism, on the number of recognised breeds that the author documented throughout the geographical regions of the world. Aldridge (2000) documented that throughout the geographical regions of Africa, Asia, Europe, Latin America/Caribbean, Pacific Islands and USA/Canada there were a total of 527 recognized horse breeds. However, this high number attributed to district breeds may be due to the incorrect classification of some crossbreeds and breed composites. According Pester (2021), there have been over 350 registered horse breeds throughout the world up to date, but this total number will constantly change due to intense selection pressures of breeding strategies placed to create new horse breeds.

The existence of many distinctive horse breeds throughout the different geographical regions of the world, barring the progressive selection strategies placed on horse populations to create desirable breeds, was due to the introduction of horses to these different regions through the colonization efforts of the Europeans (Delsol et al., 2022). The horses introduced to these regions evolved through a combination of natural selection, adapting to new environments, and artificial selection by colonists who bred them to develop specialized breeds suited for specific societal roles. According to Swart (2010), the existence of defined, modern-day horse breeds within South Africa was due to the historic colonization event of European settlers, whereby a small number of horses of a particular breed were shipped with the settlers to aid in colonization efforts.

2.3 The establishment of horse breeds in South Africa

The first documented European colonization event that occurred in South Africa was the invasion of the Dutch settlers, that docked on the port of the southern tip of South Africa referred to as the Cape of Good Hope, in 1652 (Hunt, 2005). The establishment of the Dutch settlers within the Cape (where the Indian and Atlantic Ocean waters meet) was a tactical plan by the Dutch East India Company, VOC (Vereenigde Oost-Indisch Compagnie) to be positioned between the European homelands of the settlers and the East Indies, in order to provide an efficient shipment route for spices and other trading goods. Swart (2003) indicated that horses were the first domestic livestock breed shipped to South Africa after the arrival of the Dutch, as horses played a

major role within European society and provided a means of physical labour in efforts of the shipment company construction, agricultural farming for food supply. This first shipment of horses within South Africa were crossbreds referred to as Java ponies, each composite a mixture of an Arabian horse breed and a North American breed known as the Barb, that were acquired from Arabian traders within the East Indies (Knapp, 2020). However, the breeding potential of this initial crossbred population was limited, prompting the importation of various horse breeds from around the world. These efforts included trade negotiations to acquire Andalusian stallions from South America, Persian Arabs, and English Thoroughbreds. By the early 1800s, these imported breeds were crossed with Java ponies, resulting in an improved horse population that settlers recognized as the Cape Horse (Cothran & van Dyk, 1998).

The British Empire gained strength through their warfare tactics and rapid colonization, which included the invasion of the Cape Colony that began in 1795, and eventually overthrew the Dutch East India Company site to take control over the trading route. This prevented any potential French invasion of the British colony established within India (Oliver and Oliver, 2017). The British Empire forced the Dutch settlers to abandon their establishments and propagated one of the most significant inland exploration events, within the South African country, known as the 'Great Trek' – which commenced from 1836 and onwards (Swart, 2003). The Great Trek of the Dutch inhabitants (commonly referred to as the 'Boers') allowed for the establishment of South African indigenous horse breeds through the various regional proliferation and adaption of the Boers' horses – beginning with the denotation of the Boerperd horse (Cothran and van Dyk, 1998). The below figure (**Figure 2.1**) illustrates the Great Trek inland by the Dutch settlers and the associated geographical establishment of indigenous South African horse breeds.

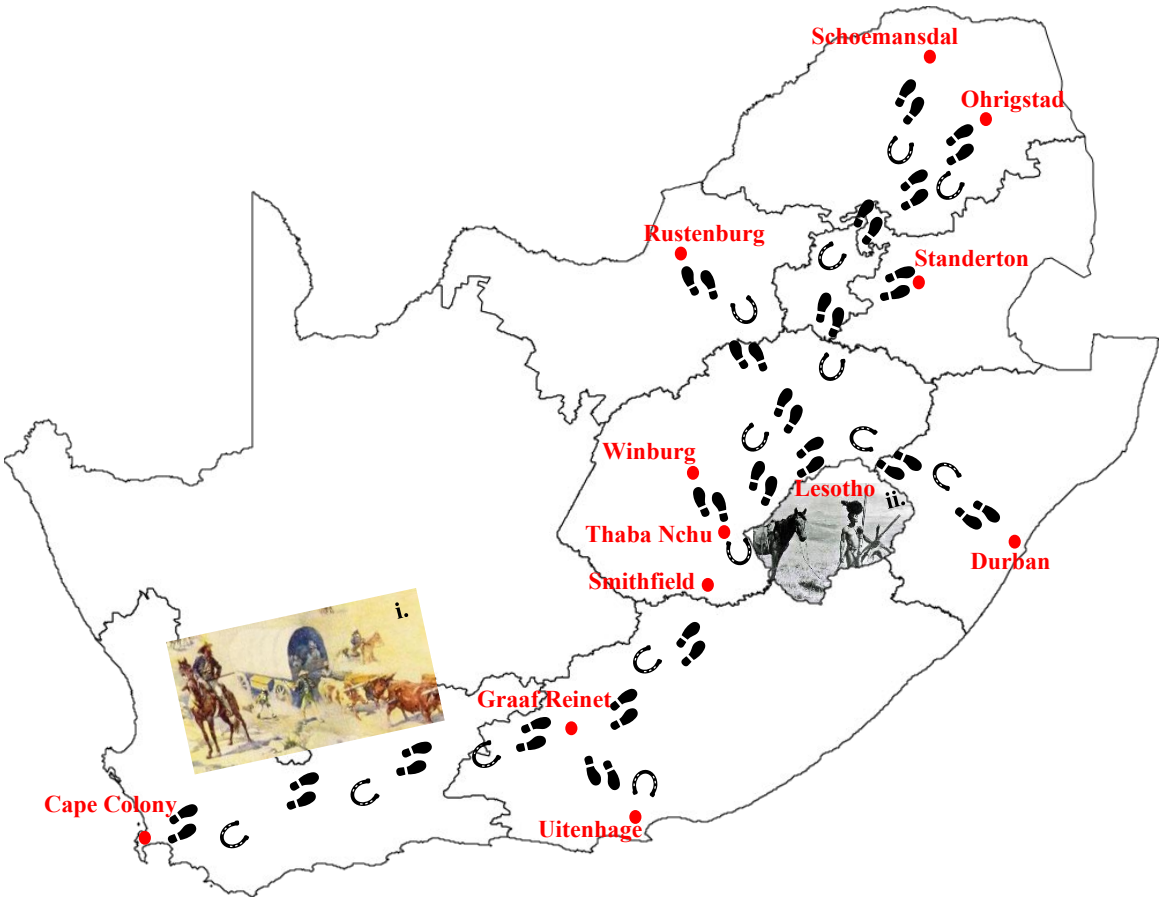


Figure 2.1 The "Groot Trek" migration pattern throughout South Africa that determined the establishment and geographical evolution of South African indigenous horse breeds

[The migration pattern of the Dutch settlers (referred to as the 'Boers') composed of some regions of settlements within the provinces of Eastern Cape (Graaf Reinet and Uitenhage), Free State (Smithfield, Thaba Nchu and Winburg), North West (Rustenburg), Limpopo (Schoemansdal and Ohrigstad), Mpumalanga (Standerton) and settlement within Durban through the trek around Lesotho. The images within the figure show the use of the Boerperd by the Dutch settlers **i.** Sourced from: https://showme.co.za/facts-about-south-africa/history-of-south-africa/the-history-of-south-africa/ and a Basotho warrior holding a Basotho pony used for warfare **ii.** Sourced from: https://imgur.com/gallery/7nPezEc .) [Adapted from "The settlers of South Africa and the expanding frontier", by Fourie, J. (2020). Stellenbosch Economic Working Papers: WP14/2020. Available: http://www.ekon.sun.ac.za/wpapers/2020/wp142020 [accessed September 08, 2022] and from "Lesotho's national transport – a brief history of the Basotho pony", by Knapp, A. (2021). The Design Train. Available: https://www.thedesigntrain.com/post/the-basotho-pony-lesotho-s-national-transport-a-brief-look-at-this-fascinating-history [accessed September 08, 2022]].

The figure above (**Figure 2.1**) presents as a map of South Africa that depicts the migration pattern of the Dutch settlers, that gave themselves the name "Voortrekker" (in English translated to "Fore trek") (Swart, 2010), which established the nomadic Boerperd breed from the originating population of the Cape Horse from the Cape colony. The Voortrekkers dispersed from the Eastern Cape, particularly from Graaff-Reinet and Uitenhage, into various regions of South Africa, bringing with them the well-adapted Boerperd horse breed. This breed had adapted to harsh conditions and survived several severe environmental challenges, including the spread of African Horse Sickness Virus (AHSV), which led to significant mortality among the migrating horse population (Nel, 2015). African Horse Sickness (AHS) is a highly infectious disease caused by several strains of the African Horse Sickness Virus (AHSV) carried and spread through the arthropod vector known

as the *Culicoides* midge, which causes respiratory and circulatory impairments in equid-specific species (Dennis et al., 2019). Due to the non-endemic origins of these Boerperd populations, the breed did not acquire immunity to this disease, resulting in significant fatalities that reduced their numbers. However, the breed was not driven to extinction, as some individuals survived and gradually developed a degree of acquired tolerance to the disease. This recovery enabled the nomadic settlers to continue their exploration further inland, which included one significant settlement within the Free state province (Thaba Nchu settlement), found in close proximity to the border of the Kingdom of Lesotho (previously deemed as Basutoland) (as seen in **Figure 2.1**).

Knapp (2020) highlighted that rebellions among ethnic tribes in Basutoland and surrounding areas, collectively known as the Mfecane wars, led to numerous raids on Voortrekker settlements. During these raids, Boerperd horses were seized and taken back to the mountainous region of Lesotho, where they were presented as gifts to the reigning king of the time. Their adaptability was shaped by the need to survive in the geographically isolated and mountainous terrain of Lesotho, where they endured extreme weather conditions and limited grazing resources (Cothran and van Dyk, 1998). This ultimately led to the development of the Basotho pony breed that is characterized by the breed's surefootedness, calm temperament, small-frame, endurance and hard-hooves (Millin, 2007).

Within the British-controlled Cape Colony, the breeding of the military Cape Horse expanded rapidly to meet the growing demand for export to other British colonies. As a result, breeding strategies prioritized both the genetic improvement of the breed and the expansion of breeding stock (Swart, 2003). This strategy was achieved through the importation of various European horse breeds that included the indigenous British breeds of the Hackney horse (Terry, 2022) and the Cleveland Bays (Richard, 2005) - both renowned for their gaited movement and refined features. The importation of military draught horses (most distinctly horses of Flemish origin) (Cothran and van Dyk, 1998) that possessed characteristics of a muscular and broad-framed body composition with an associated sensible temperament, was also utilized for the breed improvement strategy. These European breeds were crossbred with the Cape Horse breeding stock in order to achieve the ideal warhorse – characterized by their nimble ability and sturdiness. These highly improved warhorses were used by the British cavalry in one of the most infamous wars in South African history, which took place between the Republics of the inland Boers and the British Empire, deemed as the Anglo-Boer War that occurred between 1899-1902 (Caramello, 2016). To the dismay of the British Empire, this brutal war and harsh environmental factors resulted in a devastating decline in their selective crossbred warhorse population, in comparison to the minor decline in the Boerperd populations of the opposition that were advantageous in their acquired adaptability. These fatal events collectively led to the eventual extinction of the Cape Horse population, leaving only the Boerperd populations of the Republics and the Basotho pony populations of the Lesotho Kingdom – both originating from the initial Cape Horse population (Jacobs, 2011).

Due to the complete decline in the British Empire's crafted Cape Horse population, the British forces set upon exploration further inland to source superior horse supplies for importation and set their sights upon the well-adapted Basotho pony populations within Lesotho – significantly exhausting the Basotho populations from the indigenous ethnic people of Lesotho (Lekota, 2014). This rapid decline of the culturally significant Basotho

breed population caused major concerns within the Kingdom of Lesotho as their mounted culture depended solely on these horses. According to McCormack (2018), the concerns of the Lesotho people were recognized many years later and the Indigenous Livestock sector of the South African Department of Agriculture launched an initiative to prevent the extinction of the Basotho pony. This effort, along with the goal of developing a suitable South African riding horse, was undertaken at the Nooitgedachter Research Station in the Transvaal region. This conservation effort involved the selection of a small breeding group of Basotho-type ponies, that consisted of one particular stallion (named 'Vonk') and six mares from a local farmer J.A.N Cloete situated within the Cape province, as well as five Basotho-type mares from another local farmer M.G Willemse that was situated within the Free state. The Nooitgedachter Research Station collectively decided that the selection of only one stallion to cover the nine Basotho-type mares, would have undoubtedly caused high levels of inbreeding within the resurrecting population and therefore a Boerperd-type stallion (a direct descendant of the originating Cape Horse breed) was selected to be bred with a portion of the mares. The conservation efforts of the Indigenous Livestock sector and the Nooitgedachter Research Station proved to be successful through the replenishment the South African horse breed population and according to Hofmeyr and Scholtz (2017), in 1984 the Nooitgedacht became the first indigenous South African horse breed to be recognized and recorded within the South African Studbook Association (SA Studbook). This significant event led to the declaration and recognition of the three indigenous horse breeds of South Africa – the Basotho pony, the Nooitgedacht and the South African Boerperd.

2.4 Historic and modern-day economic, socio-economic and cultural societal roles of horse breeds within South Africa

South Africa is home to many diverse and refined horse breeds, which consist of our three indigenous South African breeds, as well as the vast quantity of imported horse breeds from various regions throughout the world. Pinnock (2022) states that the most recent estimated total of the domesticated horse population within South Africa consists of about 300,000 horses, of which this total population falls into three functional categories - namely into workhorses, racehorses and performance/sport horses. Workhorses are characterized by the ability to perform strenuous or manual labour for human activities, which historically played a crucial role in transportation before the era of motorization, but are still utilized for transportation purposes and agricultural practices in the modern world of today (Henry, 2021). An example of a workhorse breed within South Africa would be the indigenous Basotho pony, whereby this culturally important breed population was utilized as a mode of transport for the people of Lesotho as well as to transport goods and produce. Racehorses are characterized by their selective athletic ability for speed and endurance that are utilized within the business-orientated racing industry, whereby several light-framed horse breeds are selected and bred for specific categories of horse racing such as flat racing, harness racing and endurance racing (Jones, 2022). South Africa consists of many racehorse breeds that include the English Thoroughbred, the Arabian breed and the Saddlebred – of which all three of these breeds face different selection criteria related to the racing category. Lastly, performance horses (commonly referred to as sport horses) are characterized as horse breeds that are selected, trained and bred to perform highly athletic equestrian sports, of which some include show jumping,

eventing, dressage, equitation, vaulting and barrel racing (Hector, 2021a). South Africa consists of many horse breeds selected for these athletic sporting events, examples being our indigenous Boerperd (historically used as workhorses) and Nooitgedacht breeds, but within the equestrian sporting world warmblood importations and breeding are dominating sporting events.

The historical and modern-day societal roles of each of these indigenous and commercial horse breeds are identified below, focusing on the specific phenotypic characteristics that each breed possesses that enables the breeds to fulfil the economic, socio-economic and cultural roles within South African society.

2.4.1 Basotho pony

The Basotho pony is among the most infamous and respected indigenous horse breeds to South Africa, whereby this breed possesses a unique origin through the selective pressures of the mountainous range that defines the terrestrial ecology of the Lesotho region, previously highlighted by Cothran and van Dyk (1998). In the most recent article detailing the Basotho pony breed written by Law (2020), the author states that the population of the Basotho pony breed within Lesotho was estimated to be 98,000, with more than half of the estimated population size occupying the most rocky mountainous districts of the country. Lekota (2014) states that the phenotypic traits defining the Basotho pony breed have remained largely unchanged since its establishment in the nineteenth century. As a result, the breed's role in society has been preserved into the modern era, while also being utilized to enhance the economic and socio-economic status of the Kingdom of Lesotho. The figure (**Figure 2.2 (i.), (ii.)** and **(iii.)**) below represents the historical and modern-day phenotypic appearance of the Basotho pony breed.



Figure 2.2 The phenotypic appearance and associated utilization of the indigenous Basotho pony within the region of Lesotho

[The phenotypic appearance of the Basotho pony that defines the breed type has been maintained throughout the historical to the modern-day era which is associated with the cultural roles that this horse breed plays within society except with a growth of the economic and socio-economic aspects within modern society – indicated by the associated images. Image **i.** indicates the historical state of the Basotho pony alongside a Sotho warrior in traditional attire. Image **ii.** indicates the modern-day appearance of the Basotho pony with mounted Sotho people dressed in their modern-day cultural attire. Image **iii.** indicates the modern-day utilization of the Basotho pony for tourism purposes in order to boost the economic and socio-economic status of Lesotho.) [**i.** Sourced from: https://imgur.com/gallery/7nPezEc [accessed September 12, 2022], **ii.** Sourced from "Lesotho's national transport – a brief history of the Basotho pony", by Knapp, F. (2021). The Design Train. Available: https://www.thedesigntrain.com/post/the-basotho-pony-lesotho-s-national-transport-a-brief-look-at-this-fascinating-history [accessed September 12, 2022] and **iii.** Sourced from "Basuto pony (Basotho pony)", by Garfield, J. (2018). Horse Breeds Pictures. Available: https://www.horsebreedspictures.com/basuto-pony-basotho-pony.asp [accessed September 12, 2022]].

The phenotypic traits that depict the Basotho breed consist of a small-frame and small-stature of not more than 14.2hh (which categorizes this breed as a 'pony'), the adaption of hardy hooves in order to travel through rocky, mountainous terrain as well as the observed solid coat colours - which include bay, chestnut (red-orange colour), grey and black (Law, 2020). The Basotho pony was historically utilized by the Sotho warriors as a warhorse during the colonization period of South Africa (**i.**) and was also utilized by the Sotho people for transportation purposes – which became culturally symbolized by the Sotho people in their mounted culture (Millin, 2007). The societal roles that this breed plays within the modern era has been retained (**ii.**), in order to boost the economic and socio-economic status of Lesotho through the tourism sector, by providing trail rides to foreigners through the mountainous districts of Lesotho (**iii.**). Along with the tourism sector input, several breeding programs were established in the late 1900's to increase the population size of these workhorses, whereby Sotho farmers then took it upon themselves to continue the breeding of the Basotho pony. The Basotho pony fulfills major roles in the economic, socio-economic and cultural aspects of society – historically and presently.

2.4.2 Nooitgedacht

The Nooitgedacht horse breed was developed in 1951 to restore the existence of the fundamental Basotho pony population that verged extinction, all of which alludes to the direct translation of the Dutch/Afrikaans "Nooitgedacht" to "never-imagined" – as breeders would have never imagined how distinctive this breed would become in South African history (Smith, 2021). Knapp (2020) indicates that the Nooitgedacht horse breed was first named "Nooitgedacht pony", due to derivative of the Nooitgedacht from the selected origin nucleus of Basotho ponies, that did not reach over 15hh (hands) and therefore were categorized as ponies. However, as the breed progressed throughout the years, the resulting progeny reached heights of up to 16.2 hh and it was decided to change the name of the breed to "Nooitgedacht Horse". Burgess (2014) describes the Nooitgedacht horse breed to be versatile in their abilities as the breed is described to be agile and has the ability to be trained in several disciplines. The phenotypic traits of the historical and modern-day Nooitgedacht breed can be seen in **Figure 2.3** below.



Figure 2.3 The Nooitgedacht breed that was the first indigenous South African horse breed to be recognized and recorded within the South African Studbook Association (SA Studbook)

[The Nooitgedacht horse breed, which was once recognized as a pony breed, was the first indigenous South African horse breed to be registered within the South African Studbook Association (SA Studbook). Image **i.** shows a herd of Nooitgedacht ponies that indicates their historically small stature and previous categorization as a pony. Image **ii.** indicates one of the modern-day versatile abilities of the Nooitgedacht horse to be involved in the equine sport of show jumping.) [**i**. Sourced from "Contributing to the Nooitgedachter legacy", by Burgess, M. (2014). Farmer's Weekly. Available: https://www.farmersweekly.co.za/rural-insight/contributing-to-the-nooitgedachter-legacy/ [accessed September 13, 2022], **ii.** Sourced from "Nooities in the competition field", by McCabe, C. (2011). SA Horseman, 32-33.]

The typical phenotypic characteristics of the Nooitgedacht horse breed presents the breed's hardiness through their small but strong, muscular build that enables the breed to carry the weight of a rider over long distances, their sturdy hooves that promote 'surefootedness', as well as the evidence of their kind, doe-like eyes that associates with their soft nature (Ruhr, 2008). Image **(i.)** indicates the historical categorization of the Nooitgedacht as a pony breed that was utilized for agricultural activities alongside Dutch farmers, that included patrolling fences, hunting in the bushveld and their endurance was utilized for transport between farmlands (Smith, 2021). The Nooitgedacht horse breed is described as five-gaited, whereby identified South African gaits refer to the walk, trot, slow-gait, rack and canter (Bekker, 2009) – which enables the indigenous breed to dominate the modern-day equine sport world by partaking in show jumping (**ii.**), eventing, dressage and gymkhana. Smith (2021) drew references to the breed's economic, socio-economic and cultural value within South African society and therefore designation of the Nooitgedacht Horse Breeder's Society (http://www.nooitgedachter.co.za/) formed in 1968, dedicated to genetically improve and preserve the breed.

2.4.3 South African Boerperd

The South African (SA) Boerperd is one of the three indigenous horse breeds that represents the infamous history of South Africa, from nobility and bravery of this breed during warfare to the accumulation of natural selection adaptions and breeding strategies over the past three centuries, which has shaped the distinctive and versatile Boerperd breed of today (Coleman, 2008). Nel (2015) depicts that the development of the modern-day Boerperd took place in three phases, which includes the establishment of the original Cape Horse during the Dutch colonization in 1652, whereby influence from European importations was used to improve the Cape Horse breed. The second phase included the fixation of the genotypic composition and the resulting phenotypic manifestation of the Boerperd during the Great Trek of the Dutch settlers inland as well as the influence of a few British indigenous breeds that were bred into the populations. The third and last phase included the period after the last Anglo-Boer war whereby only the hardy horses survived after the war and the harsh environmental factors that plagued the inland bushveld.



Figure 2.4 The historical and modern-day phenotypic appearance of the South African Boerperd horse breed that was genetically shaped throughout South African history

[The historical and modern-day phenotypical appearance of the South African Boerperd directly relates to the historical workhorse utilization by the Dutch (**i**) and the progressive selective breeding of this breed for performance-related equine sports within the modern era (**ii**).) [i. Sourced from "Boer horse or Boerperd", by Kemp, K. (2020). Horse Hints. Available: http://horsehints.org/Breeds/Boer.htm [accessed September 13, 2022], ii. Sourced from "Boerperd horse", by Gainsbourg, M. (2018). Horse Breeds Pictures. Available: https://www.horsebreedspictures.com/boerperd-horse.asp [accessed September 13, 2022]].

The comparison between the historical (**i.**) and modern-day (**ii.**) phenotypical form of the South African Boerperd can be seen in **Figure 2.4**, whereby the small-framed, muscular and compact build of the Boerperd was previously selected for agricultural and transportation purposes by the Dutch farmers (Coleman, 2008). The modern-day Boerperd trait selection criteria has led to two types of Boerperd groups, in which according to Mac (2020), the groups are based on movement ability and confirmation. The classified traditional type of the Boerperd was selected to possess a higher action that accompanies their five-gaited movement, which is utilized in an equine sport known as saddle seat riding - affiliated amongst the sporting federations of South Africa. On the other hand, the common grouping of the Boerperd breed was selected for a conservatively lower

movement ability, which serves the purpose of covering ground in order to be utilized for equine sports that include show jumping, dressage and eventing. These two selected Boerperd group types were generated through the designated breeding society dedicated to the improvement of this breed that was established in 1973, deemed as the Boerperd Society of South Africa (http://www.saboerperd.com/), which propagated the recognition of this indigenous breed with the South African Studbook Association (SA Studbook) in 1980 (Nel, 2015). The efforts of the Boerperd Society of South Africa has ensured the economic, socio-economic and the cultural roles of the Boerperd breed to remain prominent within society.

2.4.4 English Thoroughbred

The English Thoroughbred is one of the world's most influential horse breeds through their domination in equine sports due to their versatility across multiple disciplines and their global impact on genetic improvement of breeds, such as their role in the development of our indigenous South African breeds (Herringer, 2004). According to Standish et al. (2011), this breed, culturally significant to the British was used in flat racing events. The first documented horse race took place Cape Town in 1797, which propagated the horse racing culture in South African society. The historical and modern-day phenotypic form of the English Thoroughbred can be seen in **Figure 1.5** below.

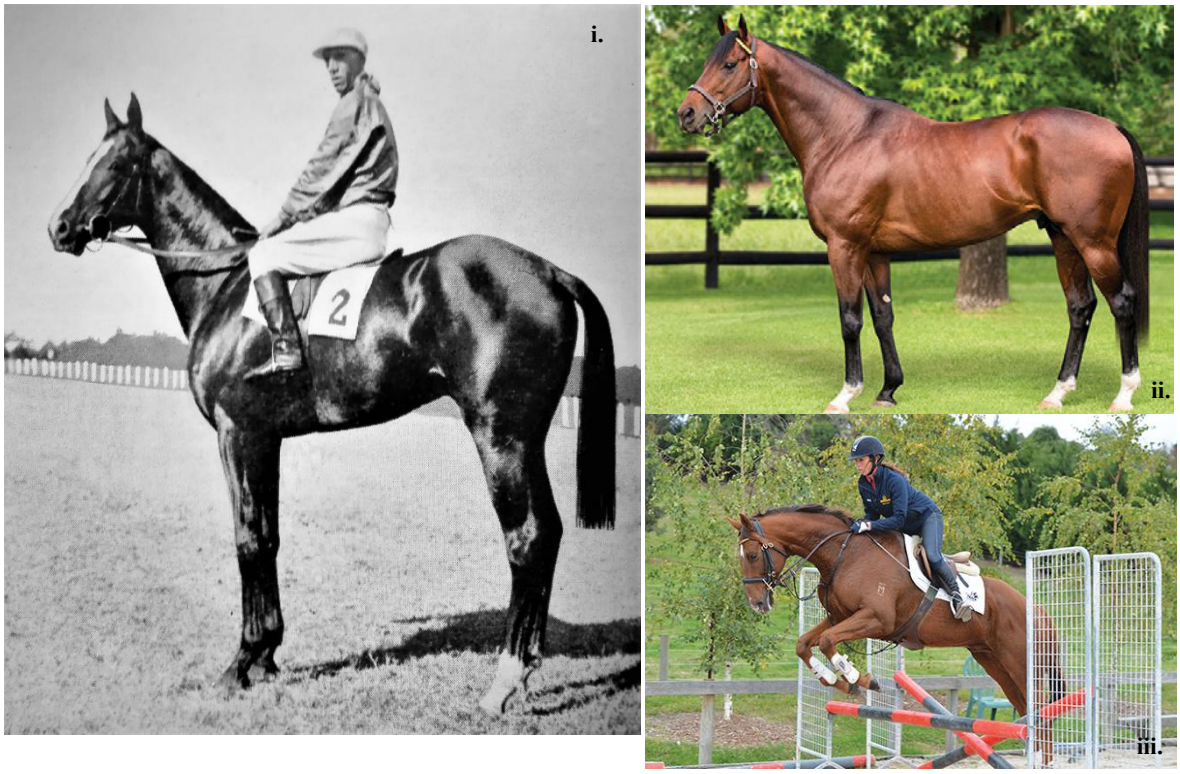


Figure 2.5 The historical and modern-day phenotypic form of the English Thoroughbred selected for specific athletic ability within the equine sporting world

[The English Thoroughbred was utilized within the British Empire colonization era in order to improve the Cape Horse that served as a warhorse to the British calvary, which proved to be unsuccessful in the first Anglo-Boer war (1899-1902), whereby the British calvary was defeated by the Boers and their hardy Boerperd horse breed. Therefore, the use of the English Thoroughbred proved to not be successful in producing hardy warhorse populations and the focus of this breed was switched to the immense athletic ability and endurance the breed was characterized by, that was exploited in order to create the horse racing industry of South Africa (which followed suit of the horseracing culture within Britain). Image **i.** illustrates the phenotypical form of the historical English Thoroughbred selected for horseracing within South Africa and lightweight rider on the horses' back - deemed as a 'jockey'. Image **ii.** indicates the phenotypical result of the modern-day selection criteria of the English Thoroughbred used in horse racing. Image **iii.** presents the multi-disciplinary ability of the English Thoroughbred that usually occurs after the horses' racing career has ended.) [**i.** Sourced from "South Africa's leading sires 1900-1971", by August, J. (2020). The South African Racehorse. Available: https://southafricanracehorse.wordpress.com/2020/05/24/south-africas-top-sires-1900-1970/ [accessed September 14, 2022], **ii.** Sourced from "The leading SA sires so far", by Whitelaw, S. (2020). Sporting Post. Available: https://www.sportingpost.co.za/breeding/sa-sires-this-season/ [accessed September 14, 2022], **iii.** Sourced from "Retraining the off-the-trach Thoroughbred as a jumper: Part two", by Hector, C. (2021b). The Horse Magazine. Available: https://www.horsemagazine.com/thm/2021/05/re-training-the-off-the-track-thoroughbred-as-a-jumper-part-two/ [accessed September 14, 2022]].

The English Thoroughbred is one of the most selectively bred horse breeds, with intense breeding programs aimed at producing progeny with desirable phenotypic traits. These traits include long, slender legs (as shown in **Figure 2.5** (**i.**) above), which enables the horse to cover large distances with each stride; well-muscled hindquarters that provide powerful propulsion; and a lean body frame that enhances aerodynamic efficiency, allowing the horse to gallop at an estimated average speed of 56 km per hour (Henry, 2020). Over centuries of intense selective breeding, the focus of trait selection in Thoroughbreds has led to the development of two distinct body types. Breeding for speed in short-distance sprints has produced compact, muscular sprinters

(**Figure 2.5** (**ii.**)), while breeding for stamina in long-distance races has resulted in lengthy, lean distance runners (**Figure 2.5** (**i**.)) (Robertson-Smith, 2012). The variation of these two racehorse types has been determined to be the result of the selection of the Myostatin (*MSTN*) gene and associated SNPs (Single Nucleotide Polymorphisms) that are responsible for the regulation of skeletal muscle growth – reported by Rooney et al. (2018). Rooney et al. (2018) identified three genotypes linked to short- and long-distance racehorse types, which influence muscle mass expressed in the progeny. The CC genotype produces progeny with large muscle mass, promoting speed and short-distance racing. The TT genotype results in lean muscle mass, enhancing stamina for long-distance races. Lastly, the heterozygous CT genotype produces progeny with moderate muscle mass, making them well suited for middle-distance races. This advantageous selective breeding strategy, along with several other trait selection criteria, has been the focus of many Thoroughbred breeding studs, governed by the Thoroughbred Breeders' Association of South Africa (https://www.tba.co.za/), which aims to ensure the preservation of the economic, socio-economic and cultural roles that this horse breed fulfils within South Africa.

2.4.5 Saddlebred Horse

The classic American Saddlebred was the resulting breed composite of the English Thoroughbred and the recently extinct Narragansett Pacer horse breed, that originated within the early eighteenth century (Salem, 2022). The precursor breeds contributed distinctive characteristics to the American Saddlebred, including stamina, speed and agility from the Thoroughbred and the three- and five-gaited ability from the Narragansett Pacer. The American Saddlebred was originally referred to as "the horse America made", whereby the breed was further improved through the use of more Thoroughbred influence within the Kentucky State, and the resulting progeny became the popular warhorse of choice for mounted generals during the civil war (1861-1865) (Cerulli, 2020). Millin (2011) owes the establishment of this American Saddlebred within South Africa to a single horseman named Claude Orphen, who imported two historically significant Saddlebred stallions "Myer's Kentucky Star 7675" and "Red Domino 8313" in 1916. These stallions were bred to selected Thoroughbred mares, as well as indigenous South African mares, in the effort to improve our indigenous breed populations. The historical and modern-day phenotypic form of the Saddlebred differs slightly, due to the selective breeding improvement of the horse throughout the past few decades, in order to perform highly competitive equestrian sport disciplines – as seen in **Figure 2.6** below.



Figure 2.6 The historical and modern-day phenotypical form of the Saddlebred that associated with their breed utilization throughout the centuries

[The illustrations above draw the comparison of the phenotypic form of the historical and modern-day breeding objectives of this American-borne Saddlebred. Image i. presents the utilization of the American Saddlebred as a warhorse during the Civil War (1861-1865). Image ii. shows the use of the South African bred Saddlebred in the Olympic, gaited equine sport known as Saddle Seat Equitation. Image iii. shows the second-most favoured Olympic equestrian sport known as harness (single and fine harness) showing that this breed is bred for. [i. Sourced from "Horses in history: Famous horses of the civil war", by Kovatch, K. (2016). Horse Nation. Available: https://www.horsenation.com/2016/02/29/horses-in-history-famous-horses-of-the-civil-war/ [accessed September 16, 2022], ii. Sourced from "Saddlebred Society of South Africa" (http://www.saddlebred.co.za/p14/about-the-breed/history-of-the-saddlebred.html) [accessed September 16, 2022], iii. Sourced from "Saddlebred Society of South Africa" (http://www.saddlebred.co.za/p14/about-the-breed/history-of-the-saddlebred.html) [accessed September 16, 2022]].

The iconic phenotypic appearance of the American-borne saddle breed (**Figure 2.6 (i)**), that was historically comprised of their solid muscle development throughout their body composition and their high-gaited position, was extremely advantageous during the cavalry attacks during the civil war and became a very noble breed within the American culture (Tavano, 2011). The modern-day phenotypic appearance (**Figure 2.6 (ii)** and **(iii)**) of the South African Saddlebred is the result of intensive breed improvement standards that focus on producing champions to dominate highly competitive equine sports. These phenotypic traits are based on gaited ability, lean-muscle build, high-arched necks, extended poll flexion of the hind legs and deep sloping shoulders (Cerulli, 2020) – all contributing to the iconic, elegant breed characteristics of the modern-day Saddlebred. Cerulli (2020) also indicates that the intensive selection of this breed has led to associated phenotypic deformities that include a prominent swayed-back, where the spine of the horse caves in from the protruding wither, that may be linked to the extended poll flexion of the hind legs. However, these breeding deformities are not detrimental to the horses during the competing phase of their life and therefore the continuation of

associated economic, socio-economic and cultural roles of this iconic breed was ensured through the formation of the Saddlebred Society of South Africa in 1949 (http://www.saddlebred.co.za/p20/the-society/saddlebred-society-of-south-africa-history.html).

The historic and modern-day economic, socio-economic and cultural roles of our three indigenous horse breeds and selected precursor horse breeds indicate their importance and provides evidence of our need to research and preserve our South African breeds.

2.5 Evaluation of previously known and future potential studies conducted on the genetic diversity of indigenous horse populations within South Africa

There have been very limited genomic studies conducted on our indigenous South African horse breed populations, due to the fact that the one documented study that examined the genetic population structure of our indigenous SA Boerperd, Nooitgedacht and Basotho pony, was conducted by Cothran and van Dyk in 1998. Other studies on these three indigenous horse breeds have primarily explored their history and establishment within South Africa, such as the work by van der Merwe and Martin (2002). However, they have not necessarily assessed the current genetic structure of these populations, which could provide valuable insights into their present state and inform potential conservation efforts through targeted breeding programs. Van der Merwe and Martin (2002) did however present support of the controversial division of the Boerperd breed into two separate breeds, known as the South African Boerperd and the Cape Boerperd, which most researchers as only one breed.

2.5.1 Previous studies conducted that assess the genetic diversity of the indigenous horse populations

Cothran and van Dyk (1998) conducted a genetic analysis study that assessed the genetic diversity of the South African Boerperd, Nooitgedacht, Basotho pony and various predecessor exotic horse breeds. Genetic diversity of these populations was observed through estimation of allelic variations of 7 blood-group loci and 10 genetic loci biochemical annotations belonging to these three sampled indigenous populations. The findings from the genetic analyses revealed observed heterozygosity (Ho) based on individual allele frequencies, using collective exotic precursor breeds – including the English Thoroughbred, Arabian, American Saddler and other exotic populations – as the reference standard ($Ho_{(mean)} = 0.377$). Among the indigenous breeds, the Boerperd exhibited the highest heterozygosity (Ho = 0.385), while the Basotho pony had the lowest (Ho = 3.47). Although this difference from the exotic population Ho(mean) is not substantial; it highlights the lower levels of heterozygosity in the Basotho pony. This reduced genetic diversity may be attributed to their isolated development within the Kingdom of Lesotho (Lekota, 2014), where high levels of inbreeding could have contributed to the increased homozygosity observed in the sampled Basotho pony population. Another major finding within this study was the strict consensus phylogenetic clustering of these three indigenous breed populations along with 52 other exotic horse breed populations, whereby the Basotho pony and Nooitgedacht breed populations showed a closer relationship between other European breeds (distinctively the English Thoroughbred). The South African Boerperd showed a closer relationship to the North American gaited breeds

(distinctively the American Saddlebred and the American Standardbred). Lastly, another major finding of the genetic analyses indicated that the three indigenous breed populations showed closer relatedness to one another, than any other of the 52 other exotic breed populations – confirming that these three indigenous populations originated from a common ancestor. These three major findings provided a basis of the genetic population structure assessment of our indigenous breed populations. However, our indigenous horse populations need further assessment and investigation through the use of high-throughput genomic technologies at a genome-wide level.

2.5.2 Future potential genomic studies that can be conducted to examine the genetic diversity of the indigenous horse breed populations

The advancement of genomic sequencing and genotyping technologies over the past few decades have allowed for extensive and accurate assessments of the genetic population structure of livestock populations, at a genome-wide level, within the Agricultural sector of South Africa (Van Marle-Koster and Visser, 2018). Advances in Single Nucleotide Polymorphism (SNP) genotyping technologies, such as the high-throughput Illumina SNP BeadChip array (Illumina, USA), have enabled the efficient generation of large-scale genomic datasets for numerous livestock species worldwide (Fan et al., 2010). Of particular relevance to this study was the development of BeadChip arrays specifically designed for domesticated horses and other equine species. The development of the Illumina EquineSNP70K BeadChip, allows for the detection of almost up to 70,000 SNPs across the equine genome, which has been utilized in several domesticated horse breed studies, such as a study conducted by Finno et al. (2018). This study used the BeadChip to identify mutations in the candidate gene *MYH1* (Myosin-1 gene), that is associated with myositis (the immune-related weakening of muscles), within sampled Quarter horse breed populations. The use of the Illumina EquineSNP70K BeadChip is not only extremely beneficial in identifying candidate genes and selection signatures found throughout the domesticated horse genome, but can also be utilized to produce large datasets simultaneously from various horse breed populations, in order to subsequently conduct multiple genetic population analyses. Some of these analyses include ancestry inference through Principal Component Analysis (PCA) (Jolliffe and Cadima, 2016), admixture proportions (Wang, 2022), assessment of inbreeding levels through Runs of Homozygosity (ROH) (Curik et al., 2014) and Extended Haplotype Homozygosity (EHH) tests for candidate gene identification to determine whether selection within a population is positive or neutral (Zhong et al., 2010). Conducting these analyses on South African horse breeds would be essential for accurately assessing their genetic structure. The findings from the determination of the genetic diversity and candidate gene identification can provide evidence for the preservation of our indigenous and rural breed populations – of which play extensive economic, socio-economic and cultural roles within South Africa.

2.6 Conclusion

The systematic dissection of the evolutionary history and subsequent establishment of the indigenous South African horse breeds within this paper highlights the fact that these breeds represent the rich history of South Africa through their historical economic, socio-economic and cultural roles played within society throughout the centuries. A comparison of the vital modern-day indigenous and subsequent crossbred rural breed

populations emphasizes the need for further studies to be conducted on these populations - in order to assess the state of genetic diversity as potential indicator for preservation efforts that may be needed.

2.7 References

Aldridge, L. (2000) in "The genetics of the horse", Hereditary 85(509). Available: https://doi.org/10.1046/j.1365-2540.2000.0827a.x [accessed September 01, 2022].

Allentoft, M.E., Sikora, M., Sjögren, K.J., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., Malaspinas, A.S., Margaryan, A., Higham, T., Chivall, D., Lynnerup, D., Harvig, L., Baron, J., Casa, P.D., Dąbrowski, P., Duffy, P.R., Ebel, A.V., Epimakhov, A., Frei, K., Furmanek, M., Gralak, T., Gromov, A., Gronkiewicz, S., Grupe, G., Hajdu, T., Jarysz, R., Khartanovich, V., Caramello, C. (2016). Equines in the Anglo-Boer and Great Wars. Horsetalk. Available: https://www.horsetalk.co.nz/2016/02/27/theirs-not-to-reason-why-horsing-the-british-army/ [accessed September 09, 2022].

August, J. (2020). South Africa's leading sires 1900-1971. The South African Racehorse. Available: https://southafricanracehorse.wordpress.com/2020/05/24/south-africas-top-sires-1900-1970/ [accessed September 14, 2022],

Bekker, L. (2009). Purity of gait. SA Horseman, 18-22.

Blocksdorf, K. (2022). Arabian horse: Breed profile. The Spruce Pets. Available: https://www.thesprucepets.com/meet-the-arabian-horse-1886131 [accessed September 15, 2022].

Burgess, M. (2014). Contributing to the Nooitgedachter legacy. Farmer's Weekly. Available: https://www.farmersweekly.co.za/rural-insight/contributing-to-the-nooitgedachter-legacy/ [accessed September 13, 2022].

Cassella, C. (2021). Genetic tracing reveals where the powerful ancestor of modern horses came from. Humans. Available: https://www.sciencealert.com/modern-horses-might-originate-from-powerful-and-docile-breed-in-ancient-russia [accessed August 28, 2022].

Cerulli, P. (2020). American Saddlebred: Horse breed profile. The Spruce Pets. Available: https://www.thesprucepets.com/american-saddlebred-horse-breed-profile-4783512 [accessed September 16, 2020].

Chappalwar, A., Sakhare, M., Katkuri, S. (2022). Breeds of Equine. Pashudhan Praharee. Available: https://www.pashudhanpraharee.com/breeds-of-equine/ [accessed August 30, 2022].

Coleman, A. (2008). The sturdy bloodline of the SA Boerperd. Farmer's Weekly. Available: https://www.farmersweekly.co.za/archive/the-sturdy-bloodline-of-the-sa-boerperd/ [accessed September 13, 2022].

Cothran, E.G., Podhajsky, A.W. (2022). Horse. *Encyclopaedia Britannica*. Available: https://www.britannica.com/animal/horse [accessed August 30, 2022].

Cothran, E. G., Van Dyk, E. (1998). Genetic analysis of three South African horse breeds. Journal of the South African Veterinary Association 69(4): 120–125.

Curik, I., Ferencakovic, M., Solkner, J. (2014). Inbreeding and runs of homozygosity: A possible solution to an old problem. Livestock Science 166, 26-34.

Delsol, N., Stucky, B.J., Oswald, J.A., Reitz, E.J., Emery, K.F., Guralnick, R. (2022). Analysis of the earliest complete mtDNA genome of a Caribbean colonial horse (Equus caballus) for 16th century Haiti. PLOS ONE 17(7). Available: https://doi.org/10.1371/journal.pone.0270600 [accessed September 01, 2022].

Dennis, S.J., Meyers, A.E., Hitzeroth, I.I., Rybicki, E.P. (2019). African Horse Sickness: A review of current understanding and vaccine development. Viruses 11(9). Available: https://doi.org/10.3390%2Fv11090844 [accessed September 08, 2022].

Eser, E., Erat, S. (2022). Domestication and origins of the modern horse breeds. Journal of Bahri Dagdas Animal Research 11(1), 51-61.

Fan, B., Du, Z.Q., Gorbach, D.M., Rothschild, M.F. (2010). Development and application of the high-density SNP arrays in genomic studies of domesticated animals. Asian-Australian Journal of Animal Science 23(7), 833-847.

Ferriss, J. (2003). The concept of strains in Arabian horses. Pyramid Society. Available: https://www.pyramidsociety.org/sites/default/files/ferriss_strainhandoutsgs_0.pdf [accessed September 15, 2022].

Finno, C.J., Gianino, G., Perumbakkam, S., Williams, Z.J., Bordbari, M.H., Gardner, K.L., Burns, E., Peng, S., Durward-Akhurst, S.A., Valberg, S.J. (2018). A missense mutation in the MYH1 is associated with susceptibility to immune-mediated myositis in Quarter horses. Skeletal Muscle 8(7). Available: https://doi.org/10.1186%2Fs13395-018-0155-0 [aaccessed September 17, 2022].

Fourie, J. (2020). The settlers of South Africa and the expanding frontier. Stellenbosch Economic Working Papers: WP14/2020. Available: http://www.ekon.sun.ac.za/wpapers/2020/wp142020 [accessed September 08, 2022].

Gainsbourg, M. (2018). Boerperd horse. Horse Breeds Pictures. Available: https://www.horsebreedspictures.com/boerperd-horse.asp [accessed September 13, 2022]

Garfield, J. (2018). Basuto pony (Basotho pony). Horse Breeds Pictures. Available: https://www.horsebreedspictures.com/basuto-pony-basotho-pony.asp [accessed September 12, 2022].

Griffin, A. (2020). What s a horse breed? Extension Horses. Available: https://horses.extension.org/what-is-a-horse-breed/ [accessed August 30, 2022].

Guimaraes, S., Arbuckle, B.S., Peters, J., Adcock, S.E., Buitenhuis, H., Chazin, H., Manaseryan, N., Uerpmann, H.P., Grange, T., Geigl, E.M. (2020). Ancient DNA shows domestic horses were introduced in the southern Caucasus and Anatolia during the Bronze Age. Science Advances 6(38). Available: https://doi.org/10.1126/sciadv.abb0030 [accessed August 28, 2022].

Hanof, P., Herrel, A., Guintard, C., Cornette, R. (2018). The impact of artificial selection on morphological integration in the appendicular skeleton of domestic horses. Journal of Anatomy 232(4), 657-673.

Hector, C. (2021a). Confirmation and the sport horse. The Horse Magazine. Available: https://www.horsemagazine.com/thm/2021/07/conformation-and-the-sport-horse/ [accessed September 12, 2022].

Hector, C. (2021b). Retraining the off-the-trach Thoroughbred as a jumper: Part two. The Horse Magazine. Available: https://www.horsemagazine.com/thm/2021/05/re-training-the-off-the-track-thoroughbred-as-a-jumper-part-two/ [accessed September 14, 2022].

Henry, M. (2019). The Arabian horse breed: Its history and facts revealed. Horse Racing Sense. Available: https://horseracingsense.com/discover-the-arabian-horse-past-to-present/ [accessed September 15, 2022].

Henry, M. (2020). Thoroughbred horse breeds: Facts, height, and characteristics. Horse Racing Sense. Available: https://horseracingsense.com/thoroughbred-horse-breed-facts/ [accessed September 15, 2022].

Henry, M. (2021). Work horses – breeds, types and characteristics. Horse Racing Sense. Available: https://horseracingsense.com/work-horses-breeds-types-characteristics/ [accessed September 12, 2022].

Herringer, P. (2004). Thoroughbred horse racing and breeding in South Africa. Thoroughbred Heritage. Available: https://www.tbheritage.com/Breeders/SABreeding/SABreeders.html [accessed September 13, 2022].

Hofmeyr, S., Scholtz, M. (2017). The Nooitgedacht-Iberian. HQ Magazine. Available: https://hqmagazine.co.za/nooitgedacht-iberian/ [accessed September 09, 2022].

Hunt, J., (2005) in "Dutch South Africa: early settlers at the Cape, 1652-1708", Troubador Publishing Ltd. Available: https://www.tandfonline.com/doi/abs/10.1111/j.1467-8306.1976.tb01070.x [accessed September 05, 2022].

Kadour, O.H. (2017). Rebel parts of Syria revive beloved Arabian horse races. Yahoo! News. Available: https://sg.news.yahoo.com/rebel-parts-syria-revive-beloved-arabian-horse-races-041936822.html [accessed September 15, 2022]

Kemp, K. (2020). Boer horse or Boerperd. Horse Hints. Available: http://horsehints.org/Breeds/Boer.htm [accessed September 13, 2022].

Khadka, R. (2010) in "Horse population, breeds and risk status in the world; A study based on Food and Agriculture Organization Database systems: FAOSTAT and DAD-IS", Lambert Academic Publishing, Germany. Available: https://www.researchgate.net/publication/261705436_Rupak_Khadka_2010_Horse_Population_Breeds_and_Risk_Status_in_the_World_A_study_based_on_Food_and_Agriculture_Organization_Database_systems_FAOSTAT_and_DAD-IS_Lambert_Academic_Publishing_Germany [accessed September 01, 2022].

Khokhlov, A., Kiss, V., Kolář, J., Kriiska, A., Lasak, I., Longhi, C., McGlynn, G., Merkevicius, A., Merkyte, I., Metspalu, M., Mkrtchyan, R., Moiseyev, V., Paja, L., Pálfi, G., Pokutta, D., Pospieszny, L., Price, T.D., Saag, L., Sablin, M., Shishlina, N., Smrčka, V., Soenov, V.I., Szeverényi, V., Tóth, G., Trifanova, S.V., Varul, L., Vicze, M., Yepiskoposyan, L., Zhitenev, V., Orlando, L., Sicheritz-Pontén, T., Brunak, S., Nielsen, R., Kristiansen, Willerslev, K. (2015). Population genomics of Bronze Age Eurasia. *Nature* 522, 167–172.

Knapp, A. (2020). Horse breeds indigenous to South Africa – a noble heritage. The Design Train. Available: https://www.thedesigntrain.com/post/horse-breeds-indigenous-to-south-africa-a-noble-heritage [accessed September 05, 2022].

Knapp, A. (2021). Lesotho's national transport – a brief history of the Basotho pony. The Design Train. Available: https://www.thedesigntrain.com/post/the-basotho-pony-lesotho-s-national-transport-a-brief-look-at-this-fascinating-history [accessed September 08, 2022].

Kovatch, K. (2016). Horses in history: Famous horses of the civil war. Horse Nation. Available: https://www.horsenation.com/2016/02/29/horses-in-history-famous-horses-of-the-civil-war/ [accessed September 16, 2022]

Jacobs, N. (2011). Review of Swart, Sandra. Riding high: Horses, humans and history in South Africa. H-Net Reviews in the Humanities and Social Sciences. Available: https://www.h-net.org/reviews/showrev.php?id=32129 [accessed September 09, 2022].

Janis, C.M., Bernor, R.L. (2019). The evolution of equid monodactyly: a review including a new hypothesis. Frontiers in Ecology and Evolution. Available: https://doi.org/10.3389/fevo.2019.00119 [accessed August 25, 2022].

Jones, O. (2022). 20 most popular racehorse breeds. PetKeen. Available: https://petkeen.com/most-popular-racehorse-breeds/ [accessed September 09, 2022].

Jansen, T., Forster, P., Levine, M.A., Oelke, H., Hurles, M., Renfrew, C., Weber, J., Olek, K. (2002). Mitochondrial DNA and the origins of the domestic horse. PNAS 99(16), 10905-10910.

Jolliffe, I.T., Cadima, J. (2016). Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society A 374(2065). Available: https://doi.org/10.1098/rsta.2015.0202 [accessed September 17, 2022].

Law, O. (2020). Breeds from around the world: Part 4 – Basotho pony. In the Saddle. Available: https://www.inthesaddle.com/breeds-from-around-the-world-part-four-basotho-pony/ [accessed September 12, 2022].

Lekota, T.A. (2014). The state of the Basotho pony in Lesotho. FAO. Available: https://www.fao.org/fileadmin/user_upload/nr/sustainability_pathways/docs/The%20State%20of%20the%20Basotho%20Pony%20in%20Lesotho.pdf [accessed September 12, 2022].

Librado, P., Fages, A., Gaunitz, C., Leonardi, M., Wagner, S., Khan, N., Hanghoi, K., Alquraishi, S.A., Alfarhan, A.H., Al-Rasheid, K., Der Sarkissian, C., Schubert, M., Orlando, L. (2016). The evolutionary origin and genetic makeup of the domestic horse. Genetics 204(2), 423-434.

Librado P., Khan N., Fages A., Kusliy M. A., Suchan T., Tonasso-Calvière L., Schiavinato S., Alioglu D., Fromentier A., Perdereau A., Aury J.-M., Gaunitz C., Chauvey L., Seguin-Orlando A., Der Sarkissian C., Southon J., Shapiro B., Tishkin A. A., Kovalev A. A., Alquraishi S., Alfarhan A. H., Al-Rasheid K. A. S., Seregély T., Klassen L., Iversen R., Bignon-Lau O., Bodu P., Olive M., Castel J.-C., Boudadi-Maligne M., Alvarez N., Germonpré M., Moskal-del Hoyo M., Wilczyński J., Pospuła S., Lasota-Kuś A., Tunia K., Nowak M., Rannamäe E., Saarma U., Boeskorov G., Lõugas L., Kyselý R., Peške L., Bălăşescu A., Dumitraşcu V., Dobrescu R., Gerber D., Kiss V., Szécsényi-Nagy A., Mende B. G., Gallina Z., Somogyi K., Kulcsár G., Gál E., Bendrey R., Allentoft M. E., Sirbu G., Dergachev V., Shephard H., Tomadini N., Grouard S., Kasparov A., Basilyan A. E., Anisimov M. A., Nikolskiy P. A., Pavlova E. Y., Pitulko V., Brem G., Wallner B., Schwall C., Keller M., Kitagawa K., Bessudnov A. N., Bessudnov A., Taylor W., Magail J., Gantulga J.-O., Bayarsaikhan J., Erdenebaatar D., Tabaldiev K., Mijiddorj E., Boldgiv B., Tsagaan T., Pruvost M., Olsen S., Makarewicz C. A., Valenzuela Lamas S., Albizuri Canadell S., Nieto Espinet A., Iborra M. P., Lira Garrido J., Rodríguez González E., Celestino S., Olària C., Arsuaga J. L., Kotova N., Pryor A., Crabtree P., Zhumatayev R., Toleubaev A., Morgunova N. L., Kuznetsova T., Lordkipanize D., Marzullo M., Prato O., Bagnasco Gianni G., Tecchiati U., Clavel B., Lepetz S., Davoudi H., Mashkour M., Berezina N. Ya., Stockhammer P. W., Krause J., Haak W., Morales-Muñiz A., Benecke N., Hofreiter M., Ludwig A., Graphodatsky A. S., Peters J., Kiryushin K. Yu., Iderkhangai T.-O., Bokovenko N. A., Vasiliev S. K., Seregin N. N., Chugunov K. V., Plasteeva N. A., Baryshnikov G. F., Petrova E., Sablin M., Ananyevskaya E., Logvin A., Shevnina I., Logvin V., Kalieva S., Loman V., Kukushkin I., Merz I., Merz V., Sakenov S., Varfolomeyev V., Usmanova E., Zaibert V., Arbuckle B., Belinskiy A. B., Kalmykov A., Reinhold S., Hansen S., Yudin A. I., Vybornov

A. A., Epimakhov A., Berezina N. S., Roslyakova N., Kosintsev P. A., Kuznetsov P. F., Anthony D., Kroonen G. J., Kristiansen K., Wincker P., Outram A., Orlando L. (2021). The origins and spread of domestic horses from the Western Eurasian steppes. Nature 598, 634–640.

Mac, Dr. (2020). Strict selection criteria are vital to Boerperd breeding. Farmer's Weekly. Available: https://www.farmersweekly.co.za/animals/horses/strict-selection-criteria-are-vital-to-boerperd-breeding/ [accessed September 13, 2022].

MacFadden, B.J., Ovideo, L.H., Seymour, G.M., Ellis, S. (2012). Fossil horses, orthogenesis, and communicating evolution in museums. Evolution: Education and Outreach 5(29). Available: https://doi.org/10.1007/s12052-012-0394-1 [accessed May 15, 2025].

Marques, A. (2022). Arabian horse breed history, characteristics and uses. Available: https://horseyhooves.com/arabian-horse-breed/ [accessed September 15, 2022].

Mattocks, S. (2017). The Norfolk and Suffolk Arabian horse group annual show. The Arabian Magazine. Available: https://www.thearabianmagazine.com/the-norfolk-and-suffolk-arabian-horse-group-annual-show [accessed September 15, 2022]

McCormack, J. (2018). PSXVI-14 The Basotho pony. A breed improvement program in the highlands of Lesotho. Journal of Animal Science 96(3). Available: https://doi.org/10.1093/jas/sky404.551 [accessed May 15, 2025].

Millin, P. (2007). The renowned Basotho pony. SA Horseman. Available: https://journals.co.za/doi/pdf/10.10520/EJC14081 [accessed September 08, 2022].

Millin, R. (2011). Saddle horse breeding in South Africa. SA Horseman, 10-13.

Nel, C. (2015). The origin of the SA Boerperd. SA Boerperd – The Ideal Sporting and Pleasure Horse. Available: http://www.saboerperd.com/a3/general/the-origin-of-the-sa-boerperd.html [accessed September 08, 2022].

Oliver, E., Oliver, W.H. (2017). The colonization of South Africa: A unique case. Theological Studies 73(3). Available: https://doi.org/10.4102/hts.v73i3.4498 [accessed September 05, 2022].

Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P.W., Avila-Arcos, M.C., Fu, Q., Kruase, J., Willerslev, E., Stone, A.C., Warinner, C. (2021). Ancient DNA analysis. Nature Reviews Methods Primers 14. Available: https://doi.org/10.1038/s43586-020-00011-0 [accessed May 15, 2025].

Pinnock, D. (2022). From rock stars to obscurity – searching for lost Thoroughbreds. Daily Maverick. Available: https://www.dailymaverick.co.za/article/2022-06-28-from-rock-stars-to-obscurity-searching-for-lost-thoroughbreds/ [accessed September 10, 2022].

Price, M. (2017). Most modern horses came from just two ancient lineages. Science. Available: https://doi.org/10.1126/science.aan7044 [accessed August 29, 2022].

Prince, A., Geor, R., Harris, K., Hoekstra, S., Gardner, S., Hudson, C., Pagan, J. (2010). Comparison of the metabolic responses of trained Arabians and Thoroughbreds during high- and low- intensity exercise. Equine Veterinary Journal 34, 95-99.

Richards, E. (2005). SPARKS is the shining light for Cleveland Bays. The Chronicle of the Horse. Available: https://www.chronofhorse.com/article/sparks-shining-light-cleveland-bays [accessed September 09, 2022].

Robertson-Smith, G. (2012). A horse needs heart and right genes to go the distance. The Sydney Morning Herald. Available: https://www.smh.com.au/sport/racing/a-horse-needs-heart-and-right-genes-to-go-the-distance-20120928-26qtg.html [accessed September 15, 2022].

Rooney, M.F., Hill EW, Kelly VP, Porter RK. (2018). The "speed gene" effect of myostatin arises in Thoroughbred horses due to a promoter proximal SINE insertion. PLoS One 13(10). Available: https://doi.org/10.1371%2Fjournal.pone.0205664 [accessed September 15, 2022].

Ropka-Molik, K., Stefaniuk-Szmukier, M., Musial, A.D., Velie, B.D. (2019). The genetics of racing performance in Arabian horses. International Journal of Genomics. Available: https://doi.org/10.1155%2F2019%2F9013239 [accessed September 15, 2022].

Ruhr, R. (2008). Nooitgedachters thrive in Namibia. SA Horseman, 39-41.

Salem, P. (2022). Breed profile: The American Saddlebred horse. Available: https://www.fei.org/stories/lifestyle/health-fitness/breed-profile-american-saddlebred-horse [accessed September 16, 2022].

Schreuder, P.J. (1915) in "The Cape Horse. It's origin, breeding and development in the Union of South Africa", PhD thesis, Cornell University, Ithaca. Available: https://archive.org/details/capehorseitsorig00schr/page/n1/mode/2up [accessed September 05, 2022].

Solounias, N., Danowitz, M., Stachtiaris, E., Khurana, A., Araim, M., Sayegh, M., Natale, J. (2018). The evolution and anatomy of the horse manus with an emphasis on digit reduction. Royal Society Open Science 5. Available: https://doi.org/10.1098/rsos.171782 [accessed August 25, 2022].

Standish, B., Boting, A., Swing, B. (2011). Strategic analysis of the sport of horse racing in South Africa (Phase 1: profile and economic contribution). Sporting Post. Available: https://www.sportingpost.co.za/horse-racing/strategic-analysis-of-the-sport-of-horse-racing-in-south-africa/#_ftn4 [accessed September 13, 2022].

Swart, S. (2010). "The world the horses made": A South African case study of writing animals into social history. International Review of Social History 55(2), 241-263.

Swart, S. (2003). Riding high – horses, power and settler society, c. 1654-1840. Kronos, Environmental History 29. Available: https://doi.org/10.2307/41056494 [accessed September 05, 2022].

Tavano, D. (2011). American Saddlebred horses. Hobby Farms. Available: https://www.hobbyfarms.com/american-saddlebred/ [accessed September 16, 2022].

Terry, M. (2022). The Hackney: the history and endangered status of this dynamic British breed. Horse and Hound. Available: https://www.horseandhound.co.uk/features/hackney-horse-factfile-795305 [accessed September 09, 2022].

Thiruvenkadan, A.K., Kandasamy, N., Panneerselvam, S. (2008). Coat colour inheritance in horses. Livestock Science 117(2), 109-129.

Todd, E.T., Fromentier, A., Sutcliffe, R., Collin, Y.R.H., Perdereau, A., Aury, J.M., Eche, C., Bouchez, O., Donnadieu, C., Wincker, P., Kalbfleisch, T., Petersen, J.L., Orlando, L. (2023). Imputed genomes of historical horses provide insights into modern breeding. iScience 26(7). Available: https://doi.org/10.1016/j.isci.2023.107104 [accessed May 15, 2025].

Usher, H.J. (1984). The charge of the light brigade and South Africa's very own pony. South African Journal of Military Studies 14(1), 26-28.

Van der Merwe, F.J., Martin, J. (2002). Four Southern African horse breeds. Agriculture 32, 57-72.

Van Marle-Koster, E., Visser, C. (2018). Genetic improvement in South African livestock: Can genomics bridge the gap between the developed and the developing sectors? Frontiers in Genetics. Available: https://doi.org/10.3389/fgene.2018.00331 [accessed September 17, 2022].

Wallner B, Palmieri N, Vogl C, Rigler D, Bozlak E, Druml T, Jagannathan V, Leeb T, Fries R, Tetens J, Thaller G, Metzger J, Distl O, Lindgren G, Rubin CJ, Andersson L, Schaefer R, McCue M, Neuditschko M, Rieder S, Schlötterer C, Brem G. (2017). Y Chromosome Uncovers the Recent Oriental Origin of Modern Stallions. Current Biology 27(13), 2029-2035.

Wang, J. (2022). Fast and accurate population admixture inference from genotype data from a few microsatellites to millions of SNPs. Hereditary 129, 79-92.

Warmuth, V., Eriksson, A., Bower, M.A., Barker, G., Barrett, E., Hanks, B.K., Li, S., Lomitashvili, D., Ochir-Goryaeva, M., Sizonov, G.V., Soyonov, V., Manica, A. (2012). Reconstructing the origin and spread of horse domestication in the Eurasian steppe. PNAS 109(21), 8202-8206.

Williams, S.C.P. (2012). Whence the domestic horse? New genetic analysis reveals where horses were first domesticated. Science. Available: https://www.science.org/content/article/whence-domestic-horse [accessed August 28, 2022].

Whitelaw, S. (2020). The leading SA sires so far. Sporting Post. Available: https://www.sportingpost.co.za/breeding/sa-sires-this-season/ [accessed September 14, 2022].

Zhong, M., Lange, K., Papp, J.C., Fan, R. (2010). A powerful score test to detect positive selection in genome-wide scans. European Journal of Human Genetics 18, 1148-1159.

Chapter 3

Understanding the diverse roles of rural, indigenous and commercial horse breeds within South African Society

3.1 Abstract

Rural, indigenous and commercial horses hold a distinctive position in livestock management due to their multi-functional roles that have evolved throughout history and society. This study investigated the functional contributions of South African horse populations, emphasizing their economic, socio-economic, and cultural significance within South African communities. A total of 110 detailed surveys were conducted and used within this study and phenotypic data sheets and photographic identification of 282 horses belonging to six different breed populations in three different provinces of South Africa, were recorded in order to determine phenotypic characteristics, that may provide insight into the roles that each breed population contribute within the community. Cross-tabulation analyses revealed all four rural horse breed populations (Basotho-type pony (BSP), East London Rural (ELR), Eastern Cape Rural (ECR) and KwaZulu-Natal (KZN)) showed significance for utilization in transportation (p = 0.001) and rural racing (p = 0.001) functional roles within rural communities. The BSP population showed significance (p = 0.001) for a unique functional role of meat production within the Sotho and Tshwane cultural communities – reinforced by hierarchical cluster analysis (HCA) whereby BSP breed formed the furthest cluster away from the remaining dataset. Closest level of similarity within the HCA was determined between the two commercial Saddlebred (SAD) and Thoroughbred (THR), whereby the low-level cluster revealed an agglomeration coefficient of 0.000 – indicating no dissimilarity of functional roles performed between the two commercial breeds. CATPCA included 6 ordinal phenotypic traits attributed to functional role suitability that yielded five clusters, whereby 4 clusters belonging to the rural breed populations overlapped indicating commonality of variable phenotypic. The last cluster belonging to the two commercial breed populations showed majority portion of THR horses clustered within the over-lapped region, however some THR individuals distributed away from cluster region along with all individuals within SAD population, owing to unique breed-specific phenotypic traits. Ordinal regression analysis indicated no significance ($p > 0.050$) of the multi-functional roles on the direct socio-economic impact of rural horse breed populations, due to determined low-income households of rural farmers but evidence of their functional roles within communities offer support and improvement of their owner's lives. In conclusion, the multi-functional roles of horses remain pivotal in supporting of livelihoods, from rural farmlands to commercial sectors, the significance of these horses are propelled by their unique ability to adapt and fulfil the evolving needs of South African communities.

Keywords: South African horse breeds; Rural farming communities; Economic impact; Cultural influence; Phenotypic-trait association; Indirect socio-economic impact

3.2 Introduction

Horses have long been intertwined with human history, whereby their purpose developed and shifted over time from the domestication of wild horse populations that were the predecessors of the modern domesticated horse (Librado et al., 2021). Before close integration with human populations, horses existed as wild prey animals within their environments and developed mechanisms of physiological responses to initiate unpredictable maneuvers or rapid flight (Hole et al., 2023), due to their subjection to an assortment of predators that included being hunted by humans. A few studies refer to the archeological evidence of horse remains found within historical human settlements, whereby specialized spears were crafted to hunt and kill horse herds (Hutson et al., 2024) as well the use of horse bones to make weapons and tools (Hoffecker et al., 2018). This depicts the observation that these wild horse populations served as a food source to early human populations. However, overtime it was hypothesized that humans observed the beneficial capabilities of horses to be utilized in constructing civilization and therefore sort out to domesticate these athletic prey animals, whereby Van Dierendonck and Sprijt (2012) emphasize that this domestication process lessened the general rapid prey responses in order to trust their captor but also maintained their existing herd social hierarchy. Koski and Spannari (2024) explore intersubjectivity between horse and human co-existence, where subjection of horses to a human-controlled environment significantly dominated the day-today decision making of survival of wild horse herds, in order fulfil multiple roles within these communities – while still maintaining cognitive independence (Brubaker and Udell, 2016) alongside their willingness to perform these roles. These historical performance roles changed as human cultures and economic growth developed that included the replacement of oxen in ploughing farming lands in agricultural practices (Lenders, 2024), chariot-style warfare (Klecel and Martyniuk, 2021), long-distance transportation that accelerated produce trade and military territorial conquest (Kysely and Peske, 2022).

Horses were first introduced into South Africa, in 1652 by Dutch settlers, as invasion agents in mounted warfare and the utilization of their draught power to establish useable land for infrastructure (Swart, 2003). Swart (2010) iterated the importance of horses shaping not only the biophysical landscapes of the colonies but how horses influenced the political and economic authority of the invasion and retaliation of conflict resigning within the country – while still utilized within agriculture and low-cost transport of produce. This economic power enabled warriors of the Basotho tribe to engage within mounted warfare to protect their landlocked country from the European colonizers (Mofuoa, 2015), mounted on top of resilient Basotho ponies that had adapted to the harsh, rocky environment of the country (Millin, 2007) which propelled the maintained independence of the Basotho people. Thereafter, the Basotho pony become a part of the cultural and traditional identity of the Basotho people where owning these horses represented a sense of nobility and capability to rear and breed these innovative animals (Phafoli and Zulu, 2012). The practical utilization of mounted horse power within Lesotho and rural areas within South Africa bordering the country, are still evident today as indirectly or directly contributing to the socio-economic and economic status of under-developed farming areas (Wadham et al., 2022) whereby horses provide a mode of transportation for people, produce distribution and agricultural powerhouses. Horses are grouped as multi-purpose livestock, which includes not only their

working or practical abilities within rural communities but their sustainability towards food security – a common consumption practice by Basotho people and other rural tribes (Danso-Abbeam et al., 2024). Sebbane et al. (2023) disseminates the fact that the consumption of horse meat has always been a normalized practice in many European cultures as a source of high-protein alternative to livestock that have a severe impact on the Fenvironment.

The establishment of Thoroughbred horseracing throughout history has now lead to a multi-million rand industry, that has significantly impacted the economic growth of South Africa (Surujlal et al., 2011). The cultural and economic influence of these sporting events influenced rural communities such as Xhosa groups in Transkei, Eastern Cape, where these groups traditionally would gather to race cattle herds (referred to as "uleqo") (Paterson, 2021) – which was then replaced by horses due to their advantageous capability of speed and power than compared to cattle. This transition improved the economic and socio-economic growth of rural communities and has evolved into significant rural racing spectacles, such as the renowned July Dundee rural racing event in the KwaZulu-Natal Province (Mhlauli and Ezeuduji, 2022), which has been pivotal in establishing South Africa as a prominent destination for tourism. The economic, socio-economic and cultural roles horses play both in rural communities and commercial breeding and racing sectors within South Africa are extremely versatile (Skinner et al., 2022; Wadham et al., 2022; de Klerk et al., 2020), whether these breed populations influence these roles directly or indirectly – their impact is significant and necessary.

The phenotypic characterization of these locally adapted horse breeds, which are highly variable due to integration of traits typical of crossbred inviduals (Monau et al., 2018), provides a vital step in the facilitation of classifying individuals within a population. Consistent traits within the population, such as conformation and body shape (Brooks et al., 2010), supports the functional characterization of these populations and underscores their value within local communities. These findings were determined within this study, in particular relation to the functional roles associated with each population. Therefore, this genomics assisted, community-driven research study aims to characterize the functional diversity of six indigenous, rural and commercial horse breed populations present within South African society today. This was achieved through the objective to conduct detailed surveys with rural farmers, breed societies and commercial horseracing training yards in order to evaluate the economic, socio-economic and cultural roles that these South African horse populations contribute towards South African society. Additionally, the study seeks to bridge the gap between traditional knowledge systems and modern genomics by integrating the perspectives and experiences of local stakeholders. By highlighting the unique adaptations, functional traits, and historical significance of these horse populations, this research underscores their value not only as economic assets but also as key contributors to cultural heritage, rural livelihoods, and community identity.

3.3 Materials and Methods

3.3.1 Detailed Surveys and Phenotypic Horse Data

A total of 110 detailed surveys were conducted and used within this study, whereby these surveys included the following sections of i) General information of respondent, ii) General information of household and iii) Horse statistics and breeding populations. A stratified sampling design was employed, wherein one district was selected per province, and within each selected district, four rural villages were targeted. This approach ensured adequate geographic representation and coverage across each district. Phenotypic data sheets and photographic identification of 282 horses belonging to six different breed populations in three different provinces of South Africa, were recorded in order to determine phenotypic characteristics, that may provide insight into the roles that each breed population contribute within the community. The sample size per breed population was influenced by the genotyping capacity of the Equine70K and Equine80K BeadChip arrays, with each array accommodating 24 individual samples (Schaefer and McCue, 2020). This constraint informed the allocation of genotyped individuals across breeds for subsequent genomic analyses. GPS coordinates of sample locations within three Provinces of South Africa were visualized in **Figure 3.1** below.

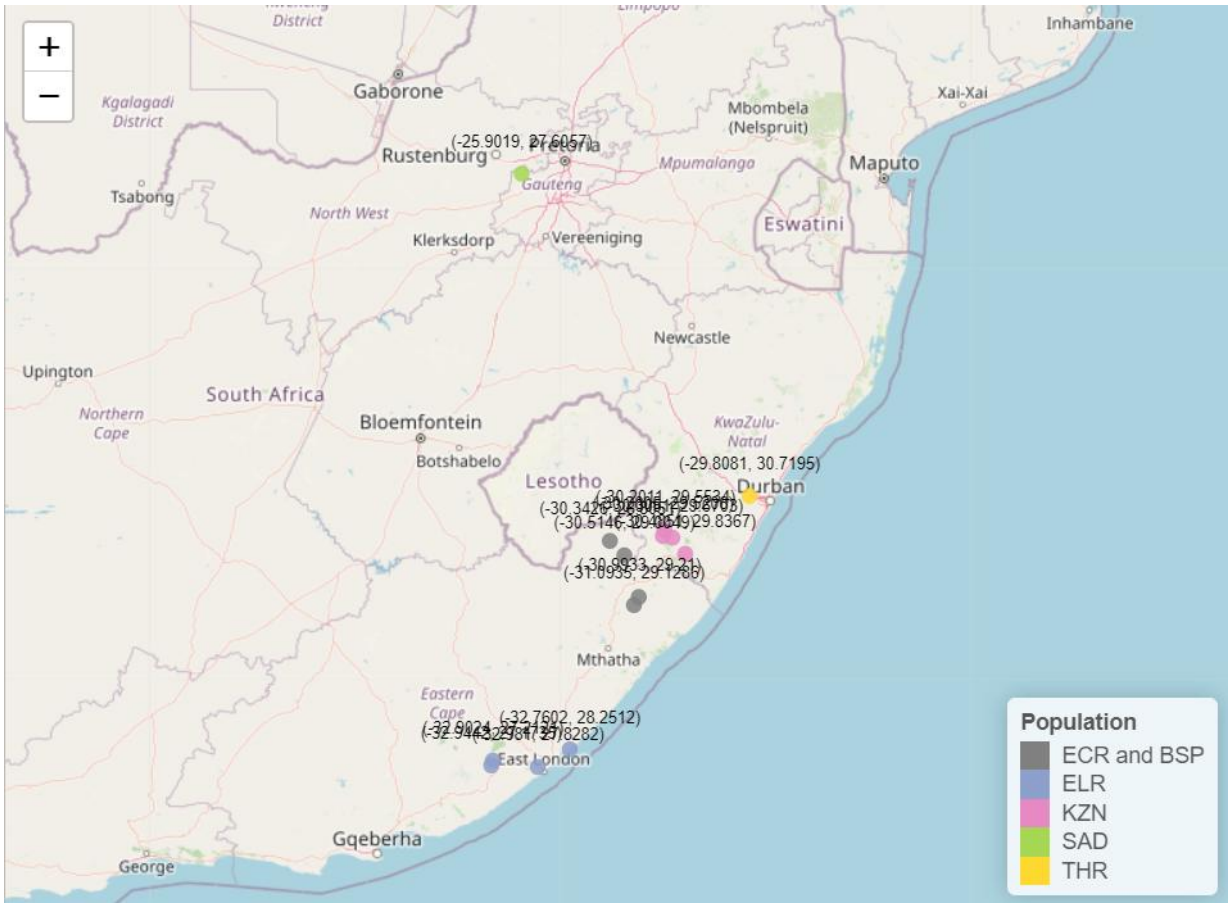


Figure 3.1 GIS map of sample regions within Eastern Cape, KwaZulu-Natal and Gauteng Provinces

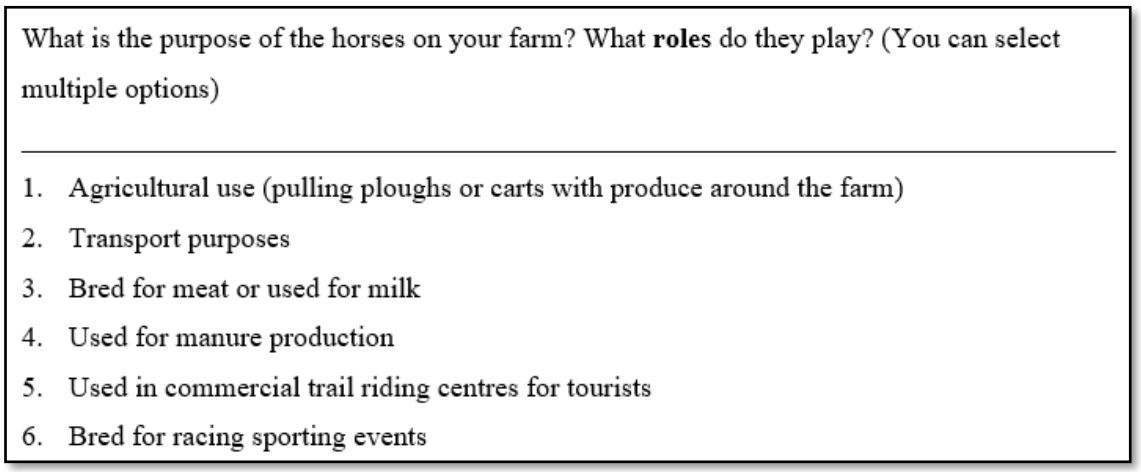
[GPS coordinates of rural hourseholds, training facility and breeding stud was indicated. This GIS map was constructed using R/Rstudio version 4.3.1 (R Core Team, 2020; Rstudio Team, 2020), applying the leaflet package (Agafonkin, 2016).]

Rural farmers were interviewed from Alfred Nzo (15 famers interviewed; n = 48 crossbred rural horses (ECR)) and Amathole (11 farmers interviewed; n = 25 crossbred rural horses (ELR)) districts in the Eastern Cape Province and from uMzinyathi (20 farmers interviewed; n = 47 crossbred horses (KZN)) district in the KwaZulu-Natal Province. Rural Sesotho farmers were interviewed in the district of Alfred Nzo (62 farmers interviewed; n = 95 indigenous Basotho-type ponies) within the Eastern Cape Province – which borders the country Lesotho where the breed was first established. A commercial horserace trainer from the Summerveld Training Centre (https://www.goldcircle.co.za/racing/training-centres/training-centres-summerveld) located in the eThekwini district of the KwaZulu-Natal Province was interviewed (1 trainer; n = 47 Thoroughbred racing horses). One commercial horse breeder was interviewed, according to the allocation of a breeding stud by the Saddlebred Society of South Africa (https://saddlebred.co.za; Swart, 2008), located in the West Rand district (1 commercial breeder; 20 Saddlebred horses) of the Gauteng Province.

Detailed surveys, phenotypic data sheets and photographic identification were conducted in accordance with the standard operating procedures approved by Onderstepoort Veterinary Research Institutional Biosafety and Biosecurity Committee (OVR-IBBC). Ethical approval for the study was obtained from the OVR Animal Ethics Committee (**AEC 24.16**).

3.3.2 Determination of Functional Roles between Communities

To evaluate the relationship between the categorical multi-functional assigned roles of each horse population located within the rural and commercial communities of five different districts, a cross-tabulation analysis was performed imploring the use of Pearson's Chi-squared measure of statistical significance from one another ($p \leq 0.05$) (Plackett, 1983; White and Korotayev, 2004), on IBM SPSS Statistics version 29.0.2.0 (IBM Corp., 2023). The functional roles assigned to horses for each household/stud farm/training facility recorded within the survey, posed six different options of roles in which multiple options could be chosen, as shown in **Figure 3.2** below.


What is the purpose of the horses on your farm? What **roles** do they play? (You can select multiple options)

1. Agricultural use (pulling ploughs or carts with produce around the farm)
2. Transport purposes
3. Bred for meat or used for milk
4. Used for manure production
5. Used in commercial trail riding centres for tourists
6. Bred for racing sporting events

Figure 3.2 Functional role options within detailed survey

[Multiple functional role options were indicated as acceptable within the household/stud farm/training facility.]

To assess the statistical significance and structure of the singular and multi-functional roles assigned to each horse breed population, Hierarchical Clustering Analysis (HCA) was employed. This approach was used to identify patterns of similarity in the roles performed by the six different horse breed populations, thereby uncovering potential functional groupings based on shared community-assigned roles (Kohn & Hubert, 2015). The analysis was conducted using IBM SPSS Statistics version 29.0.2.0 (IBM Corp., 2023). Prior to analysis, categorical role-related data from community surveys were transformed into binary (presence/absence) format to facilitate comparison across populations. Each breed population within a given community was represented by an aggregated binary profile indicating the frequency or presence of specific role types. Euclidean distance was selected as the similarity metric using the *dist()* function, which calculates pairwise distances between breed role profiles based on their binary attributes (Zhang et al., 2017). This distance matrix provided the foundation for hierarchical clustering. Clustering was performed using the between-groups linkage method (also known as average linkage between groups) implemented via the *hclust()* function. This method groups populations by minimizing the average distance between all pairs of observations in any two clusters, thus allowing for the identification of stratified groupings that reflect functional role similarity across horse populations (Sanaei et al., 2015). To evaluate the statistical robustness of identified clusters, bootstrapping was used to assess cluster stability (Fang and Wang, 2012). This metric was used to determine the reliability of the observed groupings and whether breed populations were significantly differentiated based on role assignments.

3.3.3 Phenotypic Analysis of Horse Functional Roles

Photographic identification images and categorical phenotypic traits of the six different breed populations were recorded, in order to determine the association of trait-specific phenotypes within the horse breed populations and their designated functional roles (Ribeiro et al., 2024). Trait-specific phenotypes included body height, body length, leg length, body muscle conformation, neck shape and head size with three categorical options for each trait as shown in **Table 3.1 below**.

Table 3.1 Phenotypic trait assessment for each horse breed population in relation to functional ability

[Six phenotypic traits were identified as contributing factors associated with functional roles that horses possess within different communities. Three distinct levels per variable were maintained throughout the trait observations in order to ensure distribution balance for statistical testing.]

Trait Description	Categorical Observations
1. Body Height	Short
	Medium
	Tall
2. Body Length	Short
	Medium
	Long
3. Leg length	Short
	Medium
	Long
4. Body Muscle Conformation	No muscle
	Semi-muscled
	Muscled
5. Neck Shape	Ewe
	Straight
	Arched
6. Head Size	Small
	Medium
	Large

These traits were recorded for each horse within each breed population and converted into a matrix dataset, in order to analyze patterns between the phenotypic categories of the multiple variables, through the use of a Categorical Principal Components Analysis (CATPCA) using both nominal population grouping and ordinal phenotypic variables (Atkinson, 2023), conducted within IBM SPSS Statistics version 29.0.2.0 (IBM Corp., 2023) and visualized using R/Rstudio version 4.3.1 (Rstudio Team, 2020). The dimensionality of the dataset was set to two dimensions to ensure major dimensions of variable variation attained.

3.3.4 Assessment of Socio-economic Impact in Relation to Functional Roles

The classification of monthly income (ZAR) per household/stud farm or training facility was ordinally categorized into five levels, each representing an income range as shown in **Figure 3.3** below.

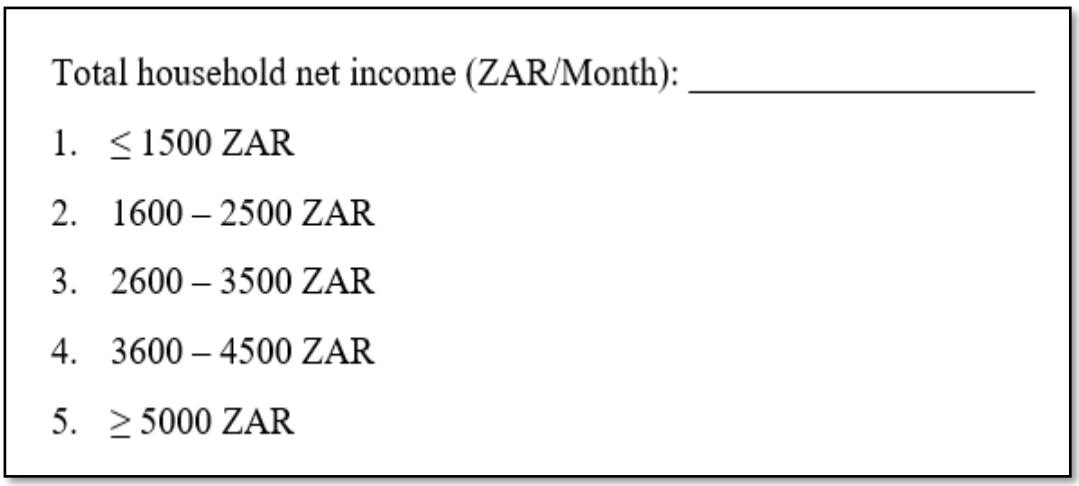


Figure 3.3 Monthly income per household/stud farm/training facility in relation to direct or indirect functional roles of horses belonging to each community

[Monthly income pertaining to direct or indirect association with the functional role of horses within each community, ranged from less than 1500 (ZAR) to over 5000 (ZAR).]

These ordinal categories were treated as the dependent variable that were tested against independent variables of multi-functional roles within each community as well as population grouping, in separate ordinal regression analyses using the Polytomous Universal Model (PLUM) procedure (DeCarlo, 2003) using IBM SPSS Statistics version 29.0.2.0 (IBM Corp., 2023). The logit link function was selected to link the ordinal response variable (monthly income), to the predictor variables (functional roles) for each breed population and expressed as odds-ratios (Smith et al., 2019). This statistical fit allowed insight into the determination of whether specific functional roles or multiple roles influenced the monthly income (ZAR) of each household/stud farm and training center involved within this study. The significance of predictors within the regression was evaluated using Wald statistics, setting significance threshold at $p \leq 0.05$ (Marshall et al., 2009).

3.4 Results

3.4.1 Determination of Functional Roles between Communities

Multi-functional roles of each horse breed population was cross-tabulated in order to determine the main functional roles representative of the different communities involved within this study, combined with associated significance test of the combined roles – shown within **Tables 3.2** to **3.6** below.

Table 3.2 Cross-comparison of a functional role in agriculture for each breed population

[**A.** Cross tabulation analysis for the comparison of functional role in agriculture within the six horse breed populations. The value of "1" represents horses that were used in agriculture and a value of "0" represents horses that are not used in agriculture. **B.** Pearson's chi-square and Likelihood ratio significance test for the use of horses in agriculture across populations.]

A

			Agriculture		
			0	1	Total
Population	BSP	Count	50	12	62
		% within Population	80.6%	19.4%	100.0%
	ECR	Count	13	2	15
		% within Population	86.7%	13.3%	100.0%
	ELR	Count	11	0	11
		% within Population	100.0%	0.0%	100.0%
	KZN	Count	11	9	20
		% within Population	55.0%	45.0%	100.0%
	SAD	Count	1	0	1
		% within Population	100.0%	0.0%	100.0%
	THR	Count	1	0	1
		% within Population	100.0%	0.0%	100.0%
Total		Count	87	23	110
		% within Population	79.1%	20.9%	100.0%

B

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	11.067[a]	5	0.050
Likelihood Ratio	12.575	5	0.028
N of Valid Cases	110		

a. 7 cells (58,3%) have expected count less than 5. The minimum expected count is ,21.

The cross-tabulation analysis indicated that only three breed populations within those regional communities that utilized horses for agricultural purposes. This includes the Basotho-type (19.400%), Eastern Cape Rural (13.300%) and KwaZulu-Natal rural (45%) breed populations whereby the most significant occurring population was that of the KwaZulu-Natal breed population. The remaining three populations indicated no use of horses for agricultural purposes. The significance test resulted in p-values of $p \leq 0.05$, indicating significance of the use of horses for agricultural purposes across the populations, as three out of the six breed populations were used for agriculture whereby this functional role only had the highest prevalence in the KZN community.

Table 3.3 Cross-comparison of a functional role in transport for each breed population

[**A.** Cross tabulation analysis for the comparison of functional role in transport within the six horse breed populations. The value of "1" represents horses that were used in transport and a value of "0" represents horses that are not used in transport. **B.** Pearson's chi-square and Likelihood ratio significance test for the use of horses in transport across populations.]

A

			Transport		
			0	1	Total
Population	BSP	Count	2	60	62
		% within Population	3.2%	96.8%	100.0%
	ECR	Count	0	15	15
		% within Population	0.0%	100.0%	100.0%
	ELR	Count	6	5	11
		% within Population	54.5%	45.5%	100.0%
	KZN	Count	0	20	20
		% within Population	0.0%	100.0%	100.0%
	SAD	Count	1	0	1
		% within Population	100.0%	0.0%	100.0%
	THR	Count	1	0	1
		% within Population	100.0%	0.0%	100.0%
Total		Count	10	100	110
		% within Population	9.1%	90.9%	100.0%

B

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	53.581[a]	5	<.001
Likelihood Ratio	34.191	5	<.001
N of Valid Cases	110		

a. 7 cells (58,3%) have expected count less than 5. The minimum expected count is ,09.

The use of horses for transport showed great significance ($p < 0.001$) across the six horse breed populations. The four horse populations belonging to rural communities showed significance in the functional role of transport, whereby the highest occurrence belonged to the Basotho-type (100%) and KwaZulu-Natal rural breed populations – indicating that all horses within that sampled population were used for transportation purposes among other functional roles. Both the commercial Thoroughbred and Saddlebred breed populations did not show any significance for the use of transport.

Table 3.4 Cross-comparison of a functional role in meat production for each breed population

[**A.** Cross tabulation analysis for the comparison of functional role in meat production within the six horse breed populations. The value of "1" represents horses that were used for meat production and a value of "0" represents horses that are not used in meat production. **B.** Pearson's chi-square and Likelihood ratio significance test for the use of horses in meat production across populations.]

A

			Meat		
			0	1	Total
Population	BSP	Count	7	55	62
		% within Population	11.3%	88.7%	100.0%
	ECR	Count	15	0	15
		% within Population	100.0%	0.0%	100.0%
	ELR	Count	11	0	11
		% within Population	100.0%	0.0%	100.0%
	KZN	Count	20	0	20
		% within Population	100.0%	0.0%	100.0%
	SAD	Count	1	0	1
		% within Population	100.0%	0.0%	100.0%
	THR	Count	1	0	1
		% within Population	100.0%	0.0%	100.0%
Total		Count	55	55	110
		% within Population	50.0%	50.0%	100.0%

B

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	85.161[a]	5	<.001
Likelihood Ratio	108.777	5	<.001
N of Valid Cases	110		

a. 4 cells (33,3%) have expected count less than 5. The minimum expected count is ,50.

The cross tabulation analysis detected only one breed population, the Basotho-type (88.700%) breed that showed a functional role of utilization for meat production and no evidence of milk production, within the Eastern Cape rural community. This was associated with great statistical significance of $p \leq 0.050$ ($p = 0.001$). There were no functional roles of horse meat production detected within the remaining horse breed populations.

Table 3.5 Cross-comparison of a functional role in rural racing for each breed population

[**A.** Cross tabulation analysis for the comparison of functional role in rural racing within the six horse breed populations. The value of "1" represents horses that were used in rural racing and a value of "0" represents horses that are not used in rural racing. **B.** Pearson's chi-square and Likelihood ratio significance test for the use of horses in rural racing across populations.]

A

			Rural_Racing		
			0	1	Total
Population	BSP	Count	55	7	62
		% within Population	88.7%	11.3%	100.0%
	ECR	Count	13	2	15
		% within Population	86.7%	13.3%	100.0%
	ELR	Count	2	9	11
		% within Population	18.2%	81.8%	100.0%
	KZN	Count	8	12	20
		% within Population	40.0%	60.0%	100.0%
	SAD	Count	1	0	1
		% within Population	100.0%	0.0%	100.0%
	THR	Count	1	0	1
		% within Population	100.0%	0.0%	100.0%
Total		Count	80	30	110
		% within Population	72.7%	27.3%	100.0%

B

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	37.504[a]	5	<.001
Likelihood Ratio	36.063	5	<.001
N of Valid Cases	110		

a. 6 cells (50,0%) have expected count less than 5. The minimum expected count is ,27.

Great statistical significance was determined ($p \leq 0.050$) across the six horse breed populations as all of the rural breed populations representing specific regional rural communities utilized horses for rural racing purposes. The largest occurrence of functional role association of rural racing was found within the East London Rural breed population (81.800%). The two commercial Thoroughbred and Saddlebred populations showed no occurrence of rural racing functional association.

Table 3.6 Cross-comparison of a functional role in commercial showing and racing for each breed population

[**A.** Cross tabulation analysis for the comparison of functional role in commercial racing and showing within the six horse breed populations. The value of "1" represents horses that were used in commercial racing and showing and a value of "0" represents horses that are not used in commercial racing and showing. **B.** Pearson's chi-square and Likelihood ratio significance test for the use of horses in commercial racing and showing across populations.]

A

			Racing_Showing		
			0	1	Total
Population	BSP	Count	62	0	62
		% within Population	100.0%	0.0%	100.0%
	ECR	Count	15	0	15
		% within Population	100.0%	0.0%	100.0%
	ELR	Count	11	0	11
		% within Population	100.0%	0.0%	100.0%
	KZN	Count	20	0	20
		% within Population	100.0%	0.0%	100.0%
	SAD	Count	0	1	1
		% within Population	0.0%	100.0%	100.0%
	THR	Count	0	1	1
		% within Population	0.0%	100.0%	100.0%
Total		Count	108	2	110
		% within Population	98.2%	1.8%	100.0%

B

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	110.000[a]	5	<.001
Likelihood Ratio	19.993	5	,001
N of Valid Cases	110		

a. 8 cells (66,7%) have expected count less than 5. The minimum expected count is ,02.

The cross-tabulation analysis revealed great significance ($p \leq 0.050$) for the determination of functional role association of commercial racing and showing in the Thoroughbred (100 %) and the Saddlebred (100 %) breed populations. There was no detection for the functional role association within the four rural community representative populations.

The observed statistical significance of multiple functional roles of horses within the representative communities were further analyzed to determine the commonality of these functional roles between the breed populations. The hierarchical cluster analysis revealed the relationship of common functional roles between the breed populations in **Figure 3.4**, **Figure 3.5** and **Table 3.7** below.

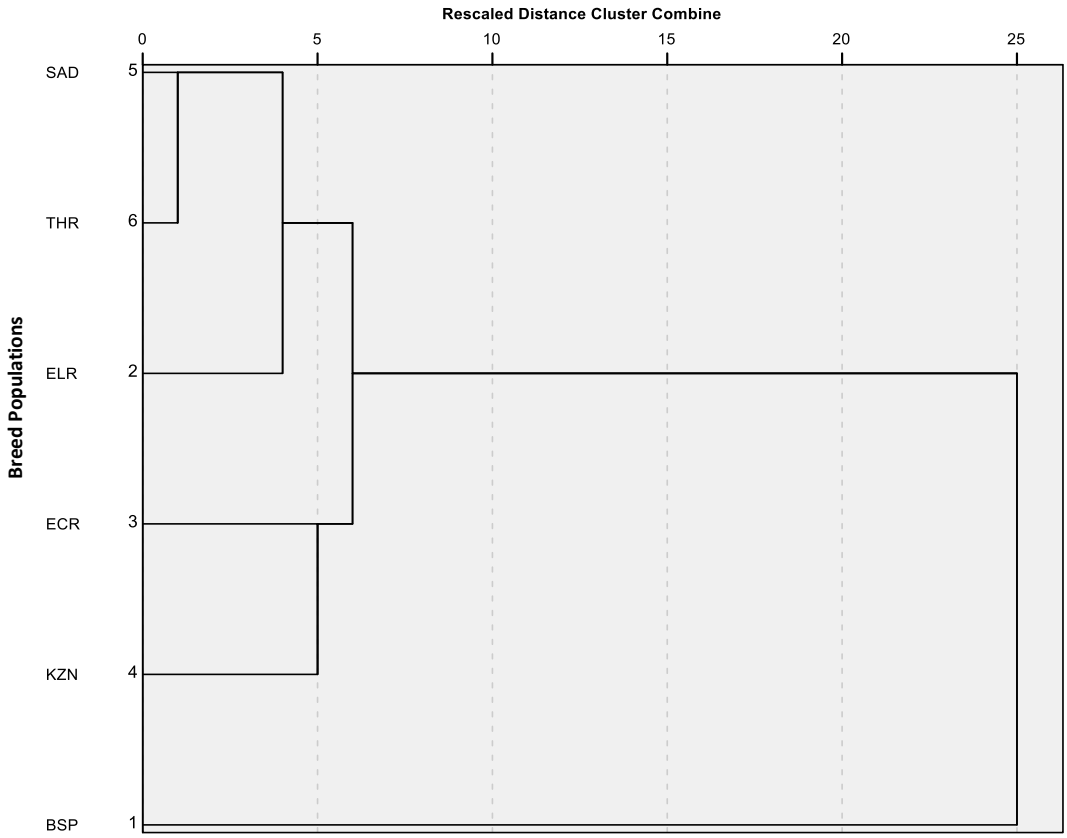


Figure 3.4 Population clustering of functional roles

[Hierarchical cluster tree was based off between-group linkage to compare multi-functional roles of the six horse populations belonging to representative communities. Rescaled Euclidean distance was used to structure the linkages between the breed populations.]

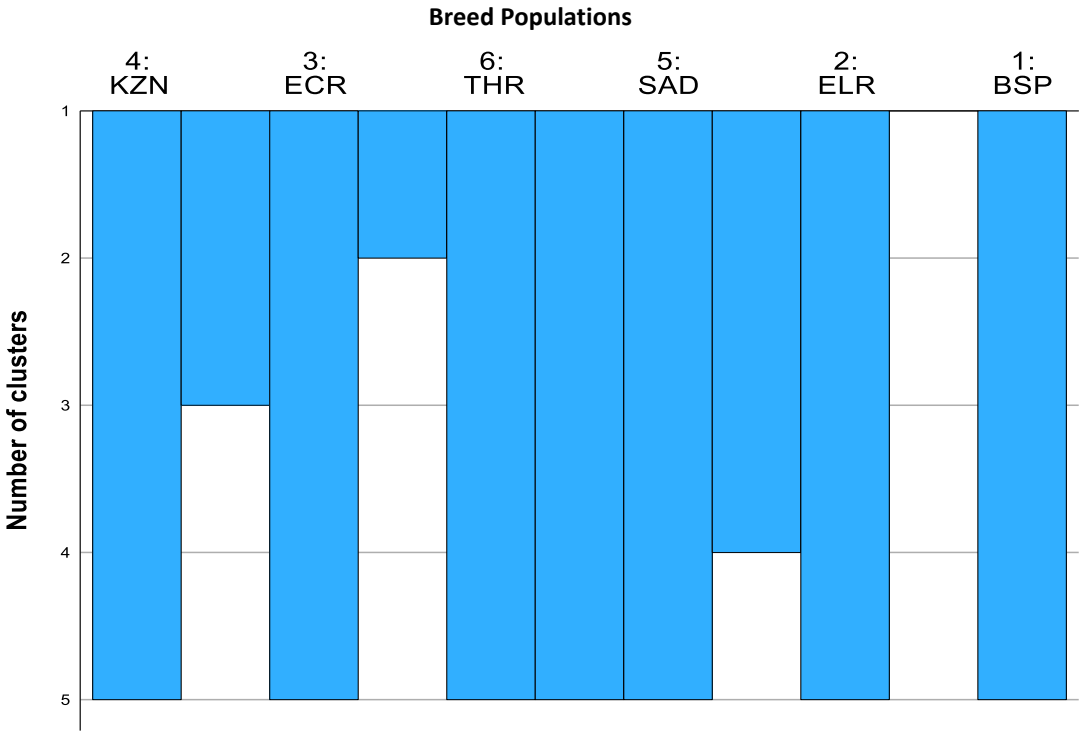


Figure 3.5 Agglomeration of clusters based on hierarchical clustering method

[The number of clusters in which each breed population was grouped was based on the agglomeration coefficient of *hclust* () function.]

Table 3.7 Agglomeration schedule of hierarchical clustering of the six breed populations

[The table below demonstrates the Stage (list of the sequence of cluster merging), cluster 1 and Cluster 2 (shows the clusters that are merged at each step), Coefficients (agglomeration coefficient that measures the distance at which clusters are joined) and Stage Cluster First Appears (tracks progress from individual data points to a single cluster).]

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	5	6	0.000	0	0	2
2	2	5	10.583	0	1	4
3	3	4	13.342	0	0	4
4	2	3	18.523	2	3	5
5	1	2	80.258	0	4	0

The closest levels of similarity was determined at the first sub-cluster in **Figure 3.4**, between the Thoroughbred (THR – 6 (**Figure 3.5**)) and Saddlebred (SAD – 5 (**Figure 3.5**)) breed populations, which indicates high similarity of the functional role assigned to the two commercial breeds of commercial racing and showing. The agglomeration coefficient of 0.000 (**Table 3.7**) suggests that there is no dissimilarity between the two groups as these populations were exclusively associated with a single functional role within their community. The East London rural (ELR – 2) breed population joins the cluster formed by SAD and THR, suggesting a broader similarity in functional traits between the three populations. However, an agglomeration coefficient of 10.583 indicates complete dissimilarity between the ELR group and the THR and SAD group cluster, as the populations do not share any form of functional roles within their separate communities. The Eastern Cape Rural (ECR – 3) and KwaZulu-Natal rural (KZN – 4) formed a lower-level cluster but the agglomeration coefficient of 13.342 suggests a higher level of dissimilarity between the two breed populations. This suggests that even though these two populations share common associated functional roles of use in agricultural practices, transport and rural racing – the number of horses utilized for each functional role was observed to be different in **Tables 3.2**, **3.3** and **3.5** above. The highest level of dissimilarity, furthest to the right in **Figure 3.4**, the Basotho-type (BSP – 1) breed population merges with all the other groups as this breed population contains the most distinct functional role compared to the other populations. This observation was suggested by the agglomeration coefficient of 80.258 in **Table 3.7** as the functional role of horse meat production associated within the representative community these horses form a part of.

3.4.2 Phenotypic Analysis of Horse Functional Roles

Photographic identification images of each horse from the six different breed populations were captured in order to associate trait-specific phenotypes associated with functional roles that were shared amongst the populations. Phenotypes were classified categorically based on the photographic identification images shown in **Figure 3.6** below.



Figure 3.6 Photographic identification demonstrating categorical trait-specific phenotypes for each horse breed population

[Trait-specific phenotypes for body height, body length, leg length, body muscle formation, neck shape and head size were categorically recorded and taken into consideration for all six horse breed populations. The above figure demonstrates where on the animal the phenotypes were determined for the **A.** Basotho-type breed; **B.** East London Rural breed; **C.** Eastern Cape Rural breed; **D.** KwaZulu-

Natal rural breed; **E.** Saddlebred breed and **F.** Thoroughbred breed.] [Sourced from the Agricultural Research Council, Onderstepoort Veterinary Research Institute (ARC-OVR).]

All three categories per trait-specific phenotypes that were hypothesized to provide insight into functional role association were analyzed in order to determine similarities between the breed populations, by structuring the phenotypes into principal components in **Figure 3.7** below.

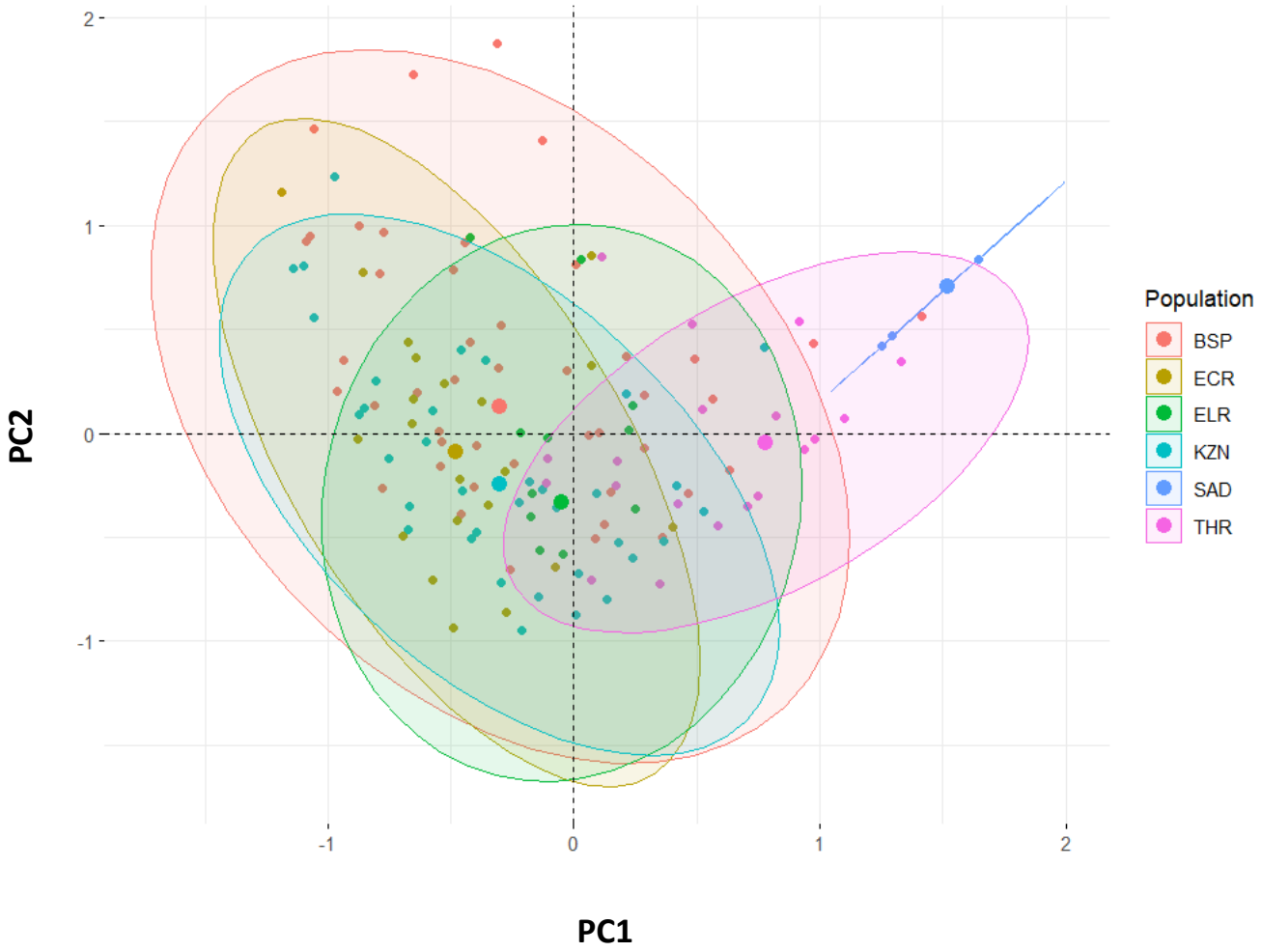


Figure 3.7 Categorical Principal Components Analysis (CATPCA) exhibiting population clustering based on phenotypic traits

[Five distinct clusters were identified within the CATPCA. The five clusters belonging to the **BSP** population (red cluster), **ECR** population (mustard cluster), **KZN** population (light blue cluster), **ELR** population (green cluster) and lastly the **SAD** and **THR** population (purple cluster) overlapped with one another.]

The above clustering (**Figure 3.7**) of nominal population grouping and ordinal phenotypic characteristics showed a total variation of 1.300% (PC1) and 0.900% (PC2). The five distinct clusters within the dimension matrix were distributed overlapping one another, with four clusters belonging to BSP (red cluster), ECR (mustard cluster), KZN (light blue cluster) and ELR (green cluster) containing individuals that were concentrated towards the middle of the dimension matrix with individuals distributed towards the upper-left quadrant. This indicates that individuals belonging to the four horse breed populations shared common trait-specific phenotypes where some individuals possessed variation within the six selected phenotypes and distributed away from the concentrated middle portion of individuals. The last cluster belonging to the THR

and SAD population contained THR individuals that overlapped with the middle sector congregation of individuals, indicating that some THR individuals possessed phenotypic traits in common with the rural and indigenous BSP, ECR, ELR and KZN individuals. This indicates the THR precursor breed influence within the three rural breed and the indigenous BSP populations, whereby similarities in a proportion of the phenotypic traits were observed. However, some THR individuals grouped towards the upper-right quadrant within the dimensional matrix indicating diversification of phenotypic traits from the middle over-lapping grouping. The SAD individuals (darker blue) grouped together within the purple cluster, in a straight line, indicating uniformity of phenotypic traits shared amongst the commercial SAD individuals. The SAD population were identified as containing the most distinctive set of phenotypic traits compared to the remaining breed clusters.

3.4.3 Assessment of Socio-economic Impact in Relation to Functional Roles

In order to assess the level of socio-economic impact of the horse breed populations found within different communities, the multi-functional roles designated to each population were taken into consideration that either indirectly or directly contributed towards the monthly income (ZAR/month) of the household/ stud farm or training center. Results shown from the analysis of this relation was conducted using an ordinal regression as shown in **Tables 3.8** to **3.11** below

Table 3.8 Comparison of household monthly income (ZAR/Month) and functional roles of Basotho-type (BSP) breed population

[The ordinal regression contained parameter estimates of Estimates (**Estimate**), Standard error (**Std. Error**), Wald's statistic (**Wald**), Degrees of freedom (**df**) and Wald significance (**Sig.**) from the regression model of dependent monthly income (ZAR/month) variable and the factorial functional roles comparison. The value of "1" represents horses that were used for the stipulated functional role and a value of "0" represents horses that are not used for the stipulated functional role. Statistical estimates that did not contain any values were functional roles that were either not associated with the specific breed population or did not cause any type of significant effect within the ordinal regression. Monthly income (ZAR) ordinals included "**1**" = ≤ 1500, "**2**" = 1600-2500, "**3**" = 2600-3500, "**4**" = 3600-4500 and "**5**" = ≥ 5000.]

		Estimate	Std. Error	Wald	df	Sig.
Threshold	[Monthy_Income = 1]	-2.137	0.656	10.608	1	0.001
	[Monthy_Income = 2]	-.739	0.552	1.796	1	0.180
	[Monthy_Income = 3]	0.949	0.557	2.905	1	0.088
	[Monthy_Income = 4]	2.723	0.682	15.924	1	<.001
Location	[Agriculture=0]	0.252	0.590	0.182	1	0.669
	[Agriculture=1]	0[a]	.	.	0	.
	[Transport=0]	-2.201	1.424	2.390	1	0.122
	[Transport=1]	0[a]	.	.	0	.
	[Meat=0]	3.089	1.014	9.281	1	0.002
	[Meat=1]	0[a]	.	.	0	.

	[Manure_Production=0]	0[a]	.	.	0	.
	[Trail_Riding=0]	0[a]	.	.	0	.
	[Rural_Racing=0]	0[a]	.	.	0	.
	[Rural_Racing=1]	0[a]	.	.	0	.
	[Racing_Showing=0]	0[a]	.	.	0	.

In **Table 3.8** above, the ordinal regression outlay the significantly highest and lowest monthly income of the Basotho-type (BSP) population community, which included the lowest income bracket of ≤ 1500 ($p < 0.001$) and the highest income bracket of 3600 – 4500 ($p < 0.001$). The effect of functional roles on the monthly household income within this community did not produce any significance, excepting the farmers that did not associate their horses with the functional role of meat production [Meat=0] which posed significance ($p < 0.002$) as these farmers were categorized in the higher earning income bracket.

Table 3.9 Comparison of household monthly income (ZAR/Month) and functional roles of East London Rural (ELR) breed population

[The ordinal regression contained parameter estimates of Estimates (**Estimate**), Standard error (**Std. Error**), Wald's statistic (**Wald**), Degrees of freedom (**df**) and Wald significance (**Sig.**) from the regression model of dependent monthly income (ZAR/month) variable and the factorial functional roles comparison. The value of "1" represents horses that were used for the stipulated functional role and a value of "0" represents horses that are not used for the stipulated functional role. . Statistical estimates that did not contain any values were functional roles that were either not associated with the specific breed population or did not cause any type of significant effect within the ordinal regression. Monthly income (ZAR) ordinals included "**1**" = ≤ 1500, "**2**" = 1600-2500, "**3**" = 2600-3500, "**4**" = 3600-4500 and "**5**" = ≥ 5000.]

		Estimate	Std. Error	Wald	Df	Sig.
Threshold	[Monthy_Income = 0]	0.251	1.141	0.048	1	0.826
	[Monthy_Income = 1]	2.197	1.448	2.303	1	0.129
	[Monthy_Income = 2]	3.664	1.717	4.553	1	0.033
Location	[Agriculture=0]	0[a]	.	.	0	.
	[Transport=0]	2.197	1.573	1.950	1	0.163
	[Transport=1]	0[a]	.	.	0	.
	[Meat=0]	0[a]	.	.	0	.
	[Manure_Production=0]	8.153E-9	0	.	1	.
	[Manure_Production=1]	0[a]	.	.	0	.
	[Trail_Riding=0]	0[a]	.	.	0	.
	[Rural_Racing=0]	21.428	4158.575	0	1	0.996
	[Rural_Racing=1]	0[a]	.	.	0	.
	[Racing_Showing=0]	0[a]	.	.	0	.

In **Table 3.9**, the highest household monthly income (ZAR/month) that was observed as statistically significant ($p < 0.033$) fell into the income bracket of 1600-2500 [Monthly_Income = 2]. Out of the multi-functional

categorical options within the analysis, none of the functional roles outlined within the table above had any statistical significance on the monthly income within each household.

Table 3.10 Comparison of household monthly income (ZAR/Month) and functional roles of KwaZulu-Natal (KZN) rural breed population

[The ordinal regression contained parameter estimates of Estimates (**Estimate**), Standard error (**Std. Error**), Wald's statistic (**Wald**), Degrees of freedom (**df**) and Wald significance (**Sig.**) from the regression model of dependent monthly income (ZAR/month) variable and the factorial functional roles comparison. The value of "1" represents horses that were used for the stipulated functional role and a value of "0" represents horses that are not used for the stipulated functional role. . Statistical estimates that did not contain any values were functional roles that were either not associated with the specific breed population or did not cause any type of significant effect within the ordinal regression. Monthly income (ZAR) ordinals included "**1**" = ≤ 1500, "**2**" = 1600-2500, "**3**" = 2600-3500, "**4**" = 3600-4500 and "**5**" = ≥ 5000.]

		Estimate	Std. Error	Wald	df	Sig.
Threshold	[Monthy_Income = 1]	-0.402	1.907	0.044	1	0.833
	[Monthy_Income = 3]	-0.024	1.900	0.000	1	0.990
	[Monthy_Income = 4]	0.605	1.907	0.101	1	0.751
Location	[Agriculture=0]	1.104	1.257	0.771	1	0.380
	[Agriculture=1]	0[a]	.	.	0	.
	[Transport=1]	0[a]	.	.	0	.
	[Meat=0]	0[a]	.	.	0	.
	[Manure_Production=0]	0[a]	.	.	0	.
	[Trail_Riding=0]	0[a]	.	.	0	.
	[Rural_Racing=0]	0.198	1.559	0.016	1	0.899
	[Rural_Racing=1]	0[a]	.	.	0	.
	[Racing_Showing=0]	0[a]	.	.	0	.

The ordinal regression in **Table 3.10** above, did not produce any significance ($p > 0.050$) regarding the influence of the multi-functional roles assigned to the KZN rural horse population on the monthly income (ZAR/month) within each household that belonged to this community.

Table 3.11 Comparison of household monthly income (ZAR/Month) and functional roles of Eastern Cape Rural (ECR) breed population

[The ordinal regression contained parameter estimates of Estimates (**Estimate**), Standard error (**Std. Error**), Wald's statistic (**Wald**), Degrees of freedom (**df**) and Wald significance (**Sig.**) from the regression model of dependent monthly income (ZAR/month) variable and the factorial functional roles comparison. The value of "1" represents horses that were used for the stipulated functional role and a value of "0" represents horses that are not used for the stipulated functional role. Statistical estimates that did not contain any values were functional roles that were either not associated with the specific breed population or did not cause any type of significant effect within the ordinal regression. Monthly income (ZAR) ordinals included "**1**" = ≤ 1500, "**2**" = 1600-2500, "**3**" = 2600-3500, "**4**" = 3600-4500 and "**5**" = ≥ 5000.]

		Estimate	Std. Error	Wald	df	Sig.
Threshold	[Monthy_Income = 1]	-5.323	1.946	7.481	1	0.006
	[Monthy_Income = 2]	-1.648	1.541	1.143	1	0.285
	[Monthy_Income = 3]	-0.213	1.482	0.021	1	0.886
	[Monthy_Income = 4]	0.312	1.511	0.043	1	0.837
Location	[Agriculture=0]	-0.109	0.927	0.014	1	0.906
	[Agriculture=1]	0[a]	.	.	0	.
	[Transport=1]	0[a]	.	.	0	.
	[Meat=0]	0[a]	.	.	0	.
	[Manure_Production=0]	0[a]	.	.	0	.
	[Trail_Riding=0]	0[a]	.	.	0	.
	[Rural_Racing=0]	-0.753	0.996	0.572	1	0.450
	[Rural_Racing=1]	0[a]	.	.	0	.
	[Racing_Showing=0]	0[a]	.	.	0	.

In **Table 3.11** above, the ordinal regression outlined the significantly lowest ($p < 0.006$) household income (ZAR/month) within this community within the income bracket of ≤ 1500 [Monthly_Income = 1]. Regarding the influence of multi-functional roles assigned to the Eastern Cape Rural (ECR) breed population within this community, there was no statistically significant effect on the monthly income (ZAR/month) of each household.

The two commercial Saddlebred (SAD) and Thoroughbred (THR) breed populations belonging to a stud farm and training facility, respectively, did not fit the ordinal regression model as these two commercial breeds did not exhibit variation in terms of their functional roles which directly fell into the highest income bracket of ≥ 5000 [Monthly_Income = 5].

3.5 Discussion

Horses are distinct livestock animals as throughout history have been assigned multi-functional roles within different communities (Skinner et al., 2022), that have evolved throughout the ever-changing modernization of society and the direct or indirect impact of horses along with it. A unique functional role of horses also includes the utilization in equine-assisted therapy (EAT) treatment method in various mental, neurological and physical conditions (Letsoalo et al., 2024). These multi-faceted functional roles of horses maintain the economic, socio-economic and cultural impact that horses contribute towards South African Society. Within the rural farming areas of South Africa, horses are primarily utilized in rural farming practices as well as a means of transport – as statistically validated within the cross-tabulation analysis conducted. Three rural horse breed populations (**Table 3.2**), namely Basotho-type (BSP), Eastern Cape Rural (ECR) and KwaZulu-Natal (KZN) rural populations were found to be statistically significant for the use of these horses in agricultural practices with the KZN representative community showing the highest significance of the population (45 % of the population) ($p \leq 0.050$) to perform this functional role. Lewu and Assefa (2009) investigates the concentration of rural farmers within KwaZulu-Natal that contributes to the economic growth of the rural farming sector, that focuses on food production in agricultural practices of cropping systems, that lack the

input of machinery as a challenge for these small-scale farmers. However, the implementation of machinery within these small-scale farms is costly and the utilization of horse power is a cheaper alternative (Rydberg and Jansen, 2002), due to the comparison of low-input into the maintenance of adapted cross-bred rural horse breeds and farming equipment associated with horse-drawn agricultural tools. Another functional role in which rural horse breeds were utilized is in transportation, varying between transportation of individuals in place of motorized vehicles or the transportation of goods. All four rural horse breed populations, Basotho-type (BSP), Eastern Cape Rural (ECR), KwaZulu-Natal (KZN) rural and East London Rural (ELR) were found statistically significant ($p < 0.001$) in the utilization for transport means – shown within **Table 3.3**. Every individual horse belonging to the Basotho-type (BSP) breed population (100 %) were utilized for transportation means within the Alfred Nzo district in Eastern Cape - a rocky, mountainous region that does not allow motorized vehicles to travel across farmlands and therefore the implementation of horses by local famers for transportation has resolved these challenges (de Klerk et al., 2020).

Another novel functional role that this Basotho-type (BSP) horse breed were associated with was the role of meat production (**Table 3.4**), whereby 88.700% of the population had been statistically evident ($p < 0.001$) in the farmers consuming their horses for meat – but no evidence of their horses utilized for milk production. The consumption of horse meat within the Sotho (Basotho) and Tswana cultural groups has been a long-standing tradition, where their horses are not necessarily bred for meat production but instead are eaten when a horse has reached old age or is injured. These horses cannot fulfill their working role anymore – therefore not wasting food source where meat may be scarce (Soal, 1999). This unique role was highly distinct compared to the other five breed populations, as observed within the hierarchical cluster analysis (**Figure 3.4**), that determined the commonality of functional roles between the different populations within their representative communities. This population clustered the furthest away from the dataset and had an associated agglomeration coefficient of 80.258 (**Table 3.7**) – highlighting this unique, economic and cultural role. Rural racing was deemed as a significant functional role ($p < 0.001$) within the four rural horse populations (**Table 3.5**), whereby 81.800% of the East London Rural (ELR) breed population was utilized for this functional role. The recreational use of athletically-produced rural horses within Eastern Cape has gained recognition for the economic and socio-economic boost that this traditional sport has contributed towards rural communities, especially towards women-owning Small, Medium and Micro Enterprises (SMMEs) that are utilized in prominent rural races. Such races include the Annual Berlin November Horse Race that celebrates the rural tradition within the Xhosa community (Ngumbela, 2023). The hierarchical cluster analysis (**Figure 3.4**) formed a low-level cluster between the Eastern Cape Rural (ECR) and KwaZulu-Natal (KZN) rural breed populations, connecting their shared functional roles of agriculture (**Table 3.2**), transport (**Table 3.3**) and rural racing (**Table 3.5**). However, it was taken into consideration that a larger associated agglomeration coefficient of 13.342 (**Table 3.7**) that indicates a less well-formed cluster (Gere, 2023) – which may be due to the difference in the number of horses in each population.

Both the commercial Saddlebred (SAD) and Thoroughbred (THR) breed populations showed statistical significance ($p < 0.001$) in the utilization of racing and showing within the cross-tabulation analysis (**Table**

3.6) – commercial equine sports that are prominent within the South African equine sporting industry (Bailey et al., 2022; Swart, 2008). This observation was supported by the hierarchical cluster analysis (**Figure 3.4**) whereby these two commercial breed populations formed the closest level of similarity of the first sub-cluster, with an agglomeration coefficient of 0.000. The influence of athletic sporting breeds has greatly boosted the economic and socio-economic growth of South Africa, not only through competing within the high-earning sporting events but propelled industries that profit off of these commercial breeds. Suuch industries include the horse breeding industry (Scholtz et al., 2022), private equine veterinary practices (Marlow, 2010) and high-protein animal feed industry (Pienaar et al., 2021) – among many others.

Trait-specific phenotype association (**Figure 3.6**) with functional roles of each horse breed were assessed based on an ordinal category, drawing conclusions of similarities and distinctiveness between the populations, in order to emphasize the performance to assigned multi-functional roles (Pandey et al., 2011) within each community. Four clusters within the CATPCA (**Figure 3.7**) belonging to the rural breed populations overlapped, even though individuals within each rural breed are highly versatile in various phenotypic traits but they have retained their short stature (body height, body length and leg length) inherited from their common ancestor – the Cape horse (Swart, 2003). This short and stocky stature of crossbred rural horses enhances draught power for utilization in agricultural practices (Zantsi and Christian, 2023) and for transportation across rocky, mountainous farmlands (de Klerk et al., 2020) whereby their robust, short legs are able handle the impact that the uneven terrain would have on their joints. The musculature of these four rural horse breeds varied throughout the populations, whereby most horses contained defined body muscle conformation due to regular high-intensity exercise (Bryan et al., 2017) associated with working functional roles designated to the horses, but some horses possessed under-developed muscle formation. Food resources are limited within rural farming communities' due harsh environmental conditions within South Africa, that may affect quality of grazing (Getchell et al., 2002) and often farmers cannot afford to feed their horses high-protein horse feed, therefore causing lack of nutrition within these rural populations leading to decreased musculature. Another underdeveloped phenotypic trait that was prominent within the rural horse populations, was the neck shape (**Figure 3.6 A**; **B**; **C**; **D**) that was shared between the four clusters within the CATPCA (**Figure 3.7**). This neck shape, described as an ewe shape, is caused by muscle atrophy and creates a concave shape within the neck that can affect the movement of horses and caused strained back muscles (Ripolles et al., 2020). Poor nutrition has detrimental effects on skeletal muscle development within horses that can lead to certain muscular diseases (Katz et al., 2009), which poses a strong link to the occurrence of ewe-neck shapes within horses that has been documented to have a moderate trait heritability potential (0.230 – 0.340) (Ripolles et al., 2020).

The cluster containing most of the Thoroughbred horses (**Figure 3.7**) also overlapped with four rural horse breed populations, as this commercial breed to shared variation in phenotypic-trait size variation of body height, body length, leg length and head size (**Figure 3.6 F**). Makvandi-Nejad et al. (2012) profiles the body size variation within Thoroughbred horses whereby they owe 83% of size variation within this breed was due to the influence of four loci, even though breeding within Thoroughbred populations is highly controlled and selected, variation of size within the breed still occurs. However, some thoroughbred individuals were

distributed towards the upper-right quadrant of the CATPCA (**Figure 3.7**) which may be due to the combination of the phenotypic traits included within the analysis that differed from the middle cluster of the matrix. This most likely owing to the differentiation of breed-specific, well-muscled straight neck shape, with a slight curve towards the upper portion of the neck, enabling maximum head and neck flexion for racing performance (Zsoldos and Licka, 2015). The most distinct grouping with the CATPCA analysis (**Figure 3.7**) was most of the commercial Saddlebred population that were connected with a straight line, showing uniformity in their breed-specific phenotypic traits. The traits in consideration were high-arched necks, tall body length and leg length, large head size and long body length – which are all suited to their gaited showing sporting events (Staiger et al., 2016).

An ordinal regression was used to determine the effect of multi-functional roles assigned to each horse population on the household/stud farm/training facility monthly income (ZAR/Month), in order to assess the direct socio-economic impact that these breed populations have. The ordinal regression analysis yielded no significance ($p > 0.050$) of direct socio-economic impact within the rural breed populations of the Basotho-type (BSP) (**Table 3.8**), the East London Rural (ELR) (**Table 3.9**), KwaZulu-Natal (KZN) rural (**Table 3.10**) and Eastern Cape Rural (ECR) (**Table 3.11**) breeds. This is due to the fact that most rural households across South Africa are deemed as low-income households (Mamabolo et al., 2024) , as seen within the income range in **Figure 3.3**, where these small-holder farmers produce livestock and cropping systems for food production for the household. Although all four rural breed populations were significant for the utilization in rural racing events (**Table 3.5**), not all rural racing events are prominent races with high winning stakes. The cost of maintaining a racehorse, which requires high energy input through food source and various other supplements (Ebert and Moore-Colyer, 2020), may have an adverse effect on the socio-economic standing of a household. The socio-economic impact of these rural horse breed populations may not necessarily directly impact wealth generated within rural households, but their multi-functional abilities within communities offer support and improvement of their owner's lives, in order for the socio-economic level of households to be maintained or increased (de Klerk et al., 2020). Both the commercial Saddlebred (SAD) and Thoroughbred (THR) breed populations did not fit the ordinal regression model, as both of these breed populations exhibited socio-economic contributions that exceeded the monthly income range (ZAR/Month) used within this study. This is due to the fact that the commercial horse industry within South Africa was attributed as a multi-million-rand industry that impacts the socio-economic status of many sectors within South Africa (Swart, 2010). The economic, socio-economic, and cultural roles performed by our rural, indigenous, and commercial horse breed populations in South African society are, and will always remain, essential to sustaining our livelihoods.

3.6 Conclusion

This study highlights the significant and multi-faceted roles that horses play in South African society, encompassing economic, socio-economic, and cultural dimensions. Horses serve as indispensable assets, particularly in rural areas, where their utility spans agricultural practices, transportation, and traditional

recreational activities such as rural racing. The Basotho-type (BSP), Eastern Cape Rural (ECR), KwaZulu-Natal (KZN) Rural, and East London Rural (ELR) populations have demonstrated their versatility, with prominent evidence for use in agriculture and transportation. The study's findings also shed light on unique roles, such as the consumption of horse meat within certain cultural groups like the Sotho and Tswana cultural groups, emphasizing the integration of horses into longstanding traditions. Rural racing, particularly among the East London Rural (ELR) population, was shown to have socio-economic benefits by fostering community cohesion and empowering small enterprises, especially women-led SMMEs. However, while the rural horse populations contribute indirectly to socio-economic stability of households, their impact on direct household income appears limited, reflecting the broader challenges faced by low-income rural households. In contrast, the commercial Saddlebred (SAD) and Thoroughbred (THR) breed populations significantly influence South Africa's equine sporting industry, a thriving multi-million-rand sector. These breeds excel in high-stakes racing and showing events, boosting the economic landscape and supporting equine-related industries such as breeding, veterinary services, and feed manufacturers. This study provides a foundation for the classification of rural horse breeds based on phenotypic traits, which, although highly variable and oftern shared among populations exposed to similar environmental conditions – reflects the influence of functional roles in shaping these populations. In conclusion, the multi-functional roles of horses remain pivotal in supporting of livelihoods, from rural farmlands to commercial sectors, the significance of these horses are propelled by their unique ability to adapt and fulfil the evolving needs of South African communities.

3.7 References

Agafonkin, V. (2016). Leaflet for R. RPubs. Available: http://rstudio.github.io/leaflet/ [accessed May 17, 2025].

Atkinson, W. (2023). Charting fields and spaces quantitatively: from multiple correspondence analysis to categorical principal components analysis. Quality and Quantity 58, 829-848.

Bailey, E., Petersen, J.L., Kalbfleisch, T.S. (2022). Genetics of Thoroughbred racehorse performance. Annual Review of Animal Biosciences 10, 131-150.

Brooks, S.A., Makvandi-Nejad, S., Chu, E., Allen, J.J., Streeter, C., Gu, E., McCleery, B., Murphy, B.A., Bellone, R., Sutter, N.B. (2010). Morphological variation in the horse: defining complex traits of body size and shape. Animal Genetics 41(2), 159-165.

Brubaker, L., Udell, M.A.R. (2016). Cognition and learning in horses (*Equus caballus*): What we know and why we should ask more. Behavioural Processes 126, 121-131.

Bryan, K., McGivney, B.A., Farries, G., McGettigan, P.A., McGivney, C.L., Gough, K.F., MacHugh, D.E., Katz, L.M., Hill, E.W. (2017). Equine skeletal muscle adaptions to exercise and training: evidence of differential regulation of autophagosomal and mitochondrial components. BMC Genomics 18(595). Available: https://doi.org/10.1186/s12864-017-4007-9 [accessed December 05, 2024].

Danso-Abbeam, G., Ogundeji, A.A., Asale, M.A., Baiyegunhi, L.J.S. (2024). Effects of livestock ownership typology on household food security in rural Lesotho. GeoJournal 89. Available: https://doi.org/10.1007/s10708-024-11049-y [accessed November 22, 2024].

DeCarlo, L.T. (2003). Using the PLUM procedure of SPSS to fit unequal variance and generalized signal detection models. Behaviour Research Methods, Instruments and Computers 35(1), 49-56.

De Klerk, J.N., Quan, M., Grewar, J.D. (2020). Socio-economic impacts of working horses in urban and peri-urban areas of the Cape Flats, South Africa. Journal of the South African Veterinary Association 91. Available: https://doi.org/10.4102/ jsava.v91i0.2009 [accessed November 25, 2024].

Ebert, M., Moore-Colyer, M.J.S. (2020). The energy requirements of racehorses in training. Translational Animal Science 4(4). Available: https://doi.org/10.1093/tas/txaa196 [accessed December 05, 2024].

Fang, Y., Wang, J. (2012). Selection of the number of clusters via the bootstrap method. Computational Statistcs and Data Analysis 56(1), 468-477.

Gere, A. (2023). Recommendations for validating hierarchical clustering in consumer sensory projects. Current Research in Food Science 6. Available: https://doi.org/10.1016/j.crfs.2023.100522 [accessed December 04, 2024].

Getchell, J.K., Vatta, A.F., Motswatswe, P.W., Krecek, R.C., Moerane, R., Pell, A.N., Tucker, T.W., Leshomo, S. (2002). Raising livestock in resource-poor communities of the North West Province of South Africa – a participatory rural appraisal study. Journal of the South African Veterinary Association 73(4), 177-184.

Hoffecker, J.F., Holliday, V.T., Stepanchuk, V.N., Lisitsyn, S.N. (2018). The hunting of horse and the problem of the Aurignacian on the central plain of Eastern Europe. Quaternary International 492, 53-63.

Hole, C., Murray, R., Marlin, D., Freeman, P. (2023). Equine behavioural and physiological responses to auditory stimuli in the presence sand absence of noise-dumping ear covers. Animals (Basel) 13(9). Available: https://doi.org/10.3390/ani13091574 [accessed November 20, 2024].

Hutson, J.M., Villaluenga, A., Garcia-Moreno, A., Turner, E., Gaudzinski-Windheuser, S. (2024). Persistent predators: Zooarchaeological evidence for specialized horse hunting at Schoningen 13II-4. Journal of Human Evolution 196. Available: https://doi.org/10.1016/j.jhevol.2024.103590 [accessed November 20, 2024].

IBM Corporation. Released 2023. IBM SPSS Statistics for Windows, Version 29.0.2.0 Armonk, NY: IBM Corp.

Katz, L.M., O'Dwyer, S., Pollock, P.J. (2009). Nutritional muscular dystrophy in four-day-old Connemara foal. Irish Veterinary Journal 62. Available: https://doi.org/10.1186/2046-0481-62-2-119 [accessed December 05, 2024].

Klecel, W., Martyniuk, E. (2021). From the Eurasian steppes to the Roman circuses: A review of early development of horse breeding and management. Animals (Basel) 11(7). Available: https://doi.org/10.3390/ani11071859 [accessed November 21, 2024].

Kohn, H., Hubert, L.J. (2015). Hierarchical cluster analysis. Wiley StatsRef. Statistics Reference Online. Available: https://doi.org/10.1002/9781118445112.stat02449.pub2 [accessed November 26, 2024].

Koski, S.E., Spannari, J. (2024). Experiences of intersubjectivity and agency in human-horse relationships differ between box- and open-shed-stable environments. Anthrozoös 37(4), 687-704.

Kysely, R., Peske, L. (2022). New discoveries change existing views on the domestication of the horse and specify its role in human prehistory and history – a review. Archeologicke Rozhledy LXXIV-2022. Available: http://dx.doi.org/10.35686/AR.2022.15 [accessed November 21, 2024].

Lenders, R. (2024). 'Wild' horses in medieval and early modern landscapes of Europe. Landscape History 44(2), 15-35.

Letsoalo, D.L., Menoe, M.B., Semenya, M.V. (2024). Equine-assisted therapy in South Africa: A scoping review of applications and effectiveness. E-Journal of Humanities, Arts and Social Sciences 5(13), 2328-2343.

Lewu, F.B., Assefa, Y. (2009). Farmers' knowledge in the cropping systems of northern KwaZulu-Natal, South Africa: Current challenges and solution for sustainable future food production. African Journal of Agricultural Research 4(11), 1148-1153.

Librado P., Khan N., Fages A., Kusliy M. A., Suchan T., Tonasso-Calvière L., Schiavinato S., Alioglu D., Fromentier A., Perdereau A., Aury J.-M., Gaunitz C., Chauvey L., Seguin-Orlando A., Der Sarkissian C., Southon J., Shapiro B., Tishkin A. A., Kovalev A. A., Alquraishi S., Alfarhan A. H., Al-Rasheid K. A. S., Seregély T., Klassen L., Iversen R., Bignon-Lau O., Bodu P., Olive M., Castel J.-C., Boudadi-Maligne M., Alvarez N., Germonpré M., Moskal-del Hoyo M., Wilczyński J., Pospuła S., Lasota-Kuś A., Tunia K., Nowak M., Rannamäe E., Saarma U., Boeskorov G., Lõugas L., Kyselý R., Peške L., Bălăşescu A., Dumitraşcu V., Dobrescu R., Gerber D., Kiss V., Szécsényi-Nagy A., Mende B. G., Gallina Z., Somogyi K., Kulcsár G., Gál E., Bendrey R., Allentoft M. E., Sirbu G., Dergachev V., Shephard H., Tomadini N., Grouard S., Kasparov A., Basilyan A. E., Anisimov M. A., Nikolskiy P. A., Pavlova E. Y., Pitulko V., Brem G., Wallner B., Schwall C., Keller M., Kitagawa K., Bessudnov A. N., Bessudnov A., Taylor W., Magail J., Gantulga J.-O., Bayarsaikhan J., Erdenebaatar D., Tabaldiev K., Mijiddorj E., Boldgiv B., Tsagaan T., Pruvost M., Olsen S., Makarewicz C. A., Valenzuela Lamas S., Albizuri Canadell S., Nieto Espinet A., Iborra M. P., Lira Garrido J., Rodríguez González E., Celestino S., Olària C., Arsuaga J. L., Kotova N., Pryor A., Crabtree P., Zhumatayev R., Toleubaev A., Morgunova N. L., Kuznetsova T., Lordkipanize D., Marzullo M., Prato O., Bagnasco Gianni G., Tecchiati U., Clavel B., Lepetz S., Davoudi H., Mashkour M., Berezina N. Ya., Stockhammer P. W., Krause J., Haak W., Morales-Muñiz A., Benecke N., Hofreiter M., Ludwig A., Graphodatsky A. S., Peters J., Kiryushin K. Yu., Iderkhangai T.-O., Bokovenko N. A., Vasiliev S. K., Seregin N. N., Chugunov K. V., Plasteeva N. A., Baryshnikov G. F., Petrova E., Sablin M., Ananyevskaya E., Logvin A., Shevnina I., Logvin V., Kalieva S., Loman V., Kukushkin I., Merz I., Merz V., Sakenov S., Varfolomeyev V., Usmanova E., Zaibert V., Arbuckle B., Belinskiy A. B., Kalmykov A., Reinhold S., Hansen S., Yudin A. I., Vybornov A. A., Epimakhov A., Berezina N. S., Roslyakova N., Kosintsev P. A., Kuznetsov P. F., Anthony D., Kroonen G. J., Kristiansen K., Wincker P., Outram A., Orlando L. (2021). The origins and spread of domestic horses from the Western Eurasian steppes. Nature 598, 634-640.

Makvandi-Nejad, S., Hoffman, G.E., Allen, J.J., Chu, E., Gu, E., Chandler, A.M., Loredo, A.I., Bellone, R.R., Mezey, J.G., Brooks, S.A., Sutter, N.B. (2012). Four loci explain 83% of size variation in the horse. PLOS One 7(7). Available: https://doi.org/10.1371/journal.pone.0039929 [accessed December 05, 2024].

Mamabolo, M., Chaminuka, P., Machethe, C. (2024). District analysis of poverty dynamics in rural South Africa. African Journal of Science, Technology, Innovation and Development 15(6). Available: https://hdl.handle.net/10520/ejc-aa_ajstid_v15_n6_a690 [accessed December 05, 2024].

Marlow, C.H.B. (2010). A brief history of equine private practice in South Africa. Journal of the South African Veterinary Association 81(4), 190-200.

Marshall, A., Altman, D.G., Holder, R.L., Royston, P. (2009). Combining estimates of interest in prognositic modelling studies after multiple imputation: current practice and guidelines. BMC Medical Research. Available: https://doi.org/10.1186/1471-2288-9-57 [accessed November 27, 2024].

Mhlauli, O., Ezeuduji, I.O. (2022). When a tourism event goes rustic: Marketing a rural horse racing event in KwaZulu-Natal. Journal of Hospitality, Tourism and Leisure 11(1), 156-168.

Millin, P. (2007). The renowned Basotho pony. SA Horseman, 31-34.

Mofuoa, K.V. (2015). The exemplary ethical leadership of the King Moshoeshoe of Basotho of Lesotho in the nineteenth century Southern Africa. Journal of Public Administration and Governance 5(3), 21-35.

Monau, P.I., Visser, C., Nsoso, S.J., van Marle-Koster, E. (2018). Phenotypic and genetic characterization of indigenous Tswana goats. South African Journal of Animal Science 48(5), 925-934.

Ngumbela, X.G. (2023). Empowering rural women through tourism entrepreneurship: The case of the Annual Berlin November horse racing in the Eastern Cape Province. International Journal of Social Science Research and Review 6(11), 195-207.

Pandey, S.K., Arora, V.K., Gael, R., Singh, R. (2001). Genetic and phenotypic study of some production traits of Hariana Cattle. Indian Journal of Animal Research 35, 129-131.

Paterson, C. (2021). Notes on the origin of 'the chase': Artefacts of an indigenous racing tradition in Transkei. Kronos 47(1), 36-59.

Phafoli, L.S., Zulu, N.S. (2012). The complexity of the cultural identity of Basotho in Lesotho. LWATI: A Journal of Contemporary Research 9(1), 238-253.

Pienaar, L., Meyer, F., Otterman, H., Davids, T. (2021). Deep-dive of the animal feed industry in South Africa: Strategic analysis. Report commissioned by the animal feed manufacturers association of South Africa. Pretoria: Bureau for Food and Agricultural Policy. Available: https://www.researchgate.net/publication/363670158_The_South_African_Feed_Industry_A_Strategic_Perspective [accessed December 04, 2024].

Plackett, R.L. (1983). Karl Pearson and the chi-squared test. International Statistical Review 51, 59-72.

Ribeiro, N.L., Rodrigues de Medeiros, G., Vieira da Silva, N., Soares, K.O., Arandas, J.K.G., Vieira do Nascimento, G., Ribeiro, M.N. (2024). Phenotypic characterisation and production system of the Nordestino horse form a multivariate perspective. Scientific Reports 14. Available: https://doi.org/10.1038/s41598-023-51018-y [accessed November 26, 2024].

Ripolles, M., Sanchez-Guerrero, M.J., Perdomo-Gonzalez, D.I., Azor, P., Valera, M. (2020). Survey of risk factors and genetic characterization of ewe neck in a world population of Pura Raza Espanol horses. Animals 10(10). Available: https://doi.org/10.3390/ani10101789 [accessed December 05, 2024].

RStudio Team. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA; 2020. http://www.rstudio.com/. Accessed 23 Sep 2024.

Rydberg, T., Jansen, J. (2002). Comparison of horse and tractor traction using emergency analysis. Ecological Engineering 19(1), 13-28.

Sanaei, E., Seiedy, M., Momtazi, F. (2015). Evolutionary view of sexual dimorphism and shape variation in Iranian populations of Hypera postica (Coleoptera: Curculionidae). Zoomorphology. Available: https://10.1007/s00435-015-0279-2 [accessed November 26, 2024].

Schaefer, R.J., McCue, M.E. (2020). Equine genotyping arrays. Veterinary Clinics of North American Equine Practice 36(2). Available: https://doi.org/10.1016/j.cveq.2020.03.001 [accessed May 16, 2025].

Scholtz, M., Guthrie, A.J., Fosgate, G.T., Schulman, M.L. (2022). Predictors of foaling outcomes in barren and maiden Thoroughbred mares in South Africa. Journal of the South African Veterinary Association 93(2), 168-175.

Sebbane, M., Vial, C., Lmay, A. (2023). A horse on your plate? A cluster analysis of French consumers hippophagy acceptance. Meat Science 203. Available: https://www.sciencedirect.com/science/article/pii/S0309174023001262 [accessed November 22, 2024].

Skinner, J.E., Hilly, L., Li, X., Cawdell-Smith, A.J., Bryden, W.L. (2022) in "Equine production systems and the changing role of horses in society". Livestock: Production, Management Strategies and Challenges (Editors V.R. Squires and W.L. Bryden). Nova Science Publishers, NY. pp. 389-433.

Smith, T.J., Walker, D.A., McKenna, C.M. (2019). An exploration of link functions used in ordinal regression. Journal of Modern Applied Statistical Methods 18(1). Available: https://dx.doi.org/10.22237/jmasm/1556669640 [accessed November 27, 2024].

Soal, J. (1999). Meaty prejudices give food for thought. IOL. Available: https://www.iol.co.za/news/south-africa/meaty-prejudices-give-food-for-thought-20109#:~:text=Horse%20meat%20is%20part%20of,%22Meat%20is%20meat [accessed December 04, 2024].

Surujlal, J., Mosai, W., Dhurup, M. (2011). 'Sport of kings': analysis of motives for gambling on horseracing. African Journal for Physical Health Education, Recreation and Dance 17(2). Available: https://hdl.handle.net/10520/EJC121735 [accessed November 25, 2024].

Staiger, E.A., Bellone, R.R., Sutter, N.B., Brooks, S.A. (2016). Morphological variation in gaited horse breeds. Journal of Equine Veterinary Science 43, 55-65.

Swart, S. (2003). Riding high – horses, power and settler society, c. 1654-1840. Kronos, Environmental History 29. Available: https://doi.org/10.2307/41056494 [accessed September 19, 2023].

Swart, S. (2008). 'High Horses': Horses, class and socio-economic change in South Africa. Journal of South African Studies 34(1), 193-213.

Swart, S. (2010). "The world the horses made": A South African case study of writing animals into social history. International Review of Social History 55(2), 241-263.

Van Dierendonck, M., Spruijt, B.M. (2012). Coping in groups of domestic horses – Review from a social and neurobiological perspective. Applied Animal Behaviour Science 138(3), 194-202.

Wadham, H., Wallace, C., Furtado, T. (2022). Agents of sustainability: How horse and people co-create, enact and embed the good life in rural places. Sociologica 63(3), 390-414.

White, D.R., Korotayev, A. (2004). Statistical analysis of cross-tabs. Irvine: Anthropology Department, University of California. Available: https://www.researchgate.net/publication/277293916_Chapter_5_Statistical_Analysis_of_Cross-Tabs [accessed November 25, 2024].

Zantsi, S., Christian, M. (2023). Is there a case for supporting animal traction research, policy, and practice in rural South Africa? – A review. South African Journal of Agricultural Extension 51(1), 34-50.

Zhang, Z., Murtagh, F., Van Poucke, S., Lin, S., Lan, P. (2017). Hierarchical cluster analysis in clinical research with heterogeneous study population: highlighting its visualization with R. Annuals of Translational Medicine 5(4). Available: https://doi.org/10.21037/atm.2017.02.05 [accessed November 26, 2024].

Zsoldos, R.R., Licka, T. (2015). The equine neck and it function during movement and locomotion. Zoology 118(5), 346-376.

Chapter 4

Unravelling Genetic Variability in South African Indigenous and Rural Horse Breeds through SNP Genotypic Markers

4.1 Abstract

South Africa hosts a diverse range of indigenous horse breeds that thrive in both commercial breeding operations and rural farming communities. The infusion of Javanese war horses by Dutch settlers marked the genesis of the Cape Horse, involving subsequent crossbreeding, with diverse equine stocks such as Arabian, Persian, English Thoroughbred and North American Saddlebred breeds. There have been very limited genomic studies completed on these indigenous horse breed populations, utilizing foundational genomic techniques. A total of 282 indigenous (Basotho-type Pony (BSP)), rural (Eastern Cape Rural (ECR); East London Rural (ELR) and KwaZulu-Natal (KZN) rural) and commercial (SA Saddlebred (SAD); Thoroughbred (THR)) South African horse specimens were collected in KwaZulu-Natal, Eastern Cape and Gauteng provinces. Therefore, this genomics-assisted, community-driven study aims to characterize the genetic and functional diversity of the indigenous and rural horse populations within South Africa. South African PCA analysis revealed 4 distinct clusters whereby the two purebred, commercial breeds (SAD and THR) and the ELR population clustered away from the largest cluster (BSP, ECR and KZN). Global PCA demonstrated how all South African breed populations clustered together, surrounded by continental representative breed clusters, within the dataset. Ancestral proportions analysis (sNMF) indicated shared ancestry among the 4 clusters, suggesting historical connections linked to both SAD and THR precursor breed. Pairwise Fixation Index analysis demonstrated minimal population genetic differentiation between the indigenous and rural breeds, reinforced by AMOVA results, which highlighted low inter-population (F_{st} <0.060) and high intra-population (F_{IS} = 94.390%) variation. Homozygous Runs of Homozygosity (ROH) analyses revealed a smaller range of accumulative runs in Basotho pony (BSP), Eastern Cape Rural (ECR) and Thoroughbred (THR) representative populations – all consisting of a base length summation of <300Mbps. East London Rural (ELR) exhibited a variation in population run lengths (300 to 800 Mbps), indicating recent and past inbreeding or selective events occurring within the rural population. Both SA Saddlebred (SAD) and KwaZulu-Natal (KZN) breed populations exhibited a higher range of total runs. Top ROH islands/peaks (>0.700) revealed beneficial, adaptive selection signatures in KZN and ELR breed populations associated with fertility and reproduction (*NLRP13*; *NLRP5*), immunity (*FIZ1*; *SSC5D*; *ZFP90*) and growth and development traits (*EPN1*; *SNAI1*). Both commercial, purebred populations contained breed-specific selection signatures associated with the breed's sporting discipline, however, two deleterious genes (*MAGI1*; *ADGRL3*) were detected within two islands carried within the SAD population. Overall population inbreeding levels (F_{ROH}) of all breed populations exhibited a large range from $0.032 \leq F_{ROH} \leq 0.326$. SAD, ELR and KZN breed populations contained the highest instances of inbred individuals, while BSP, ECR and THR contained the lowest instances of inbred individuals. In conclusion, this study offers valuable insight into the maintained high, intra-specific genetic diversity of our

rural and indigenous populations that highlights their unique characteristics as evidence for their adaptability and survivability – noting the beneficial influence of commercial breeds.

Keywords: South Africa; Indigenous horse breeds; Rural horse breeds; Genetic diversity; Continental breed influence; ROH islands; Adaptive selection signatures; Inbreeding

4.2 Introduction

South Africa boasts diverse indigenous and rural horse breeds that thrive not only in commercial breeding operations but also in rural farming areas (Van der Merwe and Martin, 2002). The existence of these unique commercial and rural indigenous breed populations is the resulting outcome of several historical events that occurred throughout South African horse breeding history, originating from their common ancestor commonly referred to as the Cape Horse (Joubert and Bosman, 1971). One of the most significant historical events was the colonization of South Africa by Dutch settlers in 1652, when Europeans docked their ships at the southern tip of the country. Their aim was to establish a supply station for Dutch trading ships circumnavigating the Cape of Good Hope to and from the Dutch East Indies—a strategic initiative orchestrated by the Dutch East India Company (VOC) (Hunt, 2005). Swart (2003) pays tribute to the fact that horses were the first domesticated livestock species to be shipped to the Dutch colony, due to their multi-purpose ability in agricultural practices, mounted warfare and transport within and beyond the colony. These horses were referred to as "Javanese ponies", sturdy and small-statured stock imports from the Dutch colony Java (East Indies island) – a composite breed of Arabian and Mongolian horse breeds provided by Arab traders that infiltrated the East Indies trading market (Swart, 2003). Marlow (2010) outlines that further breed improvements were made to this Javanese horse population in order to meet the anthropogenic needs of the Dutch settlers, through the crossbreeding of imports from North and South American stock, English Thoroughbreds, Persian stock and Arabian stock – creating a genetically and functionally diverse resulting breed known as the Cape Horse.

The invasion of the Cape Colony by the British Empire in 1795 (Potgieter and Grundlingh, 2007) significantly influenced the diversification of the original Cape Horse population. This event forced Dutch settlers to abandon their coastal settlements and undertake the Great Trek inland (Beyer, 2002). The Dutch settlers made their way up throughout the country, with various groups splitting off along the way to form multiple settlements made in the Provinces now known as the Free State, Eastern Cape and Northern KwaZulu-Natal that all border the Kingdom of Lesotho. These historical regions harbored ethnic tribes that rebelled against the uprising of Shaka Zulu's empire (Lye, 1967). These rebellions caused an uprising known as the "Mfecane" or the "Difaqane" wars, which resulted in many raids of the Dutch settlements and the capturing of the inhabitants' horses, creating a mounted warhorse society within the ethnic rebellions of the regions (Knapp, 2021). These historical events shaped horse populations through crossbreeding and adaptation to diverse environmental pressures, ultimately contributing to the development of today's rural horse ecotypes.

The Basotho people's mounted culture was influenced by the acquisition of horses from Dutch settlers and the prized animals presented to King Moshoeshoe by warriors (Eldredge, 1988). These horses, originating from

the Cape Horse population, were isolated in the harsh mountainous terrain of Lesotho, characterized by rocky landscapes, extreme temperatures, and limited resources. Through natural selection, only the most hardy individuals survived, giving rise to the renowned Basotho pony breed, known for its endurance, sure-footedness, and low-maintenance requirements (Millin, 2007). The British military observed the strength and agility of the horses that the Basotho warriors rode, and therefore sought to acquire a significant number of these horses for their own mounted cavalry during the Anglo-Boer War (1899-1902), as well as for export as prized warhorse stock (McCormack, 2018). The selective export of the finest Basotho ponies had a detrimental impact on the breed's genetic diversity within Lesotho, leading to a significant population decline by the mid-20th century (Lekota, 2014). To address this crisis, the South African Department of Agriculture established a breeding program involving Basotho-type mares and Anglo-Boer stallions. This initiative preserved the Basotho pony breed found within Lesotho and surrounding regions of South Africa today. (Joubert and Bosman, 1971).

The development of unique horse populations within South Africa is a testament to the interplay of human ingenuity, natural selection, and the remarkable adaptability of these animals. Over generations, rural communities have shaped these populations through selective breeding, guided by cultural practices and traditional knowledge to enhance traits such as endurance, strength, and temperament for specific functional roles (de Klerk, 2020; Lekota, 2014). Simultaneously, the challenging South African environment characterized by variable terrain, harsh climatic conditions, and disease pressures (Maluleke et al., 2020)—has acted as a natural filter, favouring horses with increased hardiness and disease resistance. While limited genomic research have been conducted on these indigenous and rural breeds, Cothran and Van Dyk's (1998) study provided valuable foundational insights. The advent of modern high-throughput SNP genotyping technologies offers promising opportunities for more comprehensive and informative investigations into the genetic diversity, demographic histories, and ancestry of these fascinating populations (Balague-Dobon et al., 2022). Against this background, this study unravelled the genetic diversity of South African horses by examining the population genetic structure, demographic history, ancestry, and inbreeding levels. Through genetic analyses, we aim to gain insights into the relationships among indigenous, rural, and commercial populations (Greenbaum et al., 2016), infer ancestral origins (Suarez-Pajes et al., 2021), and assess the potential accumulation of deleterious alleles within these populations (Charlesworth, 2003). This genomics assisted, community-driven research study aims to characterize the genetic and functional diversity of the rural and indigenous horse populations within South Africa, which was achieved by the objective to genotype sampled horse populations (using Illumina EquineSNP70/80K Bead chip genotype array) in order to determine the genetic diversity of the indigenous, and rural South African horse breed populations.

4.3 Materials and Methods

4.3.1 Breeds/Populations Sampled

A total of 282 South African horse specimens (*Equus caballus*) were collected from specified rural villages, designated South African horse breeding societies and from commercial horse-race training facilities. Rural,

crossbred horse populations were sampled from Alfred Nzo (ECR; n = 48) and Amathole (ELR; n =25) districts in the Eastern Cape Province and from uMzimkhulu (KZN; n = 47) district in the KwaZulu-Natal Province. Indigenous horse breed population, Basotho-type pony (BSP; n = 95), was sampled in the rural district of Alfred Nzo within the Eastern Cape province – which borders the country Lesotho where the breed was first established. Commercial, purebred horse populations of Thoroughbred (THR; n = 47) were sampled from the Summerveld Training Centre [available: https://www.goldcircle.co.za/racing/training-centres/training-centres-summerveld/] located in the eThekwini district of the KwaZulu-Natal Province. Commercial, purebred Saddlebred (SAD; n = 20) were collected according to the allocation of breeding studs in the province of Gauteng, by the Saddlebred Society of South Africa [available: https://saddlebred.co.za/ ; est. 1942]. GPS coordinates of sample locations included within this study were visualized in **Figure 4.1** below.

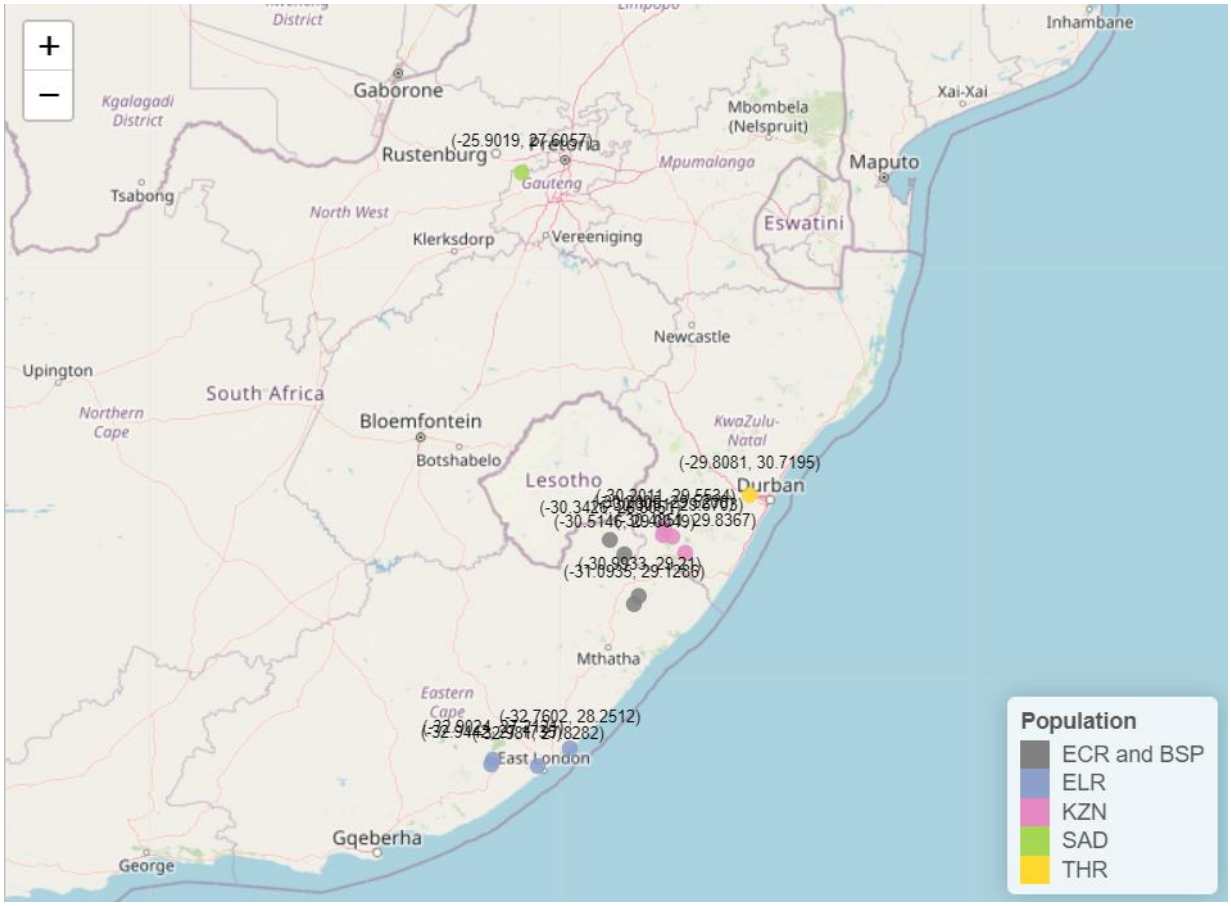


Figure 4.1 GIS map of sample regions within Eastern Cape, KwaZulu-Natal and Gauteng Provinces

[GPS coordinates of rural hourseholds, training facility and breeding stud was indicated. This GIS map was constructed using R/Rstudio version 4.3.1 (R Core Team, 2020; Rstudio Team, 2020), applying the leaflet package (Agafonkin, 2016).]

Tail-hair samples were collected in accordance with the standard operating procedures approved by Onderstepoort Veterinary Research Institutional Biosafety and Biosecurity Committee (OVR-IBBC). Ethical approval for the study was obtained from the OVR Animal Ethics Committee (**AEC 24.16**).

4.3.2 Genotyping and Quality Control

DNA was extracted from tail-hair follicles for each animal sample through the use of the NucleoMag® Tissue Kit (Macherey-Nagel) according to the user manual generated by the manufacture of the extraction kit. The concentration of extracted DNA was then quantified for each sample using the Qubit® 4.0 Fluorometer to ensure all samples surpass minimum concentration (ng/µl) requirements for genotyping.

A total of 282 horse specimens from 6 different sampled populations were genotyped using the Illumina EquineSNP70k and EquineSNP80K BeadChip (Illumina, United States), containing an array of 70, 000 SNPs and 80,000 SNPs respectively, with an average of one SNP per 35kb window [19] at the Agricultural Research Council – Biotechnology Platform (ARC-BTP) in South Africa.

PLINK version 1.90 (Purcell et al., 2007) was used to perform a Quality Control (QC) analysis on the merged, SNP genotypic data of the 6 sampled populations from South African commercial, indigenous and rural horse breeds. The QC using PLINK version 1.90 was based on the updated horse reference genome assembly - *EquCab* 3.0 (Kalbfleisch et al., 2018). The pruning parameters of the QC included SNP markers that were filtered to exclude SNP markers with a departure from Hardy-Weinberg equilibrium (HWE) of less than 10^{-6} (0.000001), SNP loci with a sample call rate of less than 90%, SNP markers with a genotyping rate of less than 90% and then SNP markers that contained a Minor Allele Frequency (MAF) of less than 5% (0.05). The remaining pruned, SNP marker dataset consisted of 79,592 SNPs and 282 horse specimens. This remaining SNP dataset was used for the following population genetic diversity analyses including F_{ST} (Fixation Index) estimation, Analysis of Molecular Variance (AMOVA), Runs of Homozygosity (ROH) detection analysis and determination of Inbreeding levels through the detection of stretches of ROH (F_{ROH}).

This dataset was then further pruned to filter out individuals that show close kinship (IBD (Identity by Descent)) ≥ 0.250, as well as to filter out SNP makers with a high Linkage Disequilibrium (LD) $r^2 = 0.500$, with a window of every 5 SNPs and a window size of 50 kb. The resulting SNP dataset consisted of 71, 595 SNPs and 282 horse specimens remaining. This filtered data was then used to complete further population genetic diversity analyses including Population Component Analysis (PCA) and sNMF analysis.

4.3.3 Population Structure and Differentiation (PCA, sNMF, Fst Pairwise, AMOVA)

Principal Component Analysis (PCA- Additive Model) was used to determine the population structure within the 6 different sampled populations, in order to ascertain the demographic history of each horse breed and the potential genetic variance between and among the populations (McVean, 2009). A subsequent PCA was conducted merging the South African indigenous, rural and commercial breeds with global horse breeds from Petersen et al. (2013), consisting of 664 horses from 34 purebred populations that represent continental breeds, originating from Asia, North America, South America and Europe. The PCA was performed using Golden Helix SNP Variation Suite (SVS) version 8.5.0 (Golden Helix Inc., 2016), firstly through the update of the

SNP marker map file based on the *EquCab* 3.0 equine genome assembly, with the inclusion of the eigenvector method in order to determine population clustering of up to 10 different components.

R/Rstudio version 4.3.1 (R Core Team, 2020; Rstudio Team, 2020) was used to infer ancestral proportions within the 6 different sampled populations, through the use of the LEA package, to conduct a sNMF analysis (Frichot et al., 2014). The sNMF sparse non-negative matrix factorization analysis was run from the estimated number of ancestral genetic clusters (*K*-value), with a minimum *K*-value = 2 and a maximum *K*-value = 10, whereby the optimum K-value considered was of that with the lowest cross-validation error. The sNMF analysis was visualized and constructed through the use of the Pophelper online website created by Francis (2017).

Population-based pairwise Fixation index (F_{st}) analysis was executed on the sampled populations, to measure the estimated genetic distance between the populations represented by rural, indigenous and commercial breeds – through the integration of the equation explained by Weir and Cockerham (1984) within Golden Helix SNP Variation Suite (SVS) version 8.5.0 (Golden Helix In., 2016). The F-statistical, Analysis of Molecular Variance (AMOVA) was utilized in order to detect population differentiation, based on prior hierarchal population clustering, among populations (F_{SC}) and within populations (F_{IS}) through the use of ARLEQUIN version 3.1 (Excoffier et al., 2005).

4.3.4 Population Genetic Diversity (Runs of Homozygosity (ROH and ROH Islands), Inbreeding and FROH)

R/Rstudio version 4.3.1 (R Core Team, 2020; Rstudio Team, 2020) was used to conduct a genome-wide scan for Runs of Homozygosity (ROH) for all six sampled breed populations, using the detectRUNS version 0.9.6 package (Biscarini et al., 2019), through the use of the sliding-window based method. The sliding-window parameters were set to a threshold = 0.500, minimum number of SNP markers in a window (minSNP) = 15, maximum number of missing runs (maxmissRun) = 1, maximum number of windows missing (maxmisswindow) = 1, the minimum length of a run measured in base pairs (minLengthBPs) = 1000 and heterozygous runs (RoHet) = 0. The accumulation of the mean number and length (bps) of ROH on each chromosome of the breed populations were determined along with the proportion of SNPs found in each run. ROH islands/peaks were determined by considering the top 30 percentile of runs found in all the breed populations, including the consideration of the highest proportions (more than 70% within the population) of SNPs found in each run. This was done in order to determine selection signatures congregated in these dense homozygous regions of the genomes of these breeds – whether due to ancestral or recent selection events (Purfield et al., 2012). A QTL analysis based on the top 30 percentile of runs was conducted, using the genome browser Ensembl (https://www.ensembl.org), in order to functionally characterize selection signatures found along these densely homozygous regions.

The detectRUNS version 0.9.6 function of R/Rstudio version 4.3.1 (R Core Team, 2020; Rstudio Team, 2020) was also used to determine inbreeding levels (F_{ROH}) within each breed population, through the use of the length of detected ROH (Mbps) compared to the total length of the genome, covered by the SNPs included

within the Illumina EquineSNP70k BeadChip (Illumina, United States) (Mbps). Inbreeding coefficients were calculated based on the implementation of the equation outlined by McQuillan et al. (2008):

$$F_{ROH} = \Sigma \frac{L_{ROH}}{L_{genome}},$$

where ΣL_{ROH} is the total length of the detected ROH and ΣL_{genome} is the total length of the equine genome covered by the SNPs found within the Illumina EquineSNP80k BeadChip (Illumina, United States). The estimated inbreeding levels (F_{ROH}) were determined and compared, on a population- and chromosomal-level, for all six sampled breed populations.

4.4 Results

4.4.1 Population Structure and Differentiation

Population structure and genetic differentiation analyses of these South African horse populations, each a representative of either indigenous, rural or commercial groups was conducted, in order to determine the genetic structure of each representative population - as well as the extent of genetic differentiation both between and within each representative population. Both PCAs (Additive model) (in **Figure 4.2** and **Figure 4.3**) and the sNMF proportion analysis (represented in **Figure 4.4**), provided the state of the population structure of each of the represented populations within the dataset. The combination of both the Analysis of Molecular Variation Analysis (AMOVA) (**Table 4.2**) and the Population-based pairwise Fixation index (F_{st}) analysis (represented in both **Table 4.1** and **Figure 4.5**) allowed for the determination of between and within representative population differentiation.

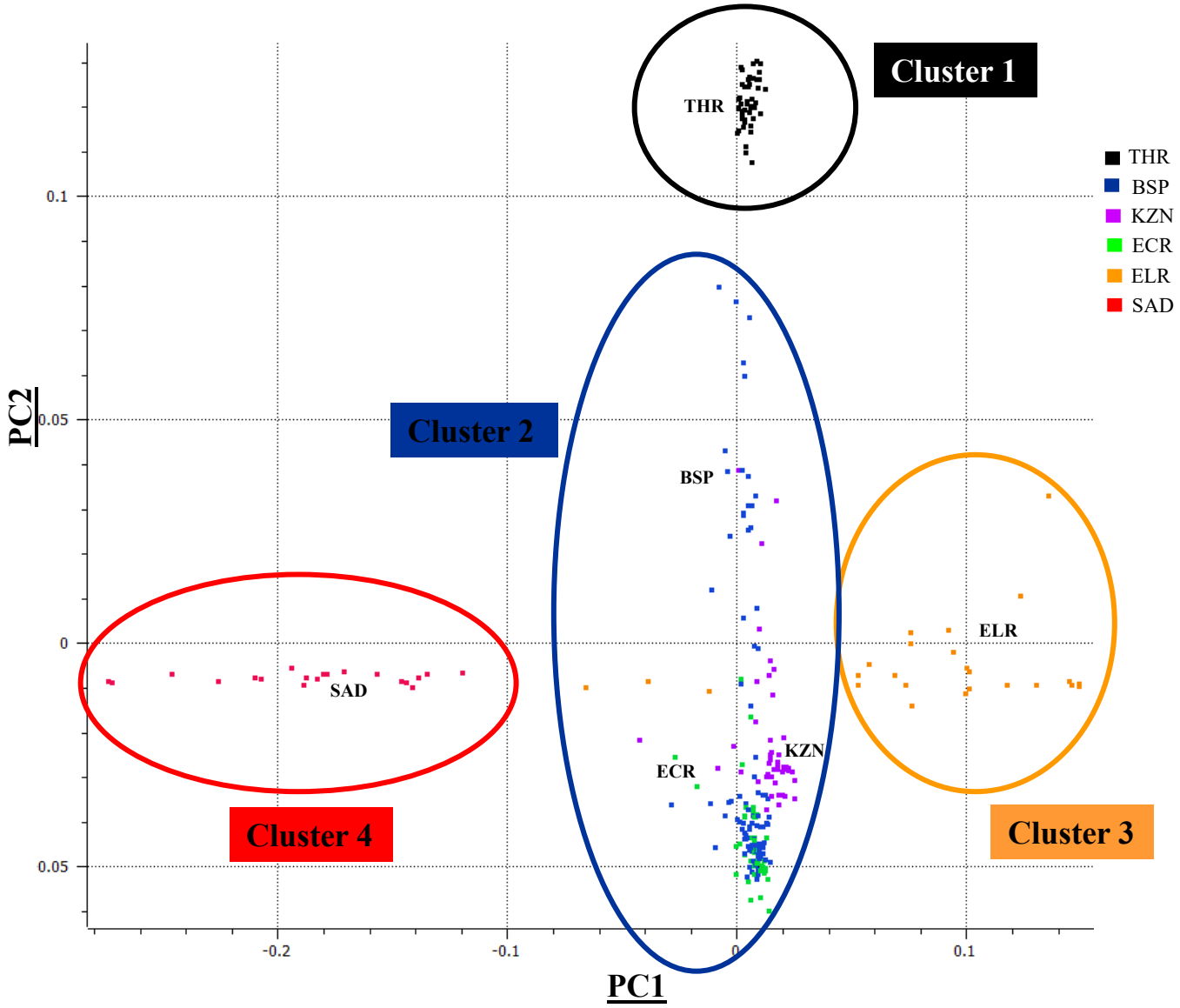


Figure 4.2 Principal Component Analysis (PCA) exhibiting population clustering

[Four distinct population clusters were identified within the PCA. The 4 clusters contained individuals from the 6 different South African horse populations as follows: **Cluster 1** exclusively contained Thoroughbred (THR) samples; **Cluster 2** contained Basotho pony (BSP), KwaZulu-Natal Rural (KZN), Eastern Cape Rural (ECR) and 3 East London Rural (ELR) samples; **Cluster 3** distinctively contained East London Rural (ELR) samples; **Cluster 4** exclusively contained SA Saddlebred samples.]

Principal component 1 (**PC1**) and principal component 2 (**PC2**) in **Figure 4.2** above, depicted a total variation of the representative population dataset as 13.020% and 39.630%, respectively. The PCA analysis indicated 4 distinct genetic clusters within the covariance matrix, whereby cluster 1 exclusively contained commercial Thoroughbred (THR) samples that showed tightly-spaced, sample arrangement within the genetic cluster expanding throughout the uppermost portion of the covariance matrix – clustered separately from the rest of the population dataset. This indicates that the South African THR breed population is a distinctly separate genetic breed from the other breed populations. Cluster 2 was found to be the largest cluster formed, comprising almost all of the rural breed populations, namely KZN, ECR and two ELR horse samples – as well as the whole BSP indigenous breed population. This finding suggests that individuals within cluster 2, sampled within different regions of South Africa, may represent a genetically distinct breed. Three BSP individuals

within cluster 2 were pulled towards the THR breed cluster (cluster 1), alluding to the precursor influence of the THR breed, that was represented strongly within the three BSP individuals. Cluster 3 exclusively contained the remaining proportion of the ELR breed population that was positioned in close proximity to the large cluster of the rural and indigenous breed populations – demonstrating close genetic relatedness amongst the rural breed populations. However, this distinct clustering may indicate the influence of a separate horse breed that contributed to the ELR sampled population, that the ECR, KZN and BSP does not contain. The last distinct cluster 4, contained only the SA Saddlebred commercial purebred that clustered away from the rest of the dataset but distributed closer to the rural and indigenous breed cluster (cluster 2) than compared to the other purebred THR population cluster (cluster 1). This indicates that the two commercial horse breed populations (THR and SAD) arose from separate ancestral populations.

The global PCA (**Figure 4.3**) showed the clustering of the South African inidgenous, rural and commercial breed populations among continental representative breed populations.

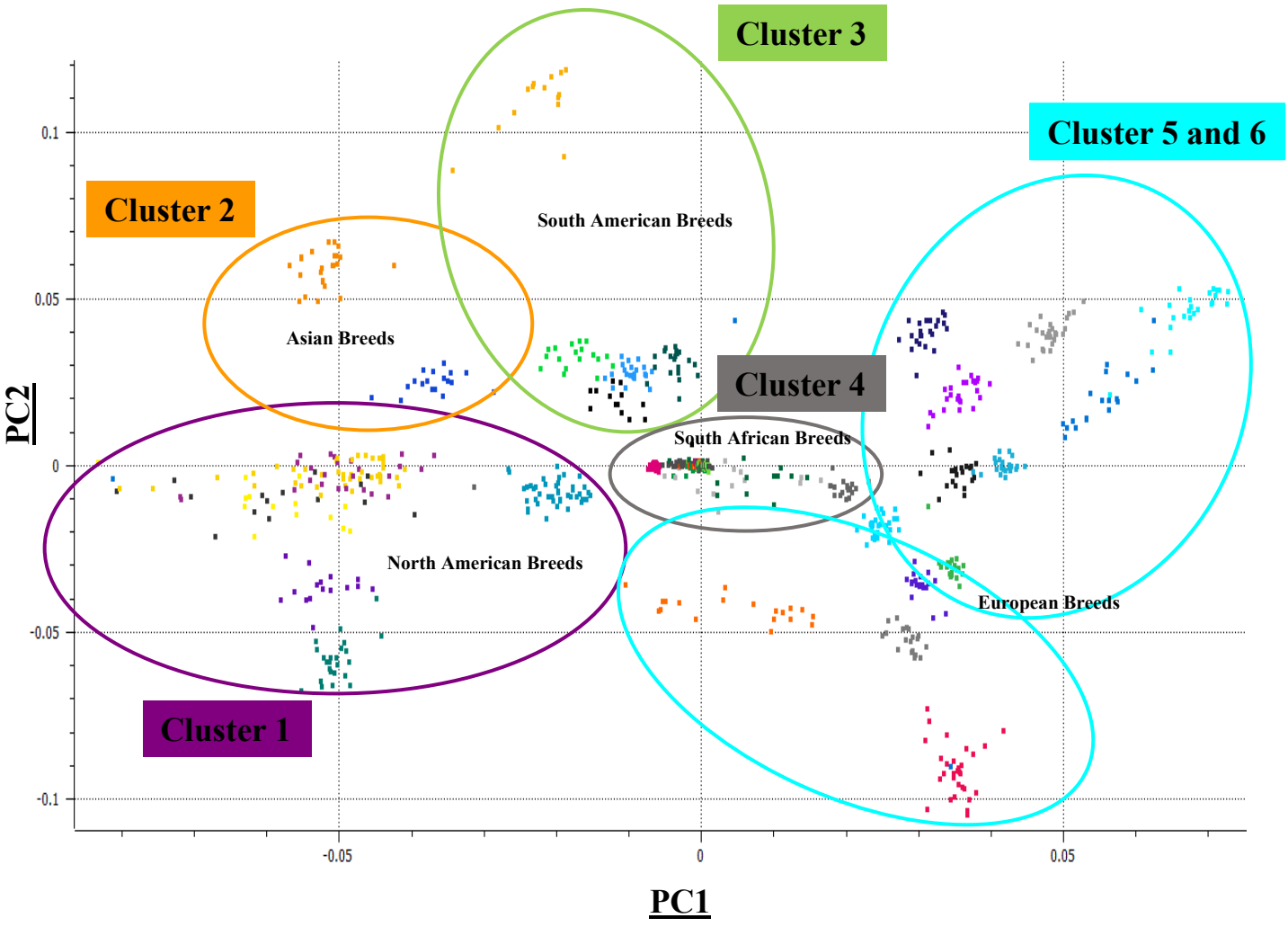


Figure 4.3 Principal Component Analysis (PCA) exhibiting population clustering including Petersen et al. (2013) breed populations (99,530 SNPs)

[Six distinct population clusters were identified within the PCA. The 6 clusters contained individuals from the 40 different global breed populations as follows: **Cluster 1** contained North American breeds (Paint, Quarter Horse, Standardbred, Morgan and Florida Cracker);

Cluster 2 contained Asian Breeds (Akhal Teke and Arabian); **Cluster 3** contained South American Breeds (Andalusian, Caspian, Mangalarga Paulista and Puerto Rican Paso Fin): **Cluster 4** contained South African breeds (BSP, ECR, ELR, KZN, THR and SAD), two Asian breeds (Tuva and Mongolian) and one European breed (New Forest Pony); **Cluster 5 and 6** contained European breeds (Franches- Montagnes, Clydesdale, Shire, Shetland, Exmoor, Miniature, Icelandic, Fell Pony, Finnhorse, North Swedish and Norwegian Fjord)]

Principal component 1 (**PC1**) and principal component 2 (**PC2**) in **Figure 4.3** above, showed a total variation of the representative population dataset as 17.660% and 6.390%, respectively. The South African indigenous, rural and commercial horse breeds all clustered together in cluster 4, which was positioned in the middle of the distance matrix – surrounded by the continental representative breed populations originating in Asia, North America, South America and Europe. Two specific Asian breeds (Tuva and Mongolian) and one European breed (New Forest Pony) clustered within the South African horse breed cluster (cluster 3).

The population structure of these six horse populations was then analyzed in closer detail, focusing on the individual classification of samples based on ancestral proportions that defined the clustering of the populations within the dataset, further. The ancestral proportions of each representative population clustering was based on the lowest cross-validation error produced from the sNMF analysis, at cluster 4 ($K = 4$; *CV error* = 0.795).

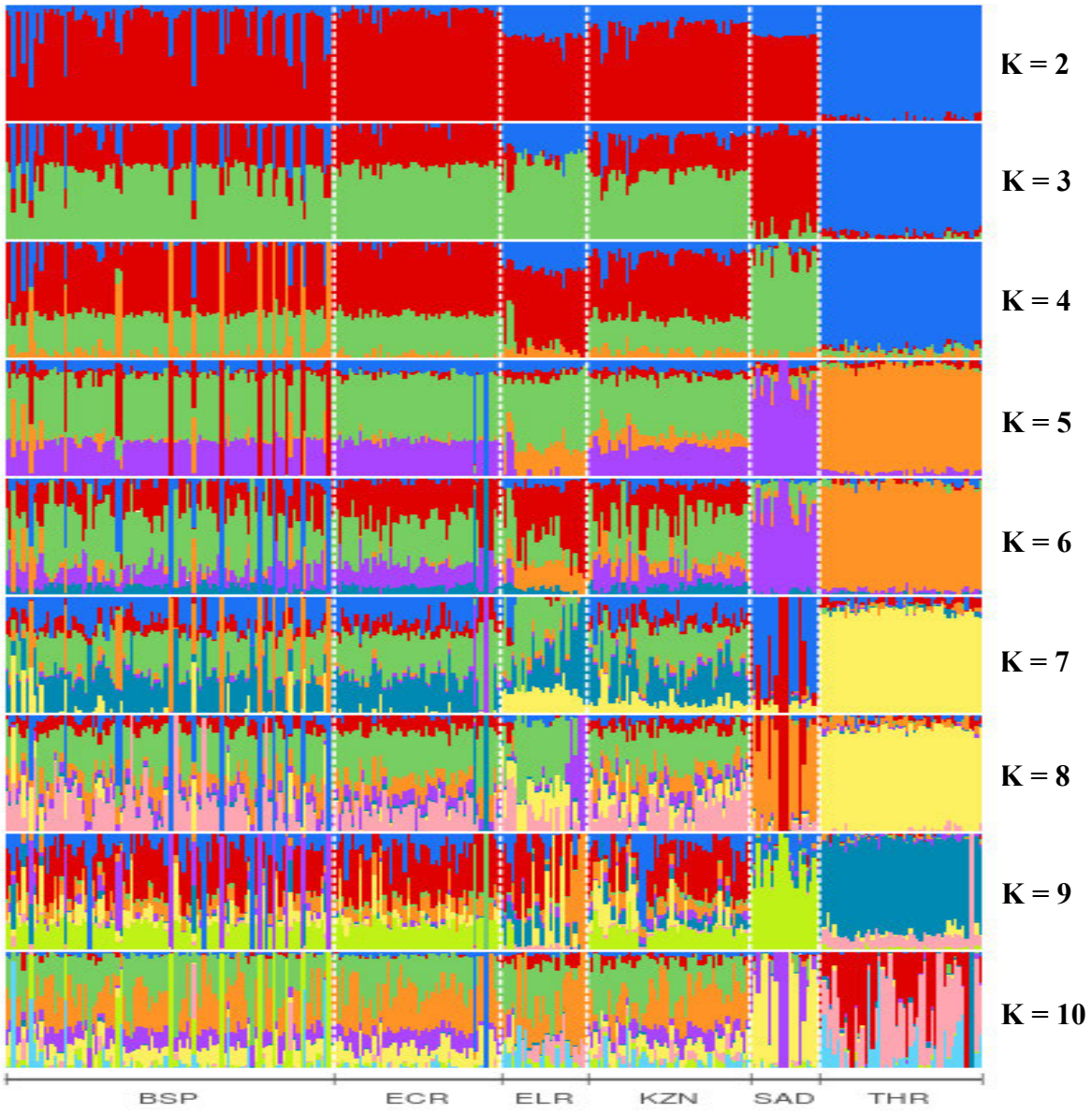


Figure 4.4 sNMF based clustering of ancestral proportions

[The clustering of the six different South African horse populations based on the ancestral clustering of K = 2 to K = 10. Each bar within the plot represents an individual horse sample of each representative population that were partitioned by K number of colours.]

At K = 4, the Thoroughbred (THR) population clustering showed a distinct, mainly uniform ancestral proportion with the inclusion of very slight, different admixed ancestral proportions. This uniform, ancestral proportion trend was maintained for the distinct Thoroughbred (THR) cluster as the K-value increases, as shown within **Figure 4.4** above. The SA Saddlebred (SAD) breed population shows three distinct ancestral proportions at K = 4, that shares commonality with the rural and indigenous horse breed populations. The clustering of both the Basotho pony (BSP), Eastern Cape Rural (ECR), East London Rural (ELR) and KwaZulu-Natal (KZN) representative populations showed to have very similar ancestral proportion plotting at K = 4, with the exception of the Basotho pony (BSP) cluster containing hyper-extended proportions, from one specific precursor influence (proportions represented in orange). All six South African indigenous, rural and commercial purebred populations shared similar ancestral proportion influence.

Population differentiation between the six representative breed populations showed a very low range of *Fst* pairwise values between the indigenous and rural populations in comparison to the commercial, purebred populations that differed significantly - seen in **Table 4.1** and represented in **Figure 4.5**.

Table 4.1 Population based *Fst* pairwise analysis

[The *Fst* pairwise values were determined between the six different horse breed populations.]

Population		BSP	ECR	ELR	KZN	THR	SAD
Indigenous	BSP	0					
Rural	ECR	0.0038	0				
	ELR	0.0051	0.0063	0			
	KZN	0.0021	0.0017	0.0039	0		
Commercial	THR	0.0814	0.1028	0.0850	0.0898	0	
	SAD	0.0626	0.0650	0.0578	0.0615	0.1460	0

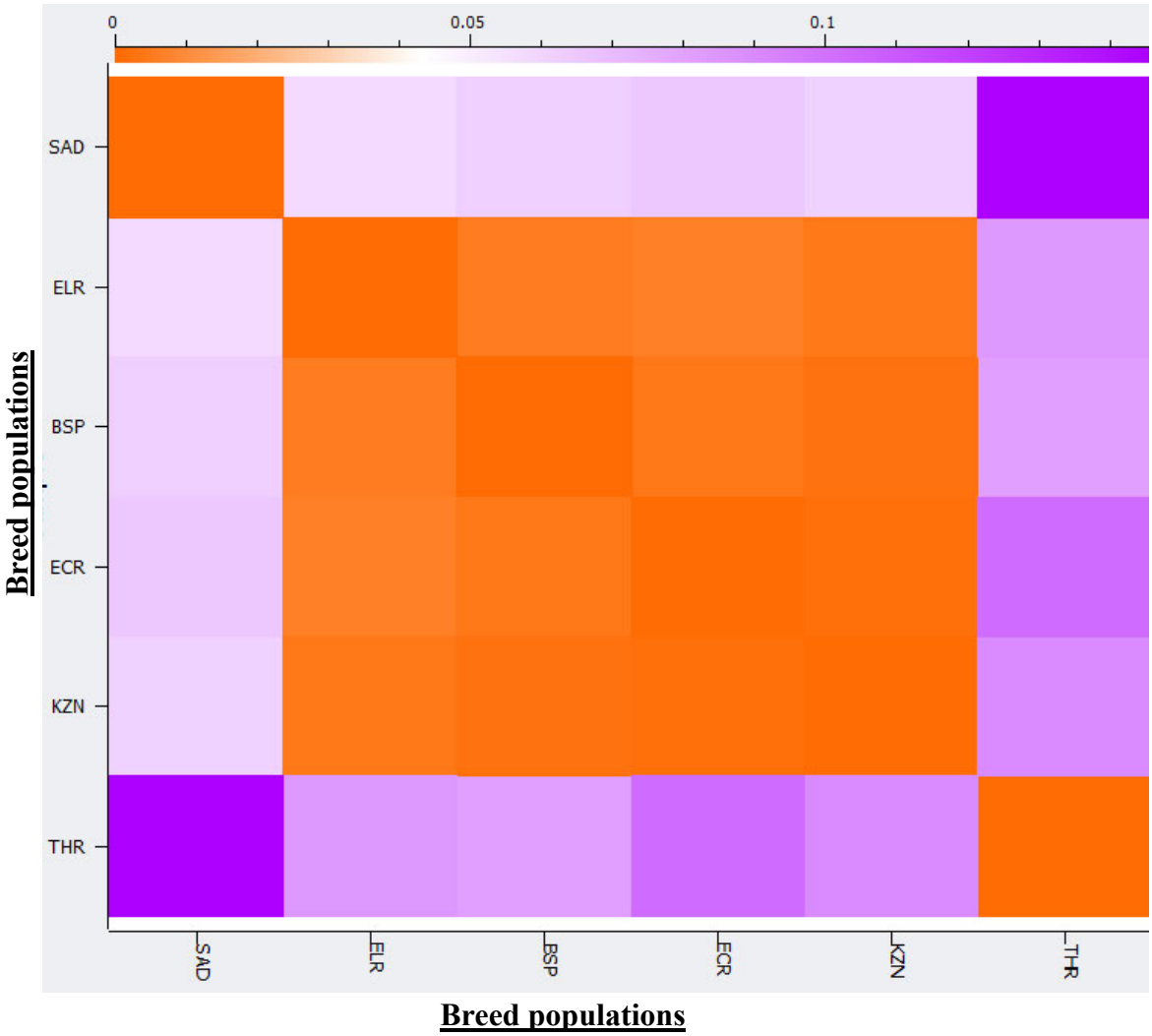


Figure 4.5 Pairwise Fst values arranged within a heatmap plot

[Values calculated from the Pairwise Fixation Index (*Fst*) analysis of genetic population differentiation between the six breed populations.]

The lowest genetic differentiation within the population differentiation analysis, with a value of *Fst* = 0.0017, was determined between the Eastern Cape Rural (ECR) and KwaZulu-Natal Rural (KZN) populations. Significantly low genetic differentiation (*Fst*) values, ranging from 0.002 – 0.005, was determined between the indigenous Basotho Pony (BSP) population and the three rural breed populations (ECR; ELR; KZN). Significant differentiation, although still a moderately low Fst value of 0.146, was found between the two purebred, commercial breed populations of the Thoroughbred (THR) and SA Saddlebred (SAD). The Thoroughbred (THR) population showed a slightly higher range of differentiation (*Fst* = 0.081 – 0.103) between the indigenous and rural breed populations than compared to the SA Saddlebred (SAD) population, which ranged from *Fst* = 0.058 – 0.065.

In order to further assess not only among population differentiation, an AMOVA (**Table 4.2**) analysis was conducted to assess within population genetic variation as well.

Table 4.2 *F*-statistical genetic differentiation within and between representative breed populations

[Analysis of Molecular Variation (AMOVA) among population (F_{CT}) and within population (F_{SC}) genetic variation, based on significance: P < 0.010, adapted from Weir and Cockerham (1984); Excoffier and Quattro (1992) and Weir (1996).]

Variation type	*df*	Sum of squares	Variance components	% Variation
Among populations ($\boldsymbol{F_{CT}}$)	5	40478.54	73.05 (*Va*)	4.43
Within populations ($\boldsymbol{F_{IS}}$)	282	438538.63	1555.10 (*Vb*)	94.39

The AMOVA results in **Table 4.2** provided further evidence of genetic differentiation between the six representative commercial, indigenous and rural breed populations collectively, to be significantly low based on *F*-statistical F_{CT} = 4.430%. This indicates that the SNP molecular markers comprised within the dataset were commonly maintained between breed populations and ultimately indicating genetic similarity among them. The AMOVA results produced a within population (F_{IS}) differentiation percentage of 94.390% within individual samples within each breed population, indicating the possession highly genetic variable markers, that each individual sample contributes towards the collective population.

4.4.2 Population Genetic Diversity

Genetic diversity between the six breed populations was determined through the accumulation of homozygous portions of the population breed genomes, drawing comparison of genetic diversity on both a population and chromosomal level. The total number of Runs of Homozygosity (ROH) in each breed population were compared to the total length (in Mbps) of those homozygous runs in **Figure 4.6** below.

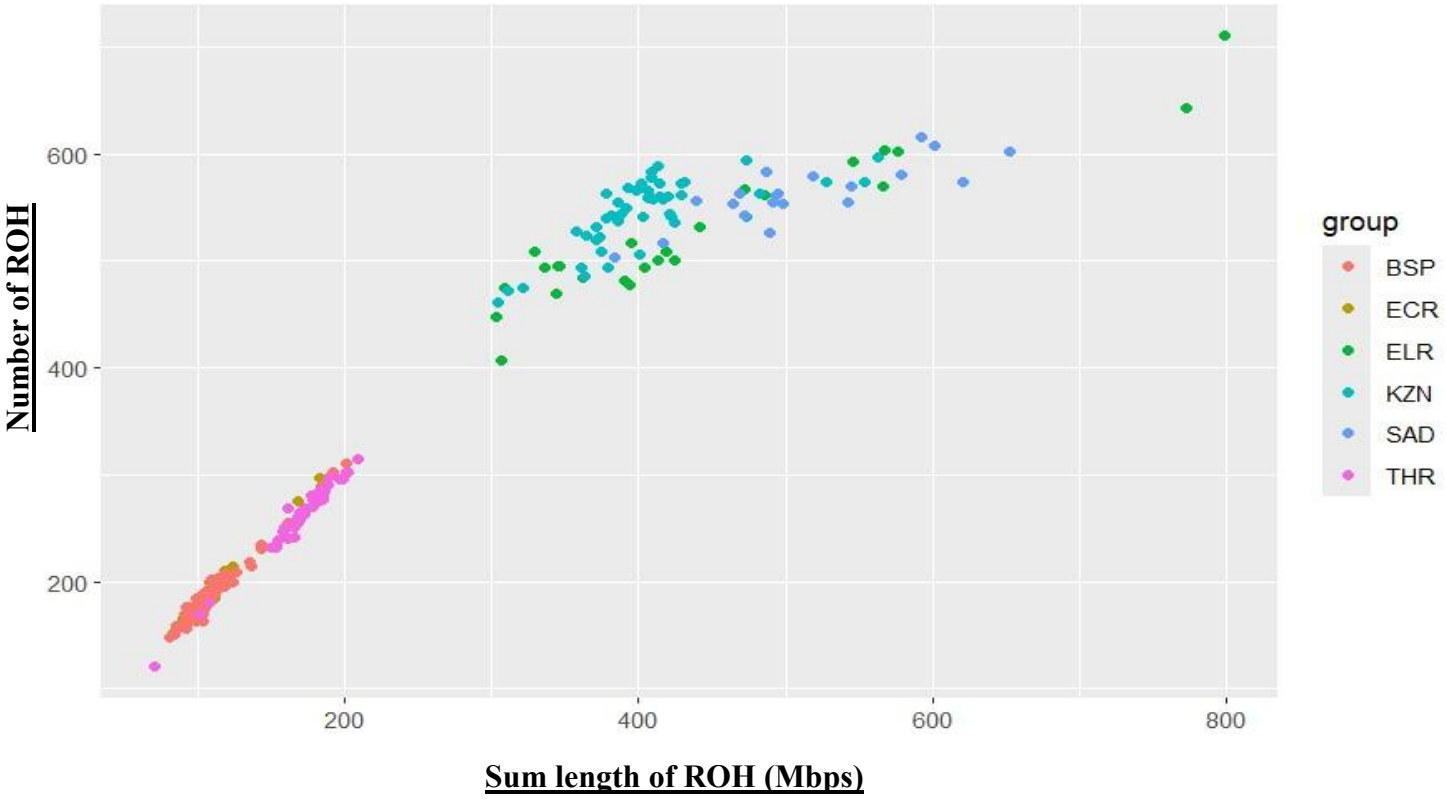


Figure 4.6 Total number of Runs of Homozygosity (ROH) stretches and associated run length (Mbps)
[The total number range of homozygous stretches found along the genome of each representative breed population, was compared to the total range length (Mbps) of the associated runs, inferring recent or ancestral inheritance of these homozygous stretches. These homozygous runs were determined based of the following "detectRUNS" function parameters of windowSize = 50, threshold = 0.050, minSNP = 20, RoHet = FALSE, maxMissWindow = 1, maxGap =10^{-6}, minLengthBPs = 250 000, minDensity = $1/10^{3}$ [SNP/Kbbps], maxOppRun = NULL and maxmissRun = NULL.]

The ROH stretches found within the Basotho pony (BSP), Eastern Cape Rural (ECR) and Thoroughbred (THR) representative populations exhibited a smaller range of total runs within the populations, as seen clustering in the left lower-side corner in **Figure 4.6**, ranging from 120 to 315 runs detected – all consisting of a base length summation of <300Mbps. This indicates that recent inbreeding events have not occured within these populations. These three different breed samples showed to have fewer homozygous runs spanning throughout their genomes that are significantly shorter in length. The East London Rural (ELR) breed population showed to have the largest range of ROH length and number congregates, ranging from 400 to >700 runs and about 300 to 800 Mbps in length. This depicts variation of homozygous proportions within the population. Both the SA Saddlebred (SAD) and KwaZulu-Natal Rural (KZN) breed populations range from the middle to uppermost right-side corner in **Figure 4.6**, with a moderate to higher range of ROH stretches and their associated run lengths.

These associated shorter, moderate and long homozygous stretches were then quantified on a chromosomal-level in each representative breed population, in order to determine their significance of inheritance, or probable cause of accumulation in each breed representative – as shown below in **Figure 4.7**.

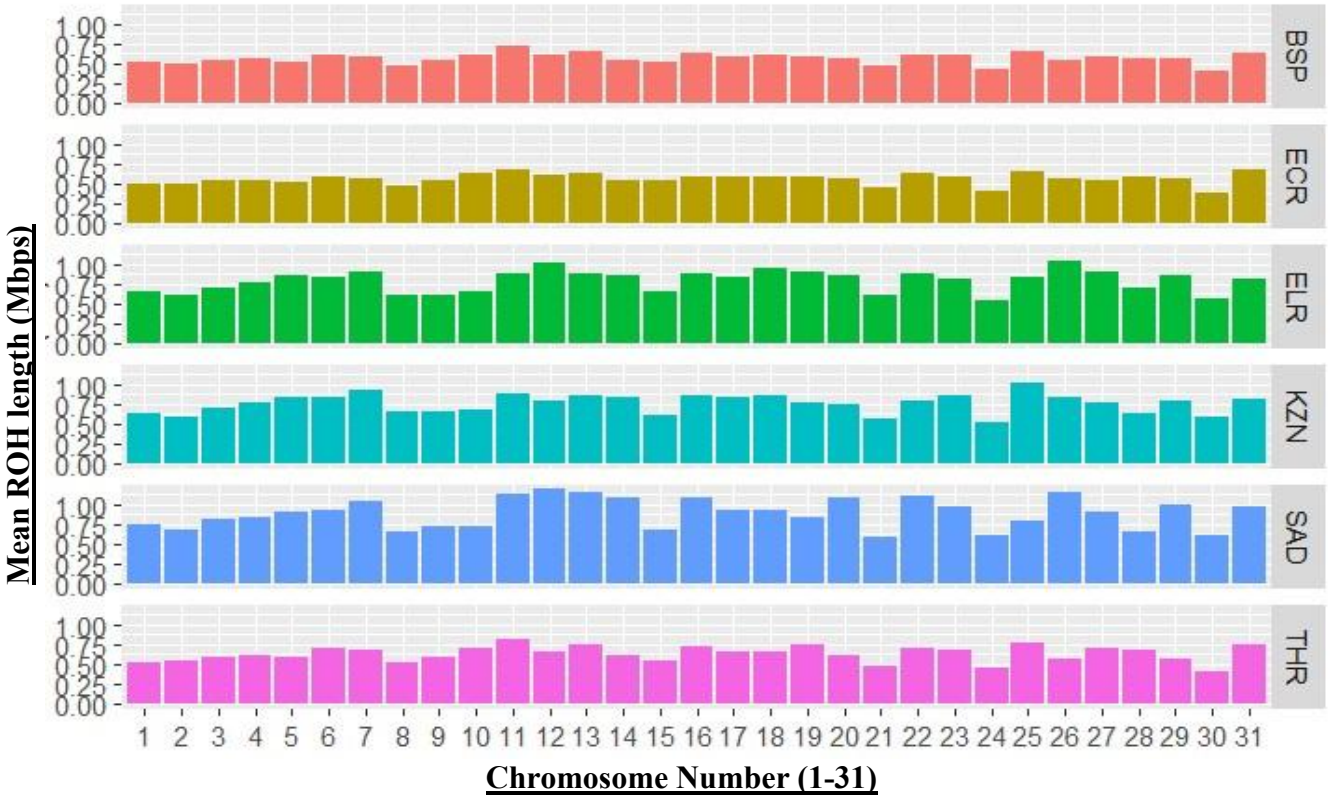


Figure 4.7 Mean length (Mbps) Runs of Homozygosity (ROH) per chromosome

[The mean length (Mbps) of Runs of homozygosity (ROH) stretches detected on autosomes 1-31 was determined for each population of the representative breeds of purebred commercial, indigenous and rural horses.]

As shown in **Figure 4.7**, each representative breed population showed significantly higher and moderately lower mean ROH lengths (Mbps) on specific chromosome numbers, as compared to the standard length of the remaining autosomes within the dataset. The highest mean ROH lengths within the population dataset of 1.190Mbps was detected on chromosome 12, within the SA Saddlebred (SAD) breed population, followed by a mean ROH length of 1.14Mbps found on chromosome 26 – belonging to the same purebred, commercial representative population. Significantly shorter mean ROH lengths of 0.380 Mbps and 0.410 Mbps, were detected exclusively within the Eastern Cape Rural (ECR) representative breed population in comparison to the rest of the population dataset, on chromosomes 30 and 24 - respectively. Both the KwaZulu-Natal (KZN) (1.020 Mbps (chr 25) – 0.530 Mbps (chr 24)) and East London Rural (ELR) (1.040 Mbps (chr 26) – 0.530 Mbps (chr 24)) rural breed populations showed a vast range of mean ROH lengths across the autosomes. Through examination of all homozygous stretches detected within the autosomes collectively, for each breed population, the Basotho Pony (BSP) and Thoroughbred (THR) populations showed the most consistent, moderate ROH lengths throughout chromosomes 1 to 31. The SA Saddlebred (SAD), East London Rural (ELR) and KwaZulu-Natal Rural (KZN) breed populations contained significantly longer ROH stretches (Mbps), as compared to the remaining populations.

Based on the number and lengths of homozygous stretches found on both a population and chromosomal-level within the population dataset, ROH islands/peaks were determined by considering only the top 30 percentile of runs within the six representative breed populations – taking into consideration that SNPs found in those ROH congregates were shared with more than 70% of the population dataset.

Table 4.3 ROH islands/peaks detected among the six representative breed populations

[The top 30 percentile (threshold = 0.700) Runs of Homozygosity (ROH) were identified within the representative breed populations, with the inclusion of the highest proportion of SNPs, found within the ROH peaks.]

Population/Breed	Chromosome Number	nROH Islands/Peaks
ELR	9	1
KZN	3, 10	2
SAD	1, 2, 3, 4, 6, 8, 9, 14, 16, 18, 20, 21, 23, 25, 30,	27
THR	1, 3, 4, 6, 11, 13, 14, 16, 17, 18, 19, 22, 28, 31	18

[**Table 4.3** contains the following information based off of the results from the top 30 percentile of runs: **Population/Breed** = Breed populations outlined within this study; **Chromosome number** = the chromosome in which the RUN is found; **nROH/Peaks Islands** = the numbers of ROH islands/peaks found on the specified chromsomes within each population.]

The top ROH islands/peaks were detected in four of the six breed populations within this study, including East London Rural (ELR), KwaZulu-Natal (KZN) rural breed populations; SA Saddlebred (SAD) and Thoroughbred (THR) commercial breed populations. The most significant number of ROH islands/peaks that were detected within the SA Saddlebred (SAD) population – with a total of 27 ROH islands. The Thoroughbred (THR) followed behind with a total of 18 detected ROH islands/peaks detected. These significant top runs may indicate important selection signatures that are inherited from generations, which may be of importance to the genetic architecture of each specific breed population.

To further assess the degree of genetic diversity among the representative breed populations, ROH-based inbreeding levels were determined in order to compare the breed populations, on both a collective genome-wide population- (**Figure 4.8**) and chromosomal-level (**Figure 4.9**).

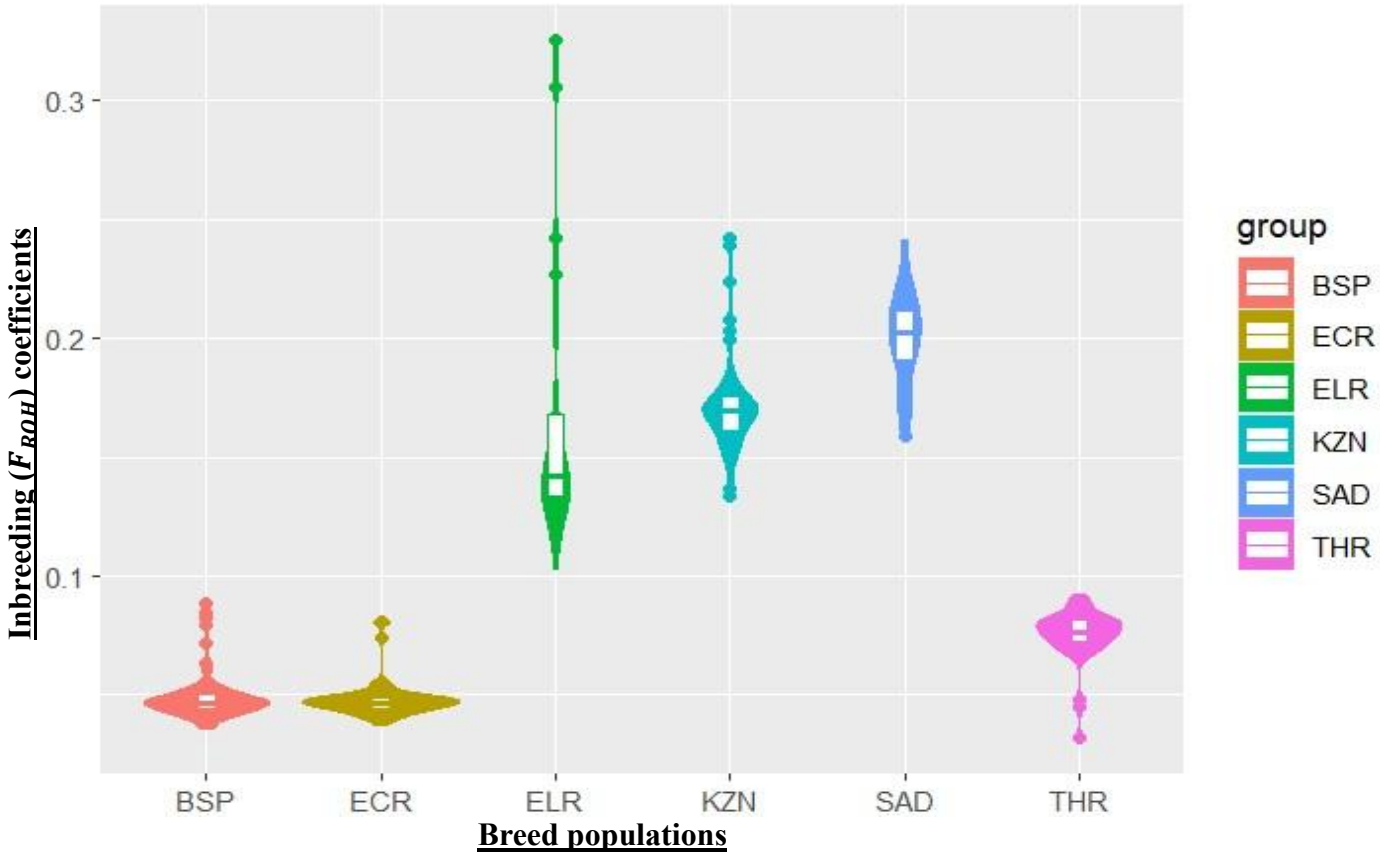


Figure 4.8 Genome-wide population inbreeding levels coefficients (F_{ROH}) based on ROH stretches

[The inbreeding coefficients (F_{ROH}) were estimated for the six representative populations of the commercial, indigenous and rural horse breeds - based on the equation proposed by McQuillan et al. (2008) that determines inbreeding levels based on the length of homozygous stretches in comparison to total genome length.]

As seen in **Figure 4.8**, the six representative breed populations exhibited a large range of low to high levels of inbreeding, ranging from $0.032 \leq F_{ROH} \leq 0.326$. A large proportion of samples within the SA Saddlebred (SAD) and KwaZulu-Natal (KZN) breed populations exhibited a higher range of inbreeding levels out of the remaining dataset, ranging from F_{ROH} = 0.133 – 0.242 and F_{ROH} = 0.169 – 0.239, respectively. The East London Rural (ELR) breed population showed the largest range of inbreeding levels throughout the population, ranging from F_{ROH} = 0.102 – 0.326, providing insight into the variation within the population. Collectively, the indigenous Basotho Pony (BSP), Eastern Cape Rural (ECR) and the purebred, commercial Thoroughbred (THR) populations that exhibited a lower range of inbreeding $F_{ROH} < 0.100$.

The above genome-wide population inbreeding level comparison was further defined on a chromosomal-level, in order to classify possible identical by descent (IBD) allele positions, in **Figure 4.9**.

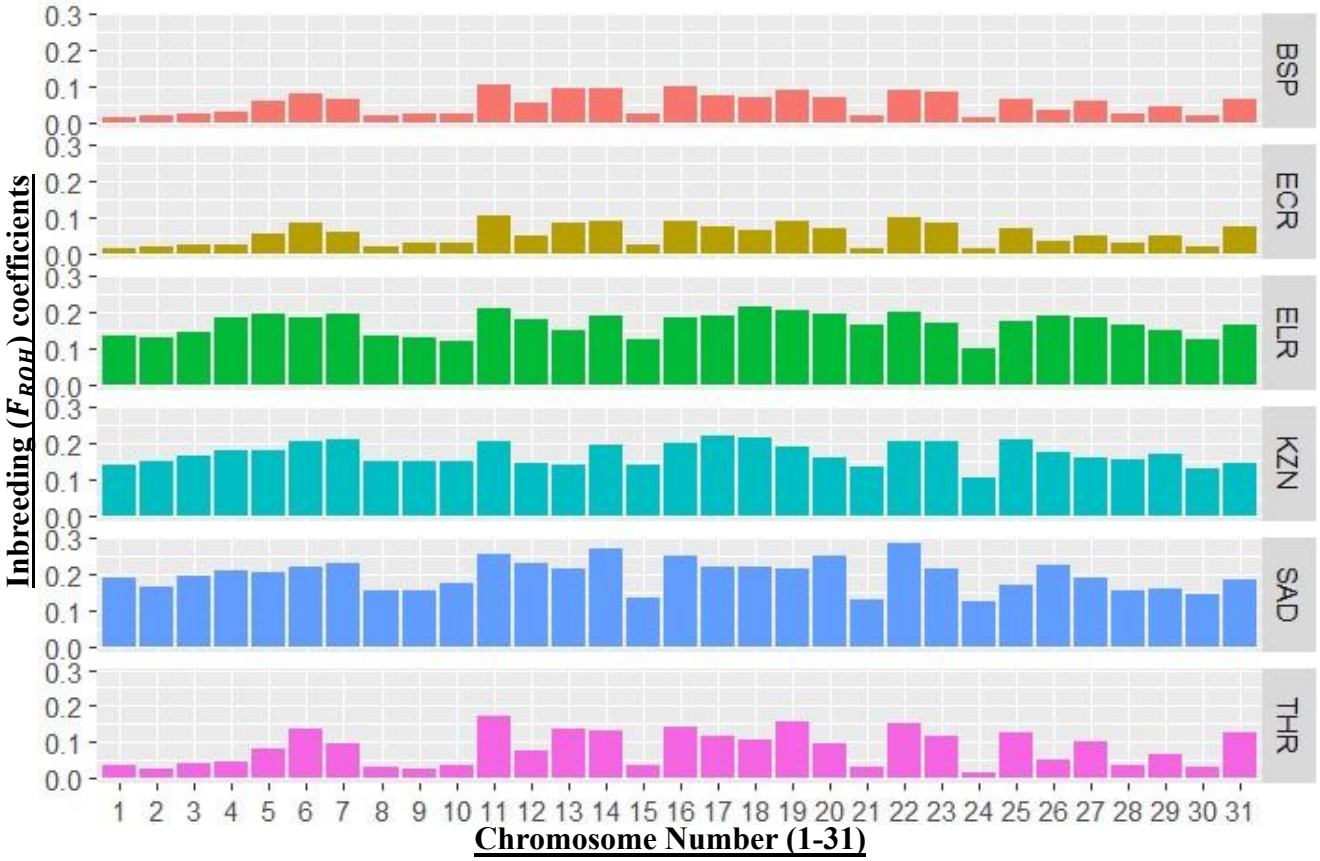


Figure 4.9 Chromosomal-level population inbreeding coefficients (F_{ROH}) based on ROH stretches

[The inbreeding coefficients (F_{ROH}) were estimated for the six representative populations of the commercial, indigenous and rural horse breeds – comparing possible identical by descent (IBD) allele positions on the 31 autosomes.]

The highest estimated inbreeding level (F_{ROH} = 0.280) (**Figure 4.9**) within the entire population dataset, was found on chromosome 22 of the SA Saddlebred (SAD) representative population, with the addition of F_{ROH} = 0.250 estimated on chromosome 14 within the same breed population, taking second place. The distribution of inbreeding levels proved to be significantly low within the Basotho Pony (BSP) and Eastern Cape Rural (ECR) breed populations, spanning across the autosomes. The significantly lowest inbreeding levels (F_{ROH} < 0.05) within the population dataset was found in the Basotho Pony (BSP) (on chromosomes 1; 2; 3; 4; 8; 9; 10; 15; 21; 24; 26; 28 and 30), Eastern Cape Rural (ECR) (on chromosomes 1; 2; 3; 4; 8; 9; 10; 15; 21; 24; 26; 28; 30) and the Thoroughbred (THR) (on chromosomes 1; 2; 3; 4; 8; 9; 10; 15; 21; 24; 28; 30) breed populations. Through the comparison of collective chromosomal F_{ROH} throughout the population dataset, the SA Saddlebred (SAD) breed population exhibited the highest consistent range of inbreeding levels autosomal-wide, followed closely by the KwaZulu-Natal (KZN) and East London Rural (ELR) breed populations.

4.5 Discussion

The investigation of the genetic diversity and population genetic architecture of these indigenous, rural and purebred commercial South African horse breeds involved within this study, enables the characterization of the evolutionary operations and influence of ecological inputs (Shen et al., 2019), that shaped these modern horse breed populations and their ability to adapt. The development of the Equine SNP70K BeadChip shortly after the release of the SNP50 BeadChip design (Schaefer and McCue, 2020), designed using a higher density

array of mapped SNP markers from 7 horse breed representatives detected using low-coverage Sanger sequences, which included an Arabian, Akhal-Teke, Andalusian, Icelandic, Quarter Horse, Standardbred and Thoroughbred (McCue et al., 2012). The release of this Equine SNP70K BeadChip has made a major impact in the high-throughput SNP processing of genomic studies conducted on many exotic, purebred horse populations including Arabian breeds (Al Abri et al., 2017), Thoroughbreds (Han et al., 2020) and Standardbreds (McCoy et al., 2019) - collectively focusing on characteristic mapped traits of the specific breeds and possible improvement. However, the utilization of the Equine SNP70K and Equine SNP80K BeadChip in this genomics assisted, community-driven research study aims to characterize the adaptive genetic and functional diversity of the rural and indigenous horse populations within South Africa and highlights probable selection events that have occurred throughout the population's demographic history.

The initial insight into the state of the population structure and differentiation of these representative breed populations was determined in **Figure 4.2**, whereby a Principal Component Analysis (PCA) yielded 4 distinct genetic clusters within the covariance matrix. Cluster 1 exclusively contained Thoroughbred (THR) breed samples that clustered away from the remaining dataset, which was expected as this purebred, commercial horse breed became a closed breeding stock (McGivney et al., 2020), due to the establishment of the Jockey Club of South Africa organization that prevented outbreeding of registered Thoroughbreds with any other horse breed (Herringer, 2004). This closed-system of breeding has accelerated selection pressures characteristic of this sport horse and limited gene pool availability (Todd et al., 2020), as indicated by the uniform ancestral proportions in the sNMF plot in **Figure 4.4** above, with minor admixed, ancestral proportion influence of known three ancestral precursor breeds (namely the Arabian breed, Barb breed and the Akhal-Teke breed (Cunningham et al., 2001)) indicated within the analysis output at $K = 4$. The additional purebred, commercial SA Saddlebred (SAD) population exclusively clustered within Cluster 4, which was spatially situated next to the indigenous and rural population Cluster 2, indicating that although this commercial breed is recognized as a distinct purebred in South Africa (Millin, 2011) – the results within **Figure 4.2** and **Figure 4.4** provides evidence that there may be a degree of genetic relatedness between these clusters. This observation may be due to the fact that American Saddlebred stallions were imported into South Africa dating back to 1917, whereby these imports were bred to indigenous South African mares in order to produce locally-adapted progeny of specific bloodlines – genetically preserved by the dedicated Saddlebred Society of South Africa established in 1949 (Swart, 2008).

Cluster 2 contained entirety of indigenous Basotho pony (BSP) representative samples and two of the rural breed populations, namely the Eastern Cape Rural (ECR) and KwaZulu-Natal (KZN) populations, indicating that these three breed samples share common allele frequencies maintained within the population as well as common ancestral proportions (Shriner, 2013) indicated within **Figure 4.4** at cluster $K = 4$. The FAO (Annexure 2, 2001) documents the animal genetic resources in developing countries whereby "locally adapted" (or rural, non-descript) and "indigenous" animal breeds or groups are sourced from one another, with variation within those ecotypes, caused by adaption to particular geographical locations. In very close proximity to Cluster 2 in **Figure 4.2**, the East London Rural (ELR) breed population formed its own cluster (Cluster 3),

supporting variation of this ecotype adapted to a different geographical location even though this population exhibited very similar ancestral proportions in **Figure 4.4** to the remaining rural breed populations. This documentation provides evidence that the locally-adapted, Eastern Cape Rural (ECR) and KwaZulu-Natal (KZN) sampled horse populations would indeed share very common adaptive genetic influence with the designated indigenous Basotho pony (BSP) in terms of ancestry and the continuous gene flow between the populations. This is due to the close geographical location of the Eastern Cape and KwaZulu-Natal Provinces that border the country of Lesotho and the well-documented migration pattern of Lesotho people into South Africa (Mokoena and Balkaran, 2018).

Ancestral proportions within **Figure 4.4** show that all six representative commercial, indigenous and rural horse breed populations share common ancestry, due to the fact that the Thoroughbred (THR) breed was a known precursor breed used in two main selection events of indigenous horse population breed improvement within South Africa. This includes the original development of the common indigenous ancestor known as the Cape horse (Cothran and Van Dyk, 1998) and the Basotho Pony breed improvement project formed in 1977 to prevent extinction of the Basotho pony (McCormack, 2018). The results produced from the population-based pairwise Fixation index (F_{st}) analysis in **Table 4.1** and represented in **Figure 4.5**, showed that differentiation values between the purebred, commercial Thoroughbred (THR) population and the indigenous and rural breed populations were below the significant differentiating threshold of $F_{st} \geq 0.150$ (Frankham et al., 2002) – indicating high genetic similarity between all five breed populations. It was expected that the Thoroughbred (THR) representative population would have significant population differentiation between the indigenous and rural populations, as this sport breed has a documented closed breeding pool since 1906 and would not be expected to have much interaction with indigenous and rural horse breeds today. Nonetheless, there are several online articles that depict the fate of race horses whose racing careers have ended or perhaps have not succeeded within the field. Some include publications made by Pinnock (2022) and Pijoos (2018), where these Thoroughbreds are either sold or given to rural communities that partake in rural racing events and therefore increase the occurrence of gene flow between these breed populations. What was expected was the differentiation *Fst* value of 0.146 in **Table 4.1** between the two commercial, purebred populations (THR and SAD), due to the fact that these two defined breeds have vastly different athletic selection criteria in order to dominate the equine sporting world within their respective fields. Thoroughbreds are bred for speed (Bower et al., 2012) and Saddlebreds are bred for gaited-showing (Vincelette, 2013).

The close genetic-relatedness between these South African horse breeds was exhibited within the global PCA in **Figure 4.3**, whereby these commercial, indigenous and rural South African breeds all clustered together in Cluster 4 in comparison to continental horse breeds sampled by Petersen et al. (2013). The cluster of South African horse breeds distributed within the middle of the matrix, closely surrounded by representative horse breeds of North America (Cluster 1), Asian (Cluster 2) South American (Cluster 3) and Europe (Cluster 5 and 6). This draws to the collective genetic contribution that all these global breeds had in the development of the South African horse breeds included within this study (Marlow, 2010). Within the South African horse breed cluster, two Asian horse breeds (Tuva and Mongolian) that Pozharskiy et al. (2023) identified as the closest

native breeds of the Kazakh horse group that were used as genetic stock resources for warhorses, owing probable cause of their genetic influence within South African horse breeds developed from warhorse founding populations introduced into South Africa. One European native horse breed (New Forest Pony) also clustered within the South African horse cluster, whereby this breed is characterized by their strength and hardiness that contributed to their capabilities as a warhorse that Falkingham (2018) reported to be used by Forest scouts during the Boer War. This provides the likelihood of surviving New Forest Pony warhorses used to breed with the South African warhorses for genetic improvement. To further explain the clustering of the South African horse breeds together in the global PCA (**Figure 4.3**), an AMOVA (**Table 4.2**) was conducted whereby a statistically low *F*-statistical among population (F_{CT}) = 4.430% parameter was estimated between the six representative breed populations. A further intra-population parameter (F_{IS}) of the AMOVA yielded a collectively high genetic variation F_{IS} = 94.390% within the six representative breed populations, indicating that although there was continuous determined gene flow among the breed populations, individuals within the populations maintained a high standard of genetic diversity (Pongratz et al., 2002) – with which was examined in closer detail.

Runs of Homozygosity (ROH) provide insight into the state of the genetic diversity maintained within populations, due to the fact that these homozygous stretches along the genome accumulate due to a few factors including consanguineous mating, genetic drift as well as selection events (Curik et al., 2014) – ultimately reducing genetic variability within populations (Pacheco et al., 2023). The lengths of these homozygous stretches also provide insight into selection events that have occurred and affected populations, whether it be due to an ancestral inheritance of homozygous genotypes or due to recent selection events that are the result from contiguous mating (Dixit et al., 2020). A summation of the total number of ROH stretches compared to the total length (Mbps) of those stretches, was detected spanning across the six representative breed populations, illustrated within **Figure 4.6**. The Basotho pony (BSP), Eastern Cape Rural (ECR) and Thoroughbred (THR) breed populations exhibited homozygous stretches ranging from 120 to 315 runs detected – all consisting of a shorter base length of <300Mbps. In the indigenous and rural breed populations, the low frequency and shorter length of homozygous stretches likely result from inheritance patterns shaped by ancient contiguous genomic regions. Over the evolutionary history of these closely related populations, recombination events, driven primarily by adaptive responses to environmental pressures, would have progressively shortened these runs of homozygosity (ROH) spanning the genome (McQuillan et al., 2008). Todd et al. (2018) highlights the fact that the Thoroughbred breed as a whole has experienced over 300 years of selective breeding but is one of the largest closed-breeding populations recorded – ultimately reducing the need for recent contiguous mating to maintain breed specific selection signatures within the population. This trend of shorter spans of ROH stretches of the Basotho pony (BSP), Eastern Cape Rural (ECR) and Thoroughbred (THR) breed populations can also be observed within **Figure 4.7**. Moderate to shorter stretches were maintained throughout the autosomes and the shortest lengths (0.380 Mbps on chromosome 30 and 0.410 Mbps on chromosome 24) out of the dataset, were detected within the Eastern Cape Rural (ECR) population.

The East London Rural (ELR) and KwaZulu-Natal (KZN) populations showed a large range of accumulative ROH lengths represented in **Figure 4.6**, spanning from 300 to 800 Mbps in length, which may be owed to recent contiguous mating events within these rural populations, or may also be due to peculiar allele recombination rates or mutations – suggested by Mastrangelo et al. (2016). The SA Saddlebred (SA) commercial breed also exhibited a higher range of accumulative ROH stretches within the population (**Figure 4.6**) and the highest detected mean ROH length of 1.190 Mbps on chromosome 12 within the dataset (**Figure 4.7**). This was expected due to the intense selective pressure placed on this commercial breed (Chen et al., 2023) and the small founding population of Saddlebreds within South Africa. These short and long ROH stretches detected within each population were further analysed for selected genomic regions that contain a high proportion of these homozygous stretches and SNP accumulation known as ROH islands/peaks (He et al., 2020). The identification of these ROH islands/peaks can reveal strong adaptive selection signatures within the population (Peripolli et al., 2018), or reveal the accumulation of deleterious recessive alleles within a population, that may affect survival and performance of each individual (Stoffel et al., 2021). Therefore, a genome-wide scan was performed to identify the top 30th percentile of ROH islands/peaks within the representative breed populations, based on the accumulation of SNPs present in more than 70% of the population dataset. ROH islands were detected in the KwaZulu-Natal (KZN), East London Rural (ELR), SA Saddlebred (SAD), and Thoroughbred (THR) populations, as shown in **Table 4.3**. A QTL analysis represented in **Table 4.4** below was utilized to define selection signatures found within the ROH islands.

Table 4.4 QTL analysis of the top 30 percentile ROH

[Most significant SNPs found within top 30 percentile ROH islands/peaks using *Ensembl* (https://www.ensembl.org).]

Population/ Breed	Chromosome Number	SNP Start position	SNP End position	Gene name	Function
ELR	9	35240634	35806600	*SNAI2*	Facilitates cartilage development *in vitro* (Chen and Gridley, 2013)
KZN	10	25070138	25528489	*ZNF865*	Regulates key cellular and molecular processes (Hussain et al., 2022)
				ZNF524	
				ZNF579	
				ZNF580	
				ZNF784	
				SBK2	Cardiac muscle organization (Van Gorp et al., 2022)
				SBK3	
				FIZ1	Plays a role in immunity (Wolf and Rohrschneider, 1999)
				EPN1	Growth performance trait (Wang et al., 2024)
				SSC5D	Plays a role in immune response (Plawinska-Czarnak et al., 2021)
				NLRP4	Plays a role in reproductive cell development (Tian et al., 2009)
				NLRP13	
				NLRP5	
KZN	3	19391644	19733671	*ZFP90*	Plays a role in immune response (Pasandideh et al., 2020)
				CDH3	Maintenance of cellular differentiation (Jose et al., 2023)

				CDH1	
SAD	1	41390317	41686476	*SGMS1*	Lipid metabolism (Ablondi et al., 2019)
				ASAH2	Lipid metabolism (Pozharskiy et al., 2023)
SAD	1	61426908	61453199	*ECD*	Plays a role in oogenesis (Kim et al., 2009)
				NUDT13	Cellular metabolism (Abdelraheim et al., 2017)
SAD	1	105139300	105289393	*IGF1R*	Regulates growth hormone levels (Kullmann et al., 2015)
				PGPEP1	Peptidase enzyme (Sigala et al., 2008)
SAD	14	10799258	10831801	*SPDL1*	Cell cycle regulation (Kodama et al., 2019)
				SLIT3	Myogenic differentiation (Zhang et al., 2009)
SAD	16	25440176	26158665	*MAGI1*	Cresty neck trait (Gmel et al., 2023)
				ADAMTS9	Reproductive traits (Hernandez-Delgado et al., 2023)
SAD	16	47827828	47977783	*TTC21A*	Sperm flagella formation (Aitken and Baker, 2020)
SAD	18	76632555	77187786	*C2CD6*	Sperm functionality (Dahan and Breitbart, 2022)
				ALS2	Motor neuron function (Daud et al., 2018)
SAD	20	50508710	50702540	*METTL8*	Skeletal muscle formation (Li et al., 2022)
SAD	21	27618930	27673548	*RICTOR*	Cell proliferation (Jebali and Dumaz, 2018)
				OSMR	Reduces inflammation (Xiao et al., 2011)
SAD	23	34744830	35273875	*TTC39B*	Plays a role in metabolism (Loux et al., 2022)
				PSIP1	Plays a role in immunity of viral infections (Marshall et al., 2007)
SAD	23	35818866	36059614	*BNC2*	Involved in spermatogenesis (Vanhoutteghem et al., 2016)
SAD	25	26643810	26805046	*OR1J1E*	Olfactory receptors (Olender et al., 2020)
				OR1J38B	
				OR1J45B	
				OR1J42	
				OR1J49B	
				OR1J46B	
SAD	3	35953765	38257938	*CDH15*	Plays a role in muscle development (Lee et al., 2020)
				CBFA2T3	Regulates lipid metabolism (Kim et al., 2024)
				NFKB1	Plays a role in immunity (Grilz-Seger et al., 2019)
				DPEP1	Plays a role in renal function and metabolic regulation (Choudhury et al., 2019)
				MC1R	Coat colour variation (Rosa et al., 2022)
				CDT1	Supports cell proliferation and genomic stability (Tsuyama et al., 2005)
				ACSF3	Associated with intramuscular fat deposits (He et al., 2022)
				SPG7	Linked to hereditary neurodegenerative disorders (Mancuso et al., 2012)
				PIEZ01	Regulates cardiac cycle under stress (Yuan et al., 2023)
				DBNDD1	Plays a role in neurological development (Worth et al., 2013)

				SPATA2L	Plays a role in regulating spermatogenesis (Zhao et al., 2017)
SAD	3	73752453	74085693	*ADGRL3*	Plays a role is neurological functioning and mutations may result in Equine Degenerative Myeloencephalopathy (EDM) (Posbergh et al., 2018)
SAD	30	7950810	8196894	*ZFP82* *ZNF793* *ZNF568* *ZNF829*	Regulates key cellular and molecular processes (Hussain et al., 2022)
SAD	4	24043651	24425079	*VSTM2A*	Involved within adipose tissue formation (Secco et al., 2017)
SAD	4	26715528	27169753	*CACNA2D1*	Muscle contraction regulation (Valberg et al., 2022)
SAD	4	34386588	34398127	*STEAP2*	Driver of cell proliferation resulting in prostate cancer (Xing et al., 2023)
				CFAP69	Promotes sperm motility (Al Abri et al., 2023)
SAD	6	41068825	41097641	*BCL2L14*	Programmed cell death of abnormal cells during oogenesis (Cao et al., 2018)
THR	1	113163157	113583818	*GABRA3* *GABRA5*	Behavioural associated traits, reduces stress (Holtby et al., 2023)
THR	13	40972223	41296615	*TNFRSF12*	Plays a role in reducing inflammation (Liao et al., 2023)
THR	14	41593213	42080012	*PAQR4*	Plays a role in mare fertility (Laseca et al., 2022)
				GDF9	Growth factor involved in ovarian follicle development and oocyte maturation (Stefaniuk-Szmukier et al., 2018)
THR	14	42327252	43102922	*IL5*	Plays a role in immune response and forms the basis of the Interleukin-5 active vaccine against insect-bit hypersensitivity (IBH) that disrupts the inflammatory response caused by this gene (Fettelschoss-Gabriel et al., 2019)
				IRF1	Plays a role in antiviral response (Kim et al., 2021)
				P4HA2	Connective tissue development (Wilhelm et al., 2023)
				CSF2	Involved in immune response (Karagianni et al., 2020)
				ACSL6	Plays a role in lipid metabolism (Lai et al., 2020)
THR	16	50088514	50517281	*STAC3*	Involved in skeletal muscle development and contraction (Zhang et al., 2014)
				ARPP21	Associated with tissue development and body size (Toro-Ospina, et al., 2024)
THR	17	20542273	21823211	*TRIM13*	Associated with muscle development (Dementieva et al., 2023)
				CAB39L	Muscle energy metabolism (Beltran et al., 2015)
THR	18	29815707	30381124	*ACVR2A*	Promotes follicular growth (da Silveira et al., 2014)
THR	22	4377551	4948733	*INSM1*	Regulates glucose metabolism (Moon et al., 2015)
				CFAP61	Facilitates flagella movement (Metzger et al., 2015)
THR	6	22000722	22688929	*GBX2*	Involved in anterior hindbrain development (Burroughs-Garcia, J., 2011)
	6	28910075	29360023	*ADA2*	Plays a role in immunity (Contreras-Aguilar et al., 2020)
				IL17RA	Drive inflammatory response that proliferates tissue damage caused by Supporting Limb Laminitis (SLL) (Cassimeris et al., 2020)

[Most significant SNPs found within top 30 percentile ROH islands/peaks using *Ensembl* (https://www.ensembl.org).]

One ROH island was detected within the East London Rural (ELR) breed population on chromosome 9 that contained an important selection signature (*SNAI2*) for cartilage growth facilitation. This is imperative for horses to protect the joints during surging growth into adulthood and to constantly absorb shock stresses

applied to joints of taxing movement being placed on them (Brama et al., 2000). The second rural breed population in which selection signatures were detected within ROH islands, were in the KwaZulu-Natal (KZN) breed population, on chromosomes 10 and 3. This breed population contained many selection signatures involved within cellular and molecular processes (*ZNF865*; *ZNF524*; *ZNF579*; *ZNF580*; *ZNF784*; *CDH3*; *CDH1*), cardiac muscle formation (*SBK2*; *SBK3*), immunity (*FIZ1*; *SSC5D*; *ZFP90*), growth traits (*EPN1*) and reproductive traits (*NLRP13*; *NLRP5*). Alam et al. (2016) highlights the importance of growth and reproductive traits within locally adapted horse populations that ensures the adaptive ability of the population within a changing environment. Immune-related genes provide insight into the pathogen resistant capability of a horse population that is exposed to harsh environmental conditions (Metzger et al., 2014) that may not have been exposed to vaccines or medication.

Within the commercial, purebred SA Saddlebred (SAD) population, there were a total of 27 ROH islands/peaks detected within the breed populations, namely on chromosomes 1, 3, 4, 6, 14, 16, 18, 20, 21, 23, 25 and 30 – as seen in **Table 4.4**. The occurrence of several metabolism genes (*SGMS1*; *ASAH2*; *TTC39B*; *CBFA2T3*; *DPEP)* within this commercial breed ensures efficient energy conversion of high-protein and high-lipid concentrated feed, in order to prevent dietary-related diseases such as equine obesity, laminitis, Equine Metabolic Syndrome (EMS) and pituitary pars intermedia dysfunction (Lewis et al., 2017). Equally as important within sporting equine breeds, fertility and reproductive traits (*ECD*; *ADAMTS9*; *TTC21A*; *C2CD6*; *BNC2*; *SPATA2L; CFAP69*; *BCL2L14*) are crucial for the continuation of bloodlines within the population that defines the economic value of the stud farms (Perdomo-Gonzalez et al., 2023). An interesting ROH island detected on chromosome 25 within the SA Saddlebred (SAD) harbours olfactory receptor genes (*OR1J1E*; *OR1J38B*; *OR1J45B*; *OR1J42*; *OR1J49B; OR1J46B*). Rorvang et al. (2020) indicates sensory and perception ability is closely linked to behaviour in horses that is important within human-to-horse relationships whereby showing equine sport relies heavily on trust between the horse and the rider. Characteristic traits of athletically selected commercial breeds include muscle formation (*METTL8*; *CDH15*; *CACNA2D1*), correct intramuscular fat deposits (*ACSF3*; *VSTM2A*), growth traits (*IGF1R*) and regulation of cardiac cycle (*SLIT3*; *PIEZ01*). These traits are essential in order for the horse's physiological adaption to training loads specific to the sport discipline, promoting stamina and maximizing oxygen consumption (de Meeus d'Argebteuil et al., 2021). The SA Saddlebred (SAD) is a gait-selected sporting horse, whereby high-arched long, slim necks are mandatory for their breed description (Staiger et al., 2016) and therefore a well-established gene found on chromosome 16 within the equine genome has been selected to produce this morphological trait. The *MAGI1* gene has been previously associated with the 'cresty-neck' shape observed within lighter-framed horses, depicting fat deposits made on the dorsal segment of the neck that although may be a desirable trait, the presence of this gene within homozygous regions details risk factors for metabolic disorders within the population (Gmel et al., 2023). Another degenerative, heritable gene identified within the SA Saddlebred (SAD) population is the *ADGRL3* gene in which studies such as Posbergh et al. (2018) have associated mutations within this gene to cause Equine Degenerative Myeloencephalopathy (EDM) – a disorder that causes loss of muscle control in limbs of younger horses that can be detrimental to the development of the horse.

Comparably, the commercial Thoroughbred (THR) breed population also contained beneficial selection signatures for reproduction and fertility traits (*PAQR4*; *GDF9*; *ACVR2A*; *CFAP61*), muscle development (STAC3; *TRIM13*; *CAB39L*) and lipid metabolism (*ACSL6*) – as to the commercial SA Saddlebred (SAD) breed population. Holtby et al. (2023) describes the incorporation of behavioural traits (*GABRB3* and *GABRA5*), detected on chromosome 1, as beneficial for the Thoroughbred breed that undergoes intense training programs within the horse racing industry. The constant involvement in a high-stress environment, whereby the occurrence of these genes that have a very high chance of being inherited within the next generation, enables soundness of mind for adaption within this environment (McBride and Mills, 2012). What is most interesting about the selection signatures carried within the homozygous proportions of the Thoroughbred (THR) population are associated with traits that promote immune response to various conditions that may affect the population. The *IL5* gene was detected on chromosome 14 in **Table 4.4**, that is involved in immune response mechanisms responsible for protecting the horse against insect bites of the *Culicoides* spp., that causes an allergic skin disease known as Insect bite hypersensitivity (IBH) (Langreder et al., 2023). Although this inflammatory response acts to protect the horse against the protein components within the insect's salvia, this immune response is what causes IBH and can be detrimental to the horse. Fettelschoss-Gabriel et al. (2019) produced an active vaccination that reduces the inflammatory response promoted by the *IL5* gene towards insect bites, which ultimately reduces IBH disease symptoms and may provide long-term treatment for affected horses. A second inflammatory-promoter gene (*IL17RA*), which was detected within the Thoroughbred (THR) population, has been documented to accelerate tissue damage caused by Supporting Limb Laminitis (SLL) that causes debilitating pain within the tissue layers of the hoof lamellae – often resulting in horses being euthanized (Cassimeris et al., 2020). However, many researchers such as Moss et al. (2020), have initiated studies that focus on inhibiting the inflammatory response of the *IL17RA* that have been successful in humans and could be applied to horses. Although these selection signatures associated with immune response within the Thoroughbred (THR) population have been proven to accelerate detrimental diseases within horses, researchers have been dedicated to alleviate inflammatory responses initiated by these genes, to ensure the optimum performances of this economically important horse breed.

McQuillan et al. (2008) proposed the use of total genomic ROH length within a population, compared to the total genome length, in order to determine inbreeding levels (F_{ROH}), as a major indicator of genetic diversity within populations. The use of inbreeding levels based on homozygous run lengths, provides a broader and more accurate determination of autozygosity within a population, compared to the out-dated pedigree information based on calculated estimates (Forutan et al., 2018). **Figure 4.8**, depicts a large range of population inbreeding (F_{ROH}) levels within the East London Rural (ELR) breed population, ranging from 0.102 to 0.326 – indicating variation within the individuals of the population. Szwaczkowski et al. (2016) reports on the fact that uncontrolled mating of rural/ecotype horses within a population ultimately leads increased inbreeding levels of closely related individuals as well as a surge of heterosis introduced within the population – providing an explanation for variation seen within the East London Rural (ELR) population. Although this rural breed population contains individuals with high inbreeding levels, also depicted in higher inbreeding levels across the autosomes of this breed in **Figure 4.9**, there were no deleterious selection signatures identified within the

top ROH island seen in **Table 4.4** that could pose a threat to the overall fitness of the rural population. The very same outcome was made regarding the KwaZulu-Natal (KZN) rural population that collectively exhibited moderate population inbreeding levels (F_{ROH} ranging from 0.169 to 0.239) (Peripolli et al., 2016) within **Figure 4.8** and **Figure 4.9**. These two ROH islands detected within this population contain important adaptive traits that ensures the reproductive and survival ability of the rural breed (de Simoni Gouveia et al., 2014). The indigenous Basotho Pony (BSP) population exhibited very low inbreeding results on a population level ($F_{ROH} <$ 0.100), indicating that recent inbreeding has not occurred within the population and therefore evidence of foundational genetic diversity has been maintained (Chen et al., 2023).

As expected, the commercial SA Saddlebred (SAD) population exhibited the highest range of population inbreeding levels (F_{ROH} = 0.133 – 0.242) within **Figure 4.8** and high occurrences of inbreeding levels across the autosomes shown in **Figure 4.9** – showing specific inbreeding inflation on particular autosomes observed in **Table 4.4**. ROH islands/peaks detected on chromosome 16 (F_{ROH}= 0.250) and chromosome 3 (F_{ROH}= 0.200) each carried autosomal, deleterious genes (*MAGI1* and *ADGRL3*) that have a higher chance of inheritance within subsequent generations due to the increased recent inbreeding levels caused by intense selection practices – ultimately reducing the overall fitness and phenotypic performance of this commercial breed population (Bosse et al., 2019). A very significant outcome within **Figure 4.8** and **Figure 4.9** was the low levels of population inbreeding (F_{ROH} < 0.100), collectively accumulated within the sampled individuals of this commercial purebred population. This is particularly beneficial in this genetically diverse purebred population, which has retained advantageous selection signatures (**Table 4.4**) despite the intense selective pressures imposed by the horse racing industry (Sumreddee et al., 2020; Swindell and Bouzat, 2006).

4.6 Conclusion

Overall, the utilization of the Equine SNP70K and SNP80K BeadChip within this genomics-assisted and community-driven study served to characterize the genetic and functional diversity of prominent indigenous and rural South African horse breeds, highlighting the importance of their population demography and evolution in South Africa's rich history. Collectively, all representative indigenous, rural, and commercial horse populations exhibited low inter-population diversity, highlighting the common ancestry shared among these South African breeds, while maintaining high intra-population diversity. The detection of ROH stretches within each population revealed recent and past selection events that have occurred within the population, resulting in high inbreeding levels detected within the East London Rural (ELR) and SA Saddlebred (SAD) breed populations and low levels maintained within the indigenous Basotho Pony (BSP) and Eastern Cape Rural (ECR) population. Although the East London Rural (ELR) breed population showed evidence of recent, contiguous mating patterns, which may have been due to uncontrolled breeding strategies, only results of adaptive selection signatures were detected within ROH islands/peaks. This was not the case for the commercial SA Saddlebred (SAD) population, which showed high levels of inbreeding and associated harmful autosomal genes that may negatively affect the overall fitness of the population. The commercial Thoroughbred (THR) population maintained low levels of inbreeding, preserving beneficial selection

signatures carried on their homozygous portions, despite the intensive selective pressures placed on the sporting breed. Future research should focus on genetically and functionally characterizing rural horse populations, on a countrywide basis, whereby variability assessments on all regional populations can be made.

4.7 References

Abldelraheim, S.R., Spiller, D.G., McLennan, A.G. (2017). Mouse *NUDT13* is a mitochondrial nudix hydrolase with NAD(P)H pyrophosphohydrolase activity. Protein Journal 36(5), 425-432.

Ablondi, M., Eriksson, S., Tetu, S., Sabbioni, A., Viklund, A., Mikko, S. (2019). Genomic divergence in Swedish warmblood horses selected for equestrian disciplines. Genes 10(12). Available: https://doi.org/10.3390/genes10120976 [accessed September 21, 2024].

Agafonkin, V. (2016). Leaflet for R. RPubs. Available: http://rstudio.github.io/leaflet/ [accessed May 17, 2025].

Aitken, R.J., Baker, M.A. (2020). The role of genetics and oxidative stress in the etiology of male infertility – a unifying hypothesis. Frontiers in Endocrinology 11. Available: https://doi.org/10.3389/fendo.2020.581838 [accessed September 22, 2024].

Al Abri, M., Alfoudari, A., Mohammad, Z., Almathen, F., Al-Marzooqi, W., Al-Hajri, S., Al-Amri, M., Bahbahani, H. (2023). Assessing genetic diversity and defining signatures of positive selection on the genome of dromedary camels from the southeast of the Arabian Peninsula. Frontiers in Veterinary Science 10. Available: https://doi.org/10.3389/fvets.2023.1296610 [accessed September 24, 2024].

Al Abri, M.A., Konig von Borstel, U., Strecker, V., Brooks, S.A. (2017). Application of genomic estimation methods of inbreeding and population structure in an Arabian horse herd. Journal of Hereditary 108(4), 361-368.

Alam, M.P., Bhuiyan, M.S.A., Bhuiyan, A.K.F.H. (2016). Morphology, morphometry and reproductive characteristics of indigenous horses in Bangladesh. Journal of Bioscience and Agriculture Research 11(2), 947-954.

Balague-Dobon, L., Caceres, A., Gonzalez, J.R. (2022). Fully exploiting SNP arrays: a systematic review on the tools to extract underlying genomic structure. Briefings in Bioinformatics 23(2). Available: https://doi.org/10.1093/bib/bbac043 [accessed October 05, 2023].

Beltran, N.A.R., Meira, C.T., de Oliveria, H.N., Pereira, G.L., Silva, J.A.V., da Mota, M. D.S., Curi, R.A. (2015). Prospection of genomic regions divergently selected in cutting line of Quarter Horse in relation to racing line. Livestock Science 174, 1-9.

Beyer, G. (2002). What was the Great Trek? The Collector. Available: https://www.thecollector.com/what-was-the-great-trek/ [accessed September 19, 2023].

Biscarini, F., Cozzi, P., Gaspa, G., Marras, G. (2019). detectRUNS: An R package to detect runs of homozygosity and heterozygosity in diploid genomes. Available: https://cran.r-project.org/web/packages/detectRUNS/vignettes/detectRUNS.vignette.html#:~:text=detectRUNS%20is%20a%20R%20package,humans%20(e.g.%20McQuillan%20et%20al. [accessed October 18, 2023].

Bosse, M., Megens, H.J., Derks, M.F.L., de Cara, A.M.R., Groenen, M.A.M. (2019). Deleterious alleles in the context of domestication, inbreeding, and selection. Evolutionary Applications 12(1), 6-17.

Bower, M.A., McGivney, B.A., Campana, M.G., Gu, J., Andersson, L.S., Barrett, E., Davis, C.R., Mikko, S., Stock, F., Voronkova, V., Bradley, D.G., Fahey, A.G., Lindgren, G., MacHugh, D.E., Sulimova, G., Hill, E.W. (2012). The genetic origin and history of speed in the Thoroughbred racehorse. Nature Communications 3. Available: https://doi.org/10.1038/ncomms1644 [accessed September 30, 2024].

Brama, P.A., Tekoppele, J.M., Barneveld, A., van Weeren, P.R. (2000). Functional adaptions of equine articular cartilage: the formation of regional biochemical characteristics up to age one year. Equine Veterinary Journal 32(3), 217-221.

Burroughs-Garcia, J., Sittaramane, V., Chandrassekhar, A., Waters, S.T. (2011). Evolutionarily conserved function of Gbx2 in anterior hindbrain development. Developmental Dynamics 240(4), 828-838.

Cao, X., Wang, X., Lu, L., Li, X., Di, R., He, X., Hu, W., Zeng, X., Liu, Q., Chu, M. (2018). Expression and functional analysis of the *BCL2*-associated agonist of cell death (BAD) gene in the sheep ovary during the reproductive cycle. Frontiers in Endocri1nology 9. Available: https://doi.org/10.3389/fendo.2018.00512 [accessed September 24, 2024].

Cassimeris, L., Engiles, J.B., Galantino-Homer, H. (2020). Interleukin-17A pathways target genes are upregulated in *Equus caballus* supporting limb laminitis. PLoS One 15(12). Available: https://doi.org/10.1371%2Fjournal.pone.0232920 [accessed September 24, 2024].

Charlesworth, D. (2003). Effects of inbreeding on the genetic diversity of populations. Philosophical Transactions: Biological Sciences 358(1434), 1051-1070.

Chen, Y., Gridley, T. (2013). The Snai1 and Snai2 proteins occupy their own and each other's promoter during chondrogenesis. Biochemical and Biophysical Research Communications 435(3), 356-360.

Chen, C., Zhu, B., Tang, X., Chen, B., Liu, M., Gao, N., Li, S., Gu, J. (2023). Genome-wide assessment of runs of homozygosity by whole-genome sequencing in diverse horse breeds worldwide. Genes 14(6). Available: https://doi.org/10.3390%2Fgenes14061211 [accessed September 30, 2024].

Choudry, S.R., Babes, L., Rahn, J.J., Ahn, B.Y., Goring, K.A.R., King, J.C., Lau, A., Petri, B., Hao, X., Chojnacki, A.K., Thanabalasuruar, A., McAvoy, E.F., Tabaries, S., Schraeder, C., Patel, K.D., Siegel, P.M., Kopciuk, K.A., Schriemer, D.C., Muruve, D.A., Kelly, M.M., Yipp, B.G., Kubes, P., Robbins, S.M., Senger, D.L. (2019). Dipeptidase-1 is an adhesion receptor for neutrophil recruitment in lungs and liver. Cell 178(5), 1205-1221.

Contreras-Aguilar, M.D., Tvarijonaviciute, A., Monkeviciene, I., Martin-Cuervo, M., Gonzalez-Arostegui, L.G., Fanco-Martinez, L., Ceron, J.J., Tecles, F., Escribano, D. (2020). Characterization of total adenosine deaminase activity (ADA) and its isoenzymes in saliva and serum in health and inflammatory conditions in four different species: an analytical and clinical validation pilot study. BMC Veterinary Research 16(1). Available: https://doi.org/10.1186/s12917-020-02574-2 [accessed September 24, 2024].

Cothran, E. G., Van Dyk, E. (1998). Genetic analysis of three South African horse breeds. Journal of the South African Veterinary Association 69(4): 120–125.

Cunningham, E.P, Dooley, J.J., Splan, R.K., Bradley, D.G. (2001). Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. Animal Genetics 32(6), 360-364.

Curik, I., Ferencakovic, M., Solkner, J. (2014). Inbreeding and runs of homozygosity: A possible solution to an old problem. Livestock Science 166, 26-34.

Dahan, T., Breitbart, H. (2022). Involvement of metabolic pathways in the sperm spontaneous acrosome reaction. Theriogenology 192, 38-44.

Daud, S., Ahmed, N., Naudhani, S., Younus, M., Manzoor, S., Firyal, S., Yaqub, T., Hashmi, A.S., ul-Amin, R., Wasim, M., Umar, S. (2018). In-silico analysis of gene *ALS2* genetic variants identified in the affected horses and humans with motor neuron disease. Journal of Equine Veterinary Science 62, 60-65.

Dementieva, N., Nikitkina, E., Scherbakov, Y., Nikolaeva, O., Mitrofanova, O., Ryabova, A., Atroshchenko, M., Makhmutova, O., Zaitsev, A. (2023). The genetic diversity of stallions of different breeds in Russia. Genes 14(7). Available: https://doi.org/10.3390/genes14071511 [accessed September 24, 2024].

De Klerk, J.N., Quan, M., Grewar, J.D. (2020). Socio-economic impacts of working horses in urban and peri-urban areas of the Cape Flats, South Africa. Journal of the South African Veterinary Association 91. Available: https://doi.org/10.4102/ jsava.v91i0.2009 [accessed November 25, 2024].

De Meeus d'Argenteuil, C., Boshuizen, B., Oosterlinck, M., van de Winkel, D., de Spiegelaere, W., de Bruijn, K., Goethals, K., Vanderperren, K., Delesalle, C.J.G. (2021). Flexibility of equine bioenergentics and muscle plasticity in response to different types of training: an integrative approach, questioning existing paradigms. PLOS One 16(4). Available: https://doi.org/10.1371%2Fjournal.pone.0249922 [accessed October 01, 2024].

De Silveria, J.C., Carnevale, E.M., Winger, Q.A., Bouma, G.J. (2014). Regulation of *ACVR1* and *ID2* by cell-secreted exosomes during follicle maturation in the mare. Reproductive Biology and Endocrinology 12. Available: https://doi.org/10.1186/1477-7827-12-44 [accessed September 24, 2024].

De Simoni Gouveia, J.J., Barbosa da Silva, M.V.G., Paiva, S.R., Pinheiro de Oliveria, S.M. (2014). Identification fo selection signatures in livestock species. Genetics and Molecular Biology 37(2), 330-342.

Dixit, S.P., Singh, S., Ganguly, I., Bhatia, A.K., Sharma, A., Kumar, N.A., Dang, A.K., Jayakumar, S. (2020). Genome-wide runs of homozygosity revealed selection signatures in Bos indicus. Frontiers in Genetics 11. Available: https://doi.org.10.3389/fgene.2020.00092 [accessed November 16, 2023].

Eldredge, E.A. (1988). Land, politics, and censorship: The history of the nineteenth-century Lesotho. History in Africa 15, 191-209.

Excoffier, L., Laval, G., Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetic data analysis. Evolutionary Bioinformatics Online 1. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2658868/ [accessed October 18, 2023].

Excoffier, L., Smouse, P.E. and Quattro, J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131, 479-491.

Falkingham, L. (2018). 12 New Forest pony facts every native fan needs to know. Horse and Hound. Available: https://www.horseandhound.co.uk/features/new-forest-pony-facts-642356 [accessed September 30, 2024].

Fettelschoss-Gabriel, A., Fettelschoss, V., Olomski, F., Birkmann, K., Thoms, F., Buhler, M., Kummer, M., Zeltins, A., Kundig, T.M., Bachmann, M.F. (2019). Active vaccination against interleukin-5 as long-term treatment for insect-bite hypersensitivity. Allergy 74(3), 572-582.

Forutan, M., Mahyari, S.A., Baes, C., Melzer, N., Schenkel, F.S., Sargolzaei, M. (2018). Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. BMC Genetics 19(98). Available: https://doi.org/10.1186/s12864-018-4453-z [accessed November 16, 2023].

Frankham, R., Ballou, J.D., Briscoe, D.A. (2002). Introduction to conservation genetics. Cambridge University Press, Cambridge. Available: https://doi.org/10.1017/CBO9780511808999 [accessed November 16, 2023].

Francis, R.M. (2017). POPHELPER: an R package and web app to analyse and visualize population structure. Molecular Ecology Resources 17(1), 27-32.

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., Francois, O. (2014). Fast and efficient estimations of individual ancestry coefficients. Genetics 196(4), 973-983.

Gmel, A.I., Brem, G., Neuditschko, M. (2023). New genomic insights into the conformation of Lipizzan horses. Scientific Reports 13. Available: https://doi.org/10.1038/s41598-023-36272-4 [accessed September 22, 2024].

Greenbaum, G., Templeton, A.R., Bar-David S. (2016). Inference and analysis of population structure using genetic data and network theory. Genetics 202(4), 1299-1312.

Grilz-Seger, G., Neuditschko, M., Ricard, A., Velie, B., Lindgren, G., Mesaric, M., Cotman, M., Horna, M., Dobretsberger, M., Brem, G., Druml, T. (2019). Genome-wide homozygosity patterns and evidence for selection in a set of European and near Eastern horse breeds. Genes 10(7). Available: https://doi.org/10.3390%2Fgenes10070491 [accessed September 22, 2024].

Han, H., McGivney, B.A., Farries, G., Katz, L.M., MacHugh, D.E., Randhawa, I.A.S., Hill, E.W. (2020). Selection in Australian Thoroughbred horses acts on a locus associated with early two-year old speed. PLOS ONE 15(2). Available: https://doi.org/10.1371/journal.pone.0227212 [accessed November 15, 2023].

He, S., Di, J., Han, B., Chen, L., Liu, M., Li, W. (2020). Genome-wide scans for runs of homozygosity identifies candidate genes related to economically important traits in Chinese Merino. Animals 10(3). Available: https://doi.org/10.3390%2Fani10030524 [accessed November 16, 2023].

He, S., Fang, X., Lu, X., Liu, Y., Li, G., Zhao, Z., Li, J., Yang, R. (2022). Function identification of Bovine ACSF3 gene and its association with lipid metabolism traits in beef cattle. Frontiers in Veterinary Science 8. Available: https://doi.org/10.3389/fvets.2021.766765 [accessed September 23, 2024].

Herringer, P. (2004). Thoroughbred horse racing and breeding in South Africa. Thoroughbred Heritage. Available: https://www.tbheritage.com/Breeders/SABreeding/SABreeders.html [accessed November 16, 2023].

Hernandez-Delgado, P., Felix-Portillo, M., Martinez-Quintana, J.A. (2023). ADAMTS proteases: importance in animal reproduction. Genes 14(6). Available: https://doi.org/10.3390/genes14061181 [accessed September 22, 2024].

Holtby, A.R., Hall, T.J., McGivney, B.A., Han, H., Murphy, K.J., MacHugh, D.E., Katz, L.M., Hill, E.W. (2023). Integrative genomics analysis highlights functionally relevant genes for equine behaviour. Animal Genetics 54(4), 457-469.

Hunt, J. (2005) in “Dutch South Africa: early settlers at the Cape, 1652-1708”. Troubador Publishing Ltd. Available: https://www.tandfonline.com/doi/abs/10.1111/j.1467-8306.1976.tb01070.x [accessed September 19, 2023].

Jebali, A., Dumaz, N. (2018). The role of RICTOR downstream of receptor tyrosine kinase in cancers. Molecular Cancer 17. Available: https://doi.org/10.1186/s12943-018-0794-0 [accessed September 22, 2024].

Johanson, H.C., Chen, W., Wicking, C., Sturm, R.A. (2010). Inheritance of novel mutated allele of the OCA2 gene associated with high incidence of oculocutaneous albinism in a Polynesian community. Journal of Human Genetic 55, 103-111.

Josem C.S., Pereira, C., Ferreira, M., Andre, A., Osorio, H., Gullo, I., Carneiro, F., Oliveira, C. (2023). 3D chromatin architecture re-writing at the *CHD3/CDH1* loci contributes to E-cadherin and P-cadherin expression switch in gastric cancer. Biology 12(6). Available: https://doi.org/10.3390/biology12060803 [accessed September 21, 2024].

Joubert, D.M., Bosman, W.M. (1971). The Nooitgedacht pony. South African Journal of Science. Available: https://journals.co.za/doi/pdf/10.10520/AJA00382353_9655 [accessed September 19, 2023].

Kalbfleisch, T., Rice, Edward E.S., DePriest, M.S., Walenz, B.P., Hestand, M.S., Vermeesch, J., O'Connell, B., Fiddes, I.T., Vershinina, A., Petersen, J.L., Finno, C.J., Bellone, R., McCue, M.E., Brooks, S.A., Bailey, E., Orlando, L., Green, R.E., Miller, D., Antczak, D., MacLeod, J.N. (2018). EquCab3, an Updated Reference Genome for the Domestic Horse. BioRxiv. Available: http://dx.doi.org/10.1101/306928 [accessed October 16, 2023].

Karagianni, A.E., Lisowski, Z.M., Hume, D.A., Pirie, R.S. (2020). The equine mononuclear phagocyte system: The relevance of the horse as a model for understanding human innate immunity. Equine Veterinary Journal 53, 231-249.

Kim, D., Ha, S.K., Gonzelez, F.J. (2024). *CBFA2T3* is PPARA sensitive and attenuates fasting-induced lipid accumulation in mouse liver. Cells 13(10). Available: https://doi.org/10.3390/cells13100831 [accessed September 23, 2024].

Kim, S.K., Shakya, A.K., O’Cakllaghan, D.J. (2021). Interferon Gamma inhibits Equine Herpesvirus 1 replication in a cell line-dependent manner. Pathogens 10(4). Available: https://doi.org/10.3390%2Fpathogens10040484 [accessed September 24, 2024].

Kim, J.H., Gurumurthy C.B., Naramura, M., Zhang, Y., Dudley, A.T., Doglio, M., Band, H., Band, V. (2009). Role of mammalian ecdysoneless in cell cycle regulation. Journal or Biological Chemistry 284(39), 26402-26410.

Knapp, A. (2021). Lesotho’s national transport – a brief history of the Basotho pony. The Design Train. Available: https://www.thedesigntrain.com/post/the-basotho-pony-lesotho-s-national-transport-a-brief-look-at-this-fascinating-history [accessed September 20, 2023].

Kodama, T., Marian, T.A., Lee, H., Kodama, M., Li, J., Parmacek, M.S., Jenkins, N.A., Copeland, N.G., Wei, Z. (2019). MRTFB suppresses colorectal cancer development through regulating SPDL1 and MCAM. Proceedings of the National Academy of Sciences 116(47), 23625-23635.

Kullmann, A., Weber, P.S., Bishop, J.B., Roux, T.M., Norby, B., Burns, T.A., McCutcheon, L.J., Belknap, J.K., Geor, R.J. (2015). Equine insulin receptor and insulin-like growth factor-1 receptor expression in digital lamellar tissue and insulin target tissues. Equine Veterinary Journal 48(5), 626-632.

Lai, Z., Wu, F., Zhou, Z., Li, M., Gao, Y., Yin, G., Yu, J., Lei, C., Dang, R. (2020). Expression profiles and polymorphic identification of the ACSL1 gene and their association with body size traits in Dezhou donkeys. Archives Animal Breeding 63, 377-386.

Langreder, N., Schackermann, D., Meier, D., Becker, M., Schubert, M., Dubel, S., Reinard, T., Figge-Wegener, S., Robbach, K., Baumer, W., Ladel, S., Hust, M. (2023). Development of an inhibiting antibody against equine interleukin 5 to treat insect bite hypersensitivity of horses. Scientific Reports 13. Available: https://doi.org/10.1038/s41598-023-31173-y [accessed September 24, 2024].

Laseca, N., Demyda-Peyras, S., Valera, M., Ramon, M., Escribo, B., Perdomo-Gonzalez, D.I., Molina, A. (2022). A genome-wide association study of mare fertility in the Pura Raza Espanol horse. Animal 16(3). Available: https://doi.org/10.1016/j.animal.2022.100476 [accessed September 24, 2024].

Lee, H.Y., Kim, J.Y., Kim, K.H., Jeong, S., Cho, Y., Kim, N. (2020). Gene expression profile in similar tissues using transcriptome sequencing data of whole-body horse skeletal muscle. Genes 11(11). Available: https://doi.org/10.3390%2Fgenes11111359 [accessed September 22, 2024].

Lekota, T.A. (2014). The state of the Basotho pony in Lesotho. FAO. Available: https://www.fao.org/fileadmin/user_upload/nr/sustainability_pathways/docs/The%20State%20of%20the%20Basotho%20Pony%20in%20Lesotho.pdf [accessed September 22, 2023].

Lewis, S.L., Holl, H.M., Streeter, C., Posbergh, C., Schanbacher, B.J., Place, N.J., Mallicote, M.F., Long, M.T., Brooks, S.A. (2017). Genome wide association study reveals a risk locus for equine metabolic syndrome in the Arabian horse. Journal of Animal Science 95(3), 1071-1079.

Li, X., Bi, H., Xie, S., Cui, W. (2022). MiR-208b regulates the conservation of skeletal muscle fiber types by inhibiting *Mettl8* expression. Frontiers in Genetics 13. Available: https://doi.org/10.3389/fgene.2022.820464 [accessed September 22, 2024].

Liao, M., Liao, J., Qu, J., Shi, P., Cheng, Y., Pan, Q., Zhao, N., Zhang, X., Zhang, L., Tan, Y., Li, Q., Zhu, J.F., Li, J., Zhang, C., Cai, S.Y., Chai, J. (2023). Hepatic TNFRSF12A promotes bile acid-induced hepatocyte pyroptosis through NFkB/Caspase-1/GSDMD signalling in cholestasis. Cell Death Discovery 9. Available: https://doi.org/10.1038/s41420-023-01326-z [accessed September 24, 2024].

Loux, S., Robles, M., Chavatte-Palmer, P., de Mestre, A. (2022). Markers of equine placental differentiation: insights from gene expression studies. Society for Reproduction and Fertility 163(3). Available: https://doi.org/10.1530/REP-21-0115 [accessed September 22, 2024].

Lye, W.F. (1967). The Difaqane: The Mfecane in the Southern Sotho area, 1822-24. Journal of African History 8(1), 107-131.

Maluleke, W., Tshabalala, N.P., Barkhuizen, J. (2020). The effects of climate change on rural livestock farming: Evidence from Limpopo Province, South Africa. Asian Journal of Agricultire and Rural Development 10(2), 645-658.

Mancuso, G., Barth, E., Crivello, P., Rugarli, E.I. (2012). Alternative splicing of *Spg7*, a gene involved in hereditary spastic paraplegia, encodes a variant of paraplegin targeted to the endoplasmic reticulum. PLOS ONE 7(5). Available: https://doi.org/10.1371/journal.pone.0036337 [accessed September 23, 2024].

Marshall, H.M., Ronen, K., Berry, C., Llano, M., Sutherland, H., Saenz, D., Bickmore, W., Poeschla, E., Bushman, F.D. (2007). Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting. PLOS One. Available: https://doi.org/10.1371/journal.pone.0001340 [accessed September 22, 2024].

Marlow, C.H.B. (2010). A brief history of equine private practice in South Africa. Journal of the South African Veterinary Association 81(4), 190-200.

Mastrangelo, S., Tolone, M., Di Gerlando, R., Fontanesi, L., Sardina, M.T., Portolano, B. (2016). Genomic inbreeding estimation in small populations: Evaluation of runs of homozygosity in three local dairy cattle breeds. Animal 10(5), 746-754.

McBride, S.D., Mills, D.S. (2012). Psychological factors affecting equine performance. BMS Veterinary Research 8 (180). Available: https://doi.org/10.1186/1746-6148-8-180 [accessed November 16, 2023].

McCormack, J. (2018). PSXVI-14 the Basotho pony. A breed improvement program in the highlands of Lesotho. Journal of Animal Science 96. Available: http://dx.doi.org/10.1093/jas/sky404.551 [accessed September 21, 2023].

McCoy, A.M., Beeson, S.K., Rubin, C.J., Andersson, L., Caputo, P., Lykkjen, S., Moore, A., Piercy, R.J., Mickelson, J.R., McCue, M.E. (2019). Identification and validation of genetic variants predictive of gait in Standardbred horses. PLOS ONE 15(5). Available: https://doi.org/10.1371/journal.pgen.1008146 [accessed November 15, 2023].

McCue, M.E., Bannasch, D.L., Petersen, J.L., Gurr, J., Bailey, E., Binns, M.M., Distl, O., Guerin, G., Hasegawa, T., Hill, E.W., Leeb, T., Lindgren, G., Penedo, M.C.T., Roed, K.H., Ryder, O.A., Swinburne, J.E., Tozaki, T., Valberg, S.J., Vaudin, M., Lindblad-Toh, K., Wade, C.N., Mickelson, J.R. (2012). A high-density SNP array for the domestic horse and extant Perissodactlya: Utility for association mapping, genetic diversity, and phylogeny. PLOS Genetics 8(1). Available: https://doi.org/10.1371/journal.pgen.1002451 [accessed November 16, 2023].

McGivney, B.A., Han, H., Corduff, L.R., Katz, L.M., Tozaki, T., MacHugh, D.E., Hill, E.W. (2020). Genomic inbreeding trends, influential sire lines and selection in the global Thoroughbred horse population. Scientific Reports 10. Available: https://doi.org/10.1038/s41598-019-57389-5 [accessed January 28, 2024].

McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., MacLeod, A.K., Farrington, S.M., Rudan, P., Hayward, C., Vitart, V., Rudan, I., Wild, S.H., Dunlop, M.G., Wilson, J.F. (2008). Runs of homozygosity in European populations. AJHG 83(3), 359-372.

McVean, G. (2009). A genealogical interpretation of principal components analysis. The Public Library of Science Genetics 5(10). Available: https://doi.org/10.1371%2Fjournal.pgen.1000686 [accessed October 17, 2023].

Metzer, J., Karwath, M., Tonda, R., Beltran, S., Agueda, L., Gut, M., Gut, I.G., Distl, O. (2015). Runs of homozygosity reveal sigantures of positive selection for reproduction traits in breed and non-breed horses. BMC Genomics 16. AVailavle: https://doi.org/10.1186/s12864-015-1977-3 [accessed September 24, 2024].

Metzger, J., Tonda, R., Beltran, S., Agueda, L., Gut, M., Distl, O. (2014). Next generation sequencing gives an insight into the characteristics of highly selected breeds versus non-breed horses in the course of domestication. BMC Genomics 15. Available: https://doi.org/10.1186/1471-2164-15-562 [accessed September 30, 2024].

Millin, P. (2007). The renowned Basotho pony. SA Horseman, 31-34.

Millin, R. (2011). Saddle Horse Breeding in South Africa. SA Horseman, 10-13.

Mokoena, S.K., Balkaran, S. (2018). An exploration of constraints for free movement of people in Africa – A case of Lesotho and South Africa. African Journal of Public Affairs 10(2), 109-126.

Moon, S., Lee, J.W., Shin, D., Shin, K.Y., Kim, J., Choi, I.Y., Kim, J., Kim, H. (2015). A genome-wide scan for selective sweeps in racing horses. Asian-Australasia Journal of Animal Sciences 28(1)1), 1525-1531.

Moss, K.L., Jiang, Z., Dodson, M.E., Linardi, R.L., Haughan,. J., Gale, A.L., Grzybowski, C., ENgiles, J.E>, Stefanovski, D., Robinson, M.A., Ortved, K.F. (2020). Sustained interleukin-10 transgene expression following intra-articular AAV5-IL-10 administration to horses. Human Gene Therapy 31(1), 110-118.

Olender, T., Jones, T.E., Bruford, E., Lancet, D. (2020). A unified nomenclature for vertebrate olfactory receptors. BMC Evolutionary Biology 20. Available: https://doi.org/10.1186/s12862-020-01607-6 [accessed September 22, 2024].

Pacheco, H.A., Rossoni, A., Cecchinato, A., Penagaricano, F. (2023). Identification of runs of homozygosity associated with male fertility in Italian Brown Swiss cattle. Frontiers in Genetics 14. Available: https://doi.org/10.3389/fgene.2023.1227310 [accessed November 16, 2023].

Pasandideh, M., Gholizadeh, M., Rahimi-Mianji, G. (2020). A genome-wide association study revealed five SNPs affecting 8-month weight in sheep. Animal Genetics 51(6), 973-976.

Perdomo-Gonzalez, D.I., Sanchez-Guerro, M.J., Bartolome, E., Guedes dos Santos, R., Molina, A., Valera, M. (2023). Designing an early selection morphological traits index for reproductive efficiency in Pura Raza Espanola mares. Journal of Animal Science 102. Available: https://doi.org/10.1093/jas/skad409 [accessed Sepetember 30, 2024].

Peripolli, E., Stafuzza, N.B., Munari, D.P., Lima, A.L.F., Irgang, R., Machado, M.A., Claudio do Carmo Panetto, J., Ventura, R.V., Baldi, F., Barbosa da Silva, M.V.G. (2018). Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. BMC Genomics 19(34). Available: https://doi.org/10.1186%2Fs12864-017-4365-3 [accessed November 16, 2023].

Petersen, J.L., Mickelson, J.R., Cothran, E.G., Andersson, L.S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M.M., Borges, A.S., Brama, P., da Câmara Machado, A., Distl, O., Felicetti, M., Fox-Clipsham, L., Graves, K.T., Guérin, G., Haase, B., Hasegawa, T., Hemmann, K., Hill, E.W., Leeb, T., Lindgren, G., Lohi, H., Lopes, M.S., McGivney, B.A., Mikko, S., Orr, N., Penedo, M.C.T., Piercy, R.J., Raekallio, M., Rieder, S., Røed, K.H., Silvestrelli, M., Swinburne, J., Tozaki, T., Vaudin, M., Wade, C.M., McCue, M.E. (2013). Genetic diversity in the odern horse illustrated from genome-wide SNP data. PLOS One 8(1). Available: https://doi.org/10.1371/journal.pone.0054997 [accessed November 16, 2023].

Pijoos, I. (2018). 'They don't care for the horses', says NSPCA of rural Eastern Cape horse races. Times Live South Africa. Available: https://www.timeslive.co.za/news/south-africa/2018-11-01-they-dont-care-for-the-horses-says-nspca-of-rural-eastern-cape-horse-races/ [accessed November 16, 2023].

Pinnock, D. (2022). From rock stars to obscurity – searching for lost Thoroughbreds. Daily Maverick. Available: https://www.dailymaverick.co.za/article/2022-06-28-from-rock-stars-to-obscurity-searching-for-lost-thoroughbreds/ [accessed November 16, 2023].

Plawinska-Czarnak, J., Majewska, A., Zarzynska, J., Bogdan, J., Kaba, J., Anusz, K., Bagnicka, E. (2021). Gene expression profile in peripheral blood nuclear cells of small ruminant lentivirus-seropositive and seronegative dairy goats in their first lactation. Animals 11(4). Available: https://doi.org/10.3390/ani11040940 [accessed September 21, 2024].

Pongratz, N., Gerace, L., Michiels, N.K. (2002). Genetic differentiation within and between populations of a hermaphroditic freshwater planarian. Hereditary 89, 64-69.

Posbergh, C.J., Pollott, G.E., Southard, T.L., Divers, T.J., Brooks, S.A. (2018). A nonsynonymous change in adhesion G protein-coupled receptor L3 associated with risk for equine degenerative myeloencephalopathy in the Caspian horse. Journal of Equine Veterinary Science 70, 96-100.

Potgieter, T., Grundlingh, A. (2007). Admiral Elphinstone and the Conquest and the Defence of the Cape of Good Hope. South African Journal of Military Studies 35(2), 39-67.

Pozharskiy, A., Abdrakhmanova, A., Beishova, I., Shamshindin, A., Nametov, A., Ulyanova, T., Bekova, G., Kikebayev, N., Kovalchuk, A., Ulyanov, V., Turabayev, A., Khusnitdinova, M., Zhambakin, K., SapaKhova, Z., Shamekova, Gritsenko, D. (2023). Genetic structure and genome-wide association study of the traditional Kazakh horses. Animal 17(9). Available: https://doi.org/10.1016/j.animal.2023.100926 [accessed September 21, 2024].

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 for whole-genome associations and population-based linkage analyses. The American Journal of Human Genetics 81(3), 559-575.

Purfield, D.C., Berry, D.P., McParland, S., Bradley, D.G. (2012). Runs of homozygosity and population history in cattle. BioMed Central Genetics 13. Available: https://doi.org/10.1186/1471-2156-13-70 [accessed October 22, 2023].

Rorvang, M.V., Nielsen, B.L., McLean, A.N. (2020). Sensory abilities of horses and their importance for equitation science. Frontiers in Veterinary Science 7. Available: https://doi.org/10.3389%2Ffvets.2020.00633 [accessed September 30, 2024].

Rosa, L.P., Martin, K., Vierra, M., Lundquist, E., Foster, G., Brooks, S.A., Lafayette, C. (2022). A *KIT* variant associated with increased white spotting epistatic to *MCR1* genotype in horses. Animals 12(15). Available: https://doi.org/10.3390/ani12151958 [accessed September 22, 2024].

Schaefer, R.J., McCue, M.E. (2020). Equine genotyping arrays. Veterinary Clinics of North America: Equine Practice 36(2), 183-193.

Secco, B., Camire, E., Briere, M.A., Caron, A., Billong, A., Gelinas, Y., Lemay, A.M., Tharp, K.M., Lee, P.L., Gobeil, S., Guimond, J.V., Patey, N., Guertin, D.A., Stahl, A., Haddad, E., Marsolais, D., Bosse, Y., Birsoy, K., Laplante, M. (2017). Amplification of adipogenic commitment by *VSTM2A*. Cell Reports 18(1), 93-106.

Shen, Y., Wang, L., Fu, J., Xu, X., Yue, G.H., Li., J. (2019). Population structure, demographic history and local adaption of the grass carp. BMC Genomics 20 (467). Available: https://doi.org/10.1186/s12864-019-5872-1 [accessed November 15, 2023].

Shriner, D. (2013). Overview of admixture mapping. Current Protocols in Human Genetics 76 (1). Available: https://doi.org/10.1002/0471142905.hg0123s76 [accessed November 16, 2023].

Sigala, S., Bodei, S., Missale, C., Zani, D., Simeone, C., Cunico, S.C., Spano, P.F. (2008). Gene expression profile of prostate cancer cell lines: effect of nerve growth factor treatment. Molecular and Cellular Endocrinology 284(1), 11-20.

Slate, J., David, P., Dodds, K.G., Veenvliet, B.A., Glass, B.C., Broad, T.E., McEwan, J.C. (2004). Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: Theoretical expectations and empirical data. Heredity 93, 255-265.

Stefaniuk-Szmukier, M., Ropka-Mlik, K., Zagrajczuk, A., Piorkowska, K., Szmatola, T., Luszczynski, J., Bugno-Poniewierska, M. (2018). Genetic variability in equine *GDF9* and *BMP15* genes in Arabian and Thoroughbred mares. Annals of Animal Science 18(1), 39-52.

Suarez-Pajes, E., Díaz-de Usera, A., Marcelino-Rodríguez, I., Guillen-Guio, B., Flores, C. (2021). Genetic ancestry inference and its application for the genetic mapping of human diseases. International Journal of Molecular Science 22(13). Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8269095/pdf/ijms-22-06962.pdf [accessed October 05, 2023].

Sumreddee, P., Toghiani, S., Hay, E.H., Roberts, A., Aggrey, S.E., Rekaya, R. (2020). Runs of homozygosity and analysis of inbreeding depression. Journal of Animal Science 98(12). Available: https://doi.org/10.1093%2Fjas%2Fskaa361 [accessed November 16, 2023].

Swart, S. (2003). Riding high – horses, power and settler society, c. 1654-1840. Kronos, Environmental History 29. Available: https://doi.org/10.2307/41056494 [accessed September 19, 2023].

Swart, S. (2008). 'High Horses': Horses, class and socio-economic change in South Africa. Journal of South African Studies 34(1), 193-213.

Swart, S. (2010). "The world the horses made": A South African case study of writing animals into social history. International Review of Social History 55(2), 241-263.

Swindell, W.R., Bouzat, J.L. (2006). Selection and inbreeding depression: effects of inbreeding rate and inbreeding environment. Evolution 60(5), 1014-1022.

Szwaczkowski, T., Gregula-Kania, M., Stachruska, A., Borowska, A., Jaworski, Z., Gruszecki, T.M. (2016). Inter-and intra-genetic diversity in the Polish Konik horse: implications for the conservation program. Canadian Journal of Animal Science 96, 570-580.

Tian, X., Pascal, G., Monget, P. (2009). Evolution and functional divergence of NLRP genes in mammalian reproductive systems. BMC Evolutionary Biology 9(202). Available: https://doi.org/10.1186/1471-2148-9-202 [accessed September 21, 2024].

Todd, E.T., Hamilton, N.A., Velie, B.D., Thomson, P.C. (2020). The effects of inbreeding on covering success, gestation length and foal sex ratio in Australian Thoroughbred horses. BMC Genetics 21 (41). Available: https://doi.org/10.1186/s12863-020-00847-1 [accessed November 16, 2023].

Todd, E.T., Ho, S.Y.W., Thomson, P.C., Ang, R.A., Veilie, B.D., Hamilton, N.A. (2018). Founder-specific inbreeding depression affects racing performance in Thoroughbred horses. Scientific Reports 8. Available: https://doi.org/10.1038/s41598-018-24663-x [accessed September 30, 2024].

Toro-Ospina, A.M., Herrera-Rios, A.C., Santos, W.B., Schettini, G.P., Aristizabal, V.H.V., Claros, G.T., Morea, E.G.O. (2024). Genetic architecture and signatures of selection in the Caqueteno Creole (Colombian Native Cattle). Diversity 16(6). Available: https://doi.org/10.3390/d14100828 [accessed September 24, 2024].

Tsuyama, T., Tada, S., Watanabe, S., Seki, M., Enomoto, T. (2005). Licensing for NDA replication requires a strict sequential assembly of Cdc6 and Cdt1 onto chromatin in *Xenopus* egg extracts. Nucleic Acids Research 33(2), 765-775.

Valberg, S.J., Velez-Irizarry, D., Williams, Z.J., Henry, M.L., Iglewski, H., Herrick, K., Fenger, C. (2022). Enriched pathways of calcium regulation, cellular/oxidative stress, inflammation, and cell proliferation characterize gluteal muscle of Standardbred horses between episodes of recurrent exertional rhabdomyolysis. Genes 13(10). Available: https://doi.org/10.3390/genes13101853 [accessed September 24, 2024].

Van Gorp, P.R.R., Zhang, J., Liu, J., Tsonaka, R., Mei, H., Dekker, S.O., Bart, C.I., De Coster, T., Post, H., Heck, A.J.R., Schalij, M.J., Atsma, D.E., Pijnappels, D.A., de Vries, A.A.F. (2022). Sbk2, a newly discovered atrium-enriched regulator of sarcomere integrity. Circulation Research 131, 24-41.

Vanhoutteghem, A., Delhomme, B., Herve, F., Nondier, I., Petit, J.M., Araki, M., Araki, K., Djian, P. (2016). The importance of basonuclin 2 in adult mice and its relation to basonuclin 1. Mechanisms of Development 140, 53-73.

Wang, Q., Xu, J., Bao, M., Wang, H., Sun, X., Ji, D., Wang, J., Li, Y. (2024). Weighted gene co-expression network analysis reveals genes related to growth performance in Hu sheep. Scientific Reports 14. Available: https://doi.org/10.1038/s41598-024-63850-x [accessed September 21, 2024].

Weir, B.S., Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. Evolution 38(6), 1358-1370.

Weir, B.S. (1996). Genetic data analysis II: Methods for discrete population genetic data. Sinauer Associates, Inc. Sunderland.

Wilhelm, D., Wurtz, A., Abouelfarah, H., Sanchez, G., Bui, C., Vincourt, J.B. (2023). Tissue-specific collagen hydroxylation at GEP/GDP triplets mediated by *P4HA2*. Matrix Biology 119, 141-153.

Wolf, I., Rohrschneider, L.R. (1999). *Fiz1*, a novel Zinc Finger Protein interacting with the receptor Tyrosine Kinase *Flt3*. Cell Biology and Metabolism 274(30), 21478-21484.

Worth, D.C., Daly, C.N., Geraldo, S., Oozeer, F., Gordon-weeks, P.R. (2013). Debrin contains a cryptic F-actin-bundling activity regulated by CdK5 phosphorylation. Journal of Cell Biology 202(5), 793-806.

Van der Merwe, F., Martin, J. (2002). Four Southern African Horse Breeds. Animal Genetic Resources 32, 57-72.

Vincelette, A. (2023). The characteristics, distribution, function, and origin of alternative lateral horse gaits. Animals 13(16). Available: https://doi.org/10.3390/ani13162557 [accessed September 30, 2024].

Xiao, F., Wang, H., Fu, X., Li, Y., Ma, K., Sun, L., Gao, X., Wu, Z. (2011). Oncostatin M inhibits myoblast differentiation and regulates muscle regeneration. Cell Research 21(2), 350-364.

Xing, J., Xie, L., Qi, X., Liu, G., Akhtar, M.F., Li, X., Bou, G., Bai, D., Zhao, Y., Dugarjaviin, M., Zhang, X. (2023). Integrated analysis of transcriptome and proteome for exploring mechanism of promoting proliferation of equine satellite

cells associated with leucine. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics 48. Available: https://doi.org/10.1016/j.cbd.2023.101118 [accessed September 24, 2024].

Yuan, W., Zhang, X., Fan, X. (2023). The role of the Piezo1 mechanosensitive channel in heart failure. Current Issues in Molecular Biology 45(7), 5830-5848.

Zhang, B., Dietrich, U.M., Geng, J.G., Bicknell, R., Esko, J.D., Wang, L. (2009). Repulsive axon guidance molecule Slit3 is a novel angiogenic factor. Blood 114(9), 4300-4309.

Zhang, X., Zheng, Q., Wang, C., Zhou, H., Jiang, G., Miao, Y., Zhang, Y., Liu, Y., Li, Q., Qiu, X., Wang, E. (2017). CCDC106 promotes non-small cell lung cancer cell proliferation. Oncotarget 8(16), 26662-26670.

Zhang, Y., Cong, X., Wang, A., Jiang, H. (2014). Identification of the *STAC3* gene as a skeletal muscle-specifically expressed gene and a novel regulator of satellite cell differentiation in cattle. Journal of Animal Science 92)8), 3284-3290.

Zhou, J., Zhao, J., Xu, G., Wang, Z., Gao, J., Cui, S., Liu, J. (2017). Deletion of *Spata2* by CRISPR/Cas9n causes increased inhibin alpha expression and attenuated fertility in male mice. Biology of Reproduction 97(3), 497-513.

Chapter 5

Genome-wide Scan for Selection Signatures to Provide Insight into the Adaptive Potential of Indigenous and Rural South African Horse Populations

5.1 Abstract

Understanding adaptive selection signatures in indigenous and rural South African horse populations provides insights into how these populations have evolved under environmental, management, and cultural pressures. Collectively, 282 indigenous (Basotho Pony-type (BSP), rural (Eastern Cape Rural (ECR); East London Rural (ELR) and KwaZulu-Natal (KZN) rural) and commercial (SA Saddlebred (SAD); Thoroughbred (THR)) South African horse specimens were collected in KwaZulu-Natal, Eastern Cape and Gauteng provinces. This genomics-assisted, community-driven study aimed to characterize the genetic and functional diversity of rural and indigenous horse populations in South Africa by identifying and evaluating genomic regions under positive selection within and between these populations in order to provide insight into adaptive genetic traits, supporting conservation and sustainable management efforts for these culturally and environmentally valuable horse populations. Strong selective sweeps obtained from population-specific iHS analyses were detected on chromosomes 1, 2, 4 and 7 whereby traits linked to fertility and reproduction (*COBL*; *CFAP58*; *SPACA3*), nervous system development (*UNC13*; *ARHGAP35*; *DTNA*; *HTR1A*), growth and development (*FBN1*), feed efficiency (*GLB1L2*; *B3GAT1*; *PPARA*) and several cellular processes were observed. Cross-population XP-EHH analyses showed similar patterns of selection signatures, further supporting the fixation of these positively selected genomic regions within the populations. However, intersection of XP-EHH and iHS analyses revealed unique selection signatures functionally associated with hair follicle development (*SOS1*) in the ECR population, locomotory behaviour (*ADAM22*) in the ELR population and fear response (*SLC1A1*) in ELR vs KZN – driving adaption to unique ecological niches by developing specific morphological and behavioural adaptations that promote survival. Additionally, traits associated with blood vessel morphogenesis (*SEMA3E*; *TSPAN12*; *JAM3*) and olfactory responses (*OR7D20*; *OR1M1*) were found, reflecting adaptations to challenging environments. Comparisons with commercial breeds, namely SAD and THR, revealed no selection sweeps across the genome – indicating that artificial selection remained the dominant selection mechanism acting within the populations. This study demonstrated the utility of combining both iHS and XP-EHH analyses to elucidate population-specific and cross-population shared adaptive mechanisms, contributing to the understanding of the genetic and functional diversity in the indigenous, rural and commercial horse populations. Although the identification of positively selected genes may serve as valuable genomic targets for future breeding programs aimed at enhancing adaptability and resilience in equine populations, a major limitation in the the representation of the number of region-specific rural breed populations restricts comprehensive comparison.

Keywords: Selection signatures; iHS analyses; XP-EHH; South Africa; Indigenous horse populations; Rural horse populations

5.2 Introduction

Vast evolutionary mechanisms occur within populations that work in conjunction with one another, to ensure the survival and continuation of individuals forming a population, whereby one of the most influential of those mechanisms being positive selection – also referred to as Darwinian selection (Jose de Simoni Gouveia et al., 2008). Positive selection propagates the accumulation of advantageous alleles within populations, selectively swept from one generation to another, until it becomes fixed within that specific population (Pavlidis and Alachiotis, 2017). Koropoulis et al. (2020) suggests the fact that these selective sweeps within a population introduce traces of selection signatures within the genome, which can be detected through regions of heightened homozygosity, patterns of alleles in linkage disequilibrium (LD) as well as shifts of allele accumulation at specific sites. This facilitates the identification of genomic regions under selection that are translated into adaptive phenotypic characteristics. The detection of these selection signatures particularly within horse breed populations provide insight into the dominant evolutionary forces acting within those populations and the candidate genes in heightened frequency that produce favorable or desirable phenotypic characteristics – through which can be classified as breed-specific selection signatures (Gurgul et al., 2019).

Domestication of the horse species shifted the natural evolutionary forces acting within wild horse populations, in order to fulfill the economic and socio-economic selective pressures bought on by people for horses to be used in warfare, agricultural practices and sporting events – genetically molding specific populations (Librado et al., 2016). Through generations of these artificial selection breeding strategies focused on these horse populations, the accumulation of specific selection signatures shaped the horse breed populations present in the modern world – possessing unique breed-specific selection signatures. Commercial, purebred horse populations have been subjected to intense selective breeding that drives athletic phenotypic traits to be maintained and fixed within the population, one of the most influential being the Thoroughbred – often regarded as the most prominent breed within the horse racing industry (Lee at al., 2014). McGivney et al. (2020) conducted a recent genome-wide scan for selection signatures of Thoroughbred blood stocks across the world, uncovering significantly high inbreeding levels within these populations through generations of directional selection to accumulate favorable signatures that encode for aesthetic and agility functional traits that include metabolism, musculoskeletal ratio and behaviour. These researchers also gave light to the fact that positively-selected sweeps were also detected on fundamental sites owing to adaptive traits within the sampled breed populations that included growth and development, neurological signaling pathways and cardiovascular signaling – indicating that natural evolutionary forces continue to withstand artificial genetic shaping.

Another athletically-selected horse breed within the equine sporting world is the American Saddlebred, whereby over a century of selective breeding from the first importation of well-composed Saddlers to South Africa in the early 1900's, created the light-framed and high-arched aesthetics of the South African Saddlebred population adapted to South Africa's variable environment (Millin, 2011). Although there is a lack of genomic studies focusing on the detection of selection signatures within the South African Saddlebred populations, except for this study, researchers focus heavily on their American precursor populations. These studies include

identification of genomic regions under selection and are relatively high in frequency - such as the recessive chestnut coat colour (*MC1R* gene) (Peterson et al., 2013), early on-set lordosis (*TAF8*, *TRERF1* and *C6orf132* genes) (Cook et al., 2010) and selection for variation in 3- and 5-gaited ability (*DMRT3* gene) of this breed (Lukanova et al., 2018).

Metzger et al. (2015) concluded within their study of the detection of positively selected reproductive signatures in non-descript horse breed populations that both positive and artificial evolutionary forces, work in conjunction with one another that have shaped these breeds to maintain fixed adaptive selection signatures, and incorporate breed-specific signatures that accumulate due to selective breeding pressures. The very same inference can be made regarding the rural and indigenous horse populations belonging to South Africa, whereby collectively the rural non-descript horse populations and the Basotho-type indigenous populations, both contain a variety of selection signatures that are influenced by the adaptability, productivity and economic significance (Swart, 2010). Genetic characterization through the identification of genomic regions under positive selection provides valuable insight into historical selection pressures that have shaped the adaptive potential of each breed population (de Simoni Gouveia et al., 2014). The selection signatures identified within these regions help establish associations with advantageous phenotypic traits that enhance the adaptability of these horses to both human-mediated management and challenging environmental conditions (Tiwari et al., 2024). These findings further validate the importance of these populations as essential genetic resources within the South African context. Against this background, this genomics assisted, community-driven research study aims to characterize the genetic and functional diversity of the rural and indigenous horse populations within South Africa, which was achieved through the identification and assessment of genomic regions under positive selection found within and across populations.

5.3 Materials and Methods

5.3.1 Sample Collection

A total of 282 Equus caballus specimens were collected across South Africa from several rural villages, registered horse breeding societies, and commercial horse-racing training centres. The rural crossbred populations were sampled from three districts: Alfred Nzo (ECR; n = 48) and Amathole (ELR; n = 25) in the Eastern Cape Province and uMzimkhulu (KZN; n = 47) in KwaZulu-Natal Province. A population of the indigenous Basotho-type pony breed (BSP; n = 95) was sampled in the Alfred Nzo district of the Eastern Cape province, bordering the country of Lesotho – the breed's land of origin. Purebred Thoroughbred horses (THR; n = 47) were sampled from the Summerveld Training Centre in the eThekwini District of KwaZulu-Natal [source: https://www.goldcircle.co.za/racing/training-centres/training-centres-summerveld/]. Purebred Saddlebred horses (SAD; n = 20) were obtained from breeding studs in Gauteng Province, as per allocation by the Saddlebred Society of South Africa [source: https://saddlebred.co.za/; established in 1942 (Swart, 2008)]. GPS coordinates of sample locations included within this study were visualized in **Figure 5.1** below.

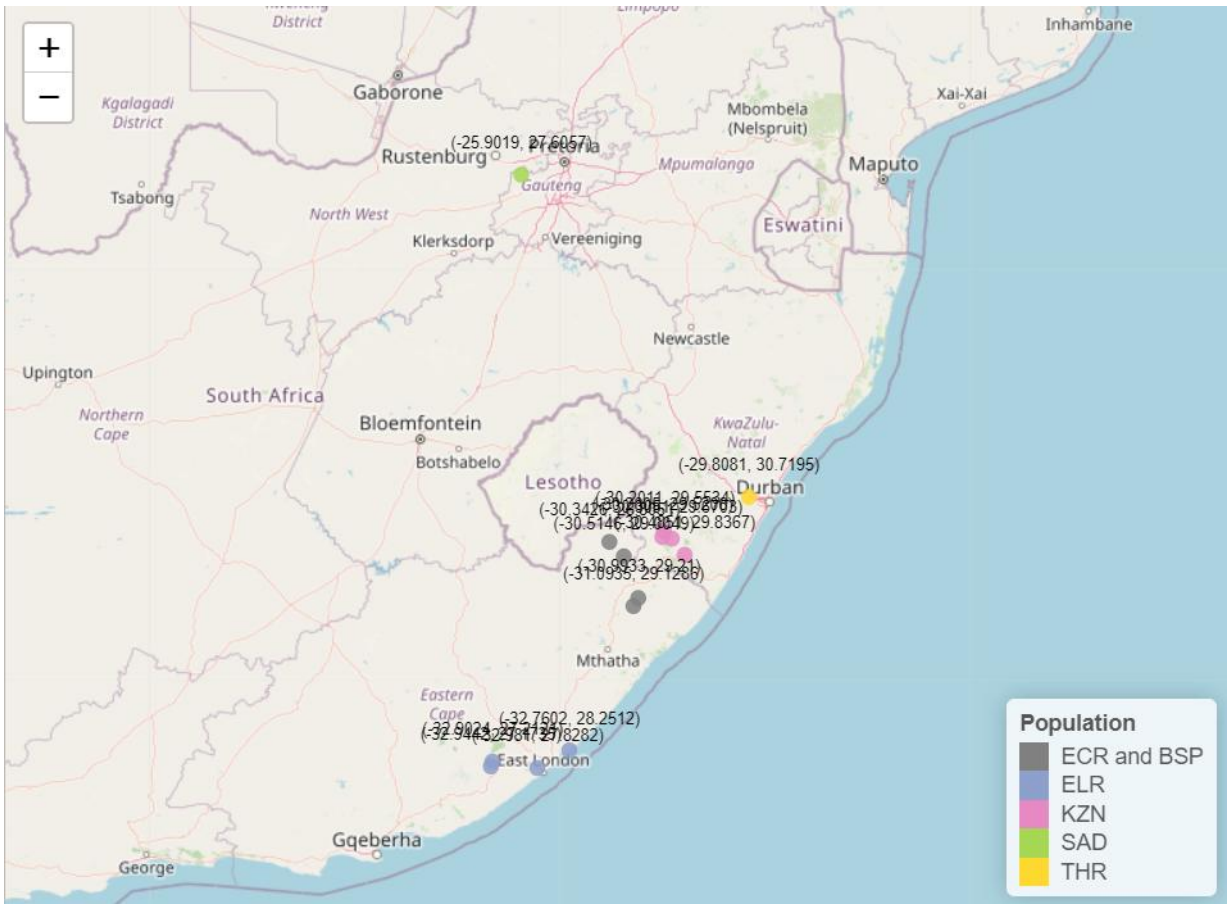


Figure 5.1 GIS map of sample regions within Eastern Cape, KwaZulu-Natal and Gauteng Provinces

[GPS coordinates of rural hourseholds, training facility and breeding stud was indicated. This GIS map was constructed using R/Rstudio version 4.3.1 (R Core Team, 2020; Rstudio Team, 2020), applying the leaflet package (Agafonkin, 2016).]

Tail-hair samples were collected in accordance with the standard operating procedures approved by Onderstepoort Veterinary Research Institutional Biosafety and Biosecurity Committee (OVR-IBBC). Ethical approval for the study was obtained from the OVR Animal Ethics Committee (**AEC 24.16**).

5.3.2 Genotyping and Quality Control

DNA was extracted from tail-hair follicles for each horse specimen according to the user manual generated by the manufacture of the NucleoMag® Tissue Kit (Macherey-Nagel). Concentrations of extracted DNA was then quantified for each sample using the Qubit® 4.0 Fluorometer to ensure all samples surpass minimum concentration (ng/µl) requirements for genotyping.

A total of 282 horse specimens from 6 different sampled populations were genotyped using the Illumina EquineSNP70k and EquineSNP80K BeadChip (Illumina, United States), containing an array of 70, 000 SNPs and 80,000 SNPs respectively, with an average of one SNP per 35kb window (Schaefer and McCue, 2020) at the Agricultural Research Council – Biotechnology Platform (ARC-BTP) in South Africa.

PLINK version 1.90 (Purcell et al., 2007) was used to perform a Quality Control (QC) analysis on the merged, SNP genotypic data of the 6 sampled populations from South African commercial, indigenous and rural horse breeds. The QC using PLINK version 1.90 was based on the updated horse reference genome assembly - *EquCab* 3.0 (Kalbfleisch et al., 2018). The pruning parameters of the QC included SNP markers that were filtered to exclude SNP markers with a departure from Hard-Weinberg equilibrium (HWE) of less than 10^{-6} (0.000001), SNP loci with a sample call rate of less than 90%, SNP markers with a genotyping rate of less than 90% and then SNP markers that contained a Minor Allele Frequency (MAF) of less than 5% (0.050). The remaining pruned, SNP marker dataset consisted of 79,592 SNPs and 282 horse specimens. This dataset was then further pruned to filter out individuals that show close kinship (IBD (Identity by Descent)) ≥ 0.25, as well as to filter out SNP makers with a high Linkage Disequilibrium (LD) $r^2 = 0.500$, with a window of every 5 SNPs and a window size of 50 kb. The resulting SNP dataset consisted of 71, 595 SNPs and 282 horse specimens remaining. This filtered data was then used to complete a genome-wide scan of recent positive selection within and across populations through Extended Haplotype Homozygosity (EHH) based tests (Sabeti et al., 2002).

5.3.3 Within population scan for selection signatures (iHS-based Scores)

Autosomes were each separately extracted from every horse breed population using PLINK version 1.90 (Purcell et al., 2007) in order to conduct haplotype phasing and imputation of SNP data contained within each chromosome file – using Beagle 5.4 version 22Jul22.46e (Browning et al., 2018).

All autosomal haplotype files were scanned (using the 'scan_hh' function) and combined to conduct integrated-haplotype score analyses (using the 'ihh2ihs' function) for each SNP within each breed population, employing the rehh package (Gautier and Vitalis, 2012), in R/Rstudio version 4.3.1 (Rstudio Team, 2020). SNP scores were converted into two-sided *p*-values using the formula piHS=−log10[1−2|Φ(iHS)−0.500|], where Φ(x) represents the Gaussian cumulative distribution function, assuming that iHS values followed a normal distribution (produced by associated Q-Q plots). A threshold of −log10(p-value)=3 was set to identify significant iHS values. A secondary Benjamini-Hochbery (Benjamini and Hochberg, 1995) correction test was applied to the significant SNPs detected within each population, to ensure reduction of FDR (False Discovery Rates) produced by iHS scores, only considering SNPs with *p*-values < 0.050. Candidate iHS regions were calculated using the 'calc_candidate_regions' function, whereby a window_size = 1 around each significant SNP, was set to include consecutive neighbouring SNPs for determination of candidate regions – a method endorsed by Voight et al. (2006).

5.3.4 Cross-population scan for selection signatures (XP-EHH)

The phased haplotype SNP data was used to determine differential selection across the indigenous and rural horse populations, with the rehh package (Gautier and Vitalis, 2012), in R/Rstudio version 4.3.1 (Rstudio Team, 2020), using the 'ies2xpehh' function. Cross-population extended haplotype homozygosity (XP-EHH)

statistical method (Sabeti et al., 2007) was applied based on XP-EHH = ln ($\frac{I_A}{I_B}$), whereby I_A denotes the EHH integration estimates pertaining to genetic distance in population A and I_B denotes the EHH estimate pertaining to genetic distance in population B. The same threshold of −log10(*p*-value)=3 was set to identify significant SNPs across the populations, whereby reduction of FDR produced by the XP-EHH scores were also ensured through Benjamini-Hochbery (Benjamini and Hochberg, 1995) correction test – only considering SNPs with *p*-values < 0.050. Candidate regions determined by the cross-population SNP scores were determined using the 'calc_candidate_regions' function through a window_size = 1 (Voight et al., 2006).

5.3.5 Candidate gene identification, functional annotation and enrichment analysis

QTL analysis was conducted in order to identify and functionally annotate adaptive selection signatures resulted from both iHS and XP-EHH scores of significance. Population-specific selection signatures yielded from the iHS analyses were assessed and compared to cross-population selection signatures from the XP-EHH analysis – in order to determine overlapping genomic regions. Ensembl genome browser (https://www.ensembl.org) (Harrison et al., 2024) was utilized to identify genes in which significant SNPs were mapped according to the EquCab3 genome assembly (Kalbfleisch et al., 2018). The identified genes were functionally annotated and enriched using the DAVID analysis platform (https://david.ncifcrf.gov/) (Sherman et al., 2022) whereby gene lists were uploaded and functions of adaptive potential were considered.

5.4 Results

5.4.1 Within population scan for selection signatures

Genome-wide scan for the distribution of iHS values spanning the genome for each horse breed population was conducted as shown in **Figure 5.2** below.

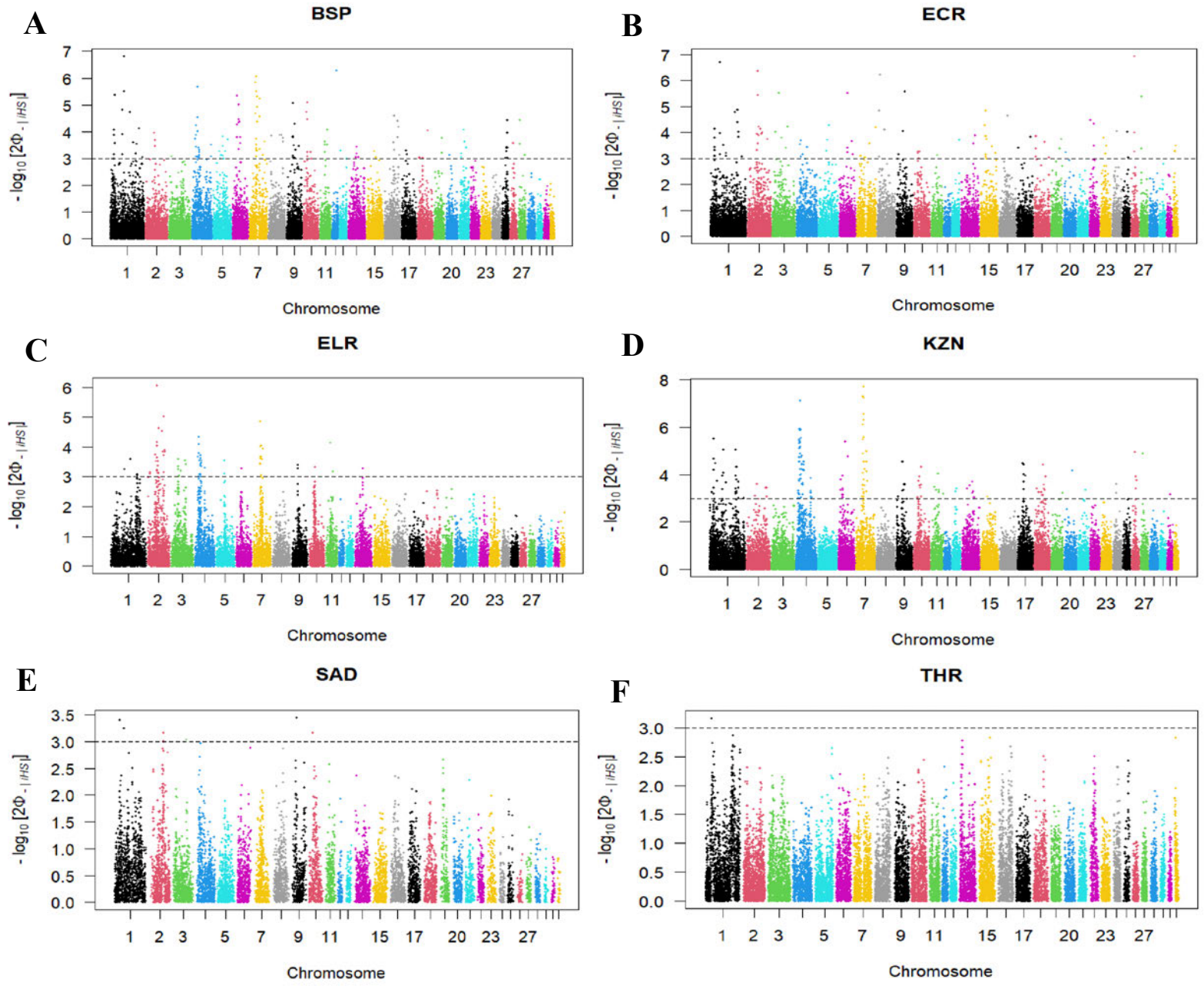


Figure 5.2 Genome-wide scan for distribution of iHS scores across six horse breed populations

[iHS scores of significant SNPs were detected above a threshold = 3 of the Indigenous (**A.** Basotho pony (BSP)); Rural (**B.** Eastern Cape Rural (ECR); **C.** East London Rural (ELR); **D.** KwaZulu-Natal Rural (KZN)) and Commercial (**E.** Saddlebred (SAD); **F.** Thoroughbred (THR) breed populations.]

The iHS score threshold shown above detected 143, 108, 81, 177, 7 and 1 selection signatures in BSP, ECR, ELR, KZN, SAD and THR breed populations - respectively. All identified selective sweeps exhibited positive iHS scores, indicating strong selection on the ancestral alleles, with no detection of negative iHS scores indicating selection of derived alleles. The presence of selection signatures exceeding the iHS threshold of 3 suggests very strong and recent positive selection within the breed populations. The distribution of selection signatures across the autosomes of each breed population were not distributed evenly, however, the indigenous

BSP and two rural populations (ECR and KZN) showed a wider range of significant SNPs across the genome. Strong selective sweeps within each breed population were considered based on larger SNP clusters detected across autosomes. Strong sweeps were detected on chromosomes 1, 4 and 7 within both the BSP and KZN breed populations, whereby a single SNP within the BSP population on chromosome 12 had a very high associated iHS score (6.30) than compared to the remaining detected SNPs. It was also observed that chromosome 4 within the KZN population, contained the most dense selective sweep of 49 SNPs. The ECR population had the largest range of significant SNPs across a total of 26 chromosomes than compared to the other horse breed populations, while only strong selective sweeps were detected on chromosomes 1 and 2. Fairly dense selective sweeps were found within the ELR population, whereby 23 SNPs clustered on chromosome 2 and 26 SNPs were clustered on chromosome 4. The two commercial breed populations contained very few significant selection signatures obtained through the iHS genome-wide scan, whereby only 5 SNPs (on chromosomes 1, 2, 3, 9 and 10) were detected within the SAD breed and only 1 SNP (on chromosome 1) within the THR breed. There were no strong selective sweeps detected within both commercial breed populations and were therefore not considered in the cross-population (XP-EHH) analyses.

5.4.2 Cross-population scan for selection signatures

To identify genomic regions under selection between two populations, XP-EHH analyses were presented in **Figure 5.3** and **Figure 5.4** below, comparing the indigenous and rural horse breed populations – whereby several significant selection sweeps were detected.

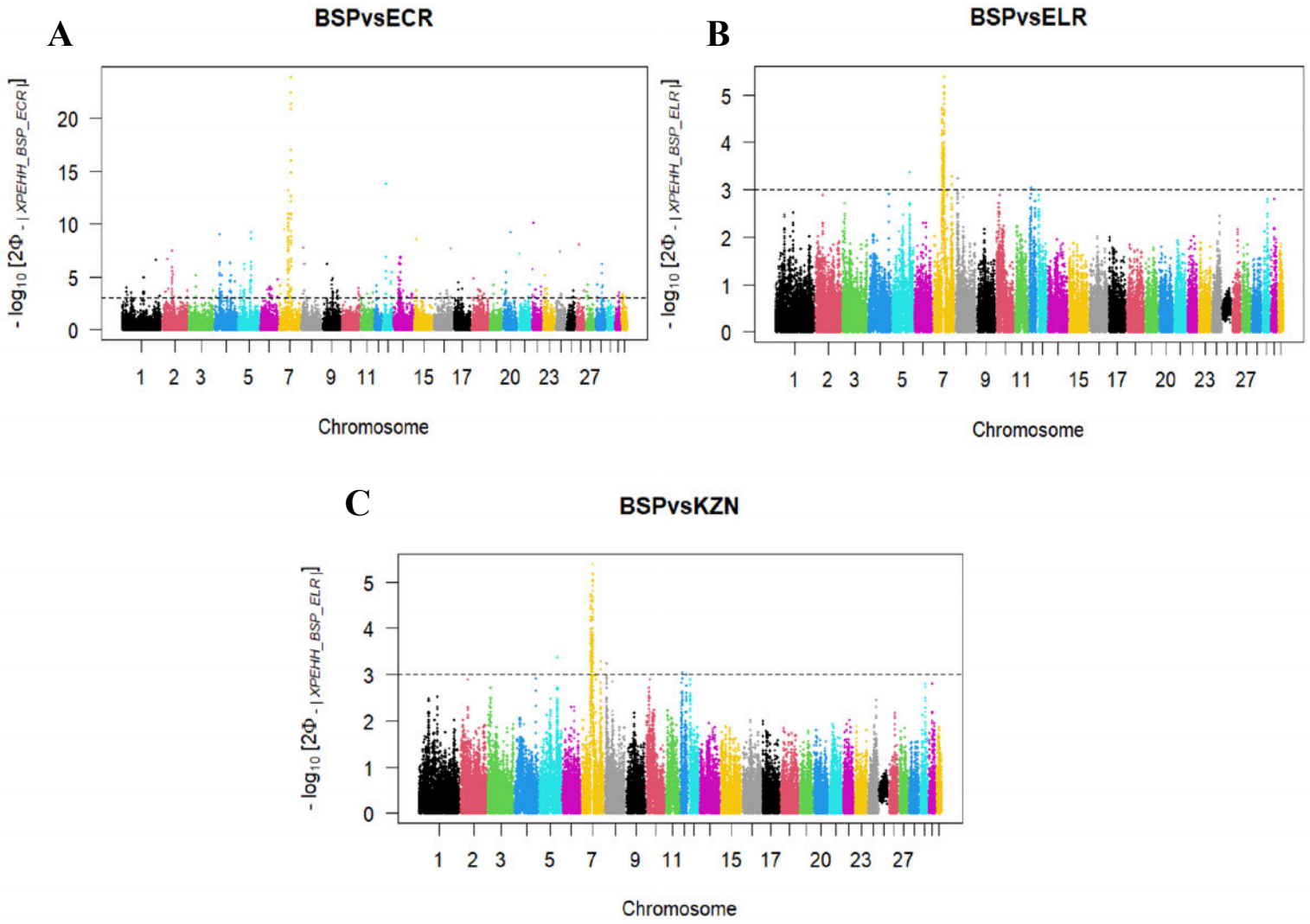


Figure 5.3 Genome-wide scan for distribution of XP-EHH scores across indigenous BSP breed and rural breed populations

[XP-EHH scores of significant SNPs were detected above threshold = 3. The indigenous Basotho Pony (BSP) was compared to **A**. Eastern Cape Rural (ECR), **B.** East London Rural (ELR) and **C.** KwaZulu-Natal (KZN) rural.]

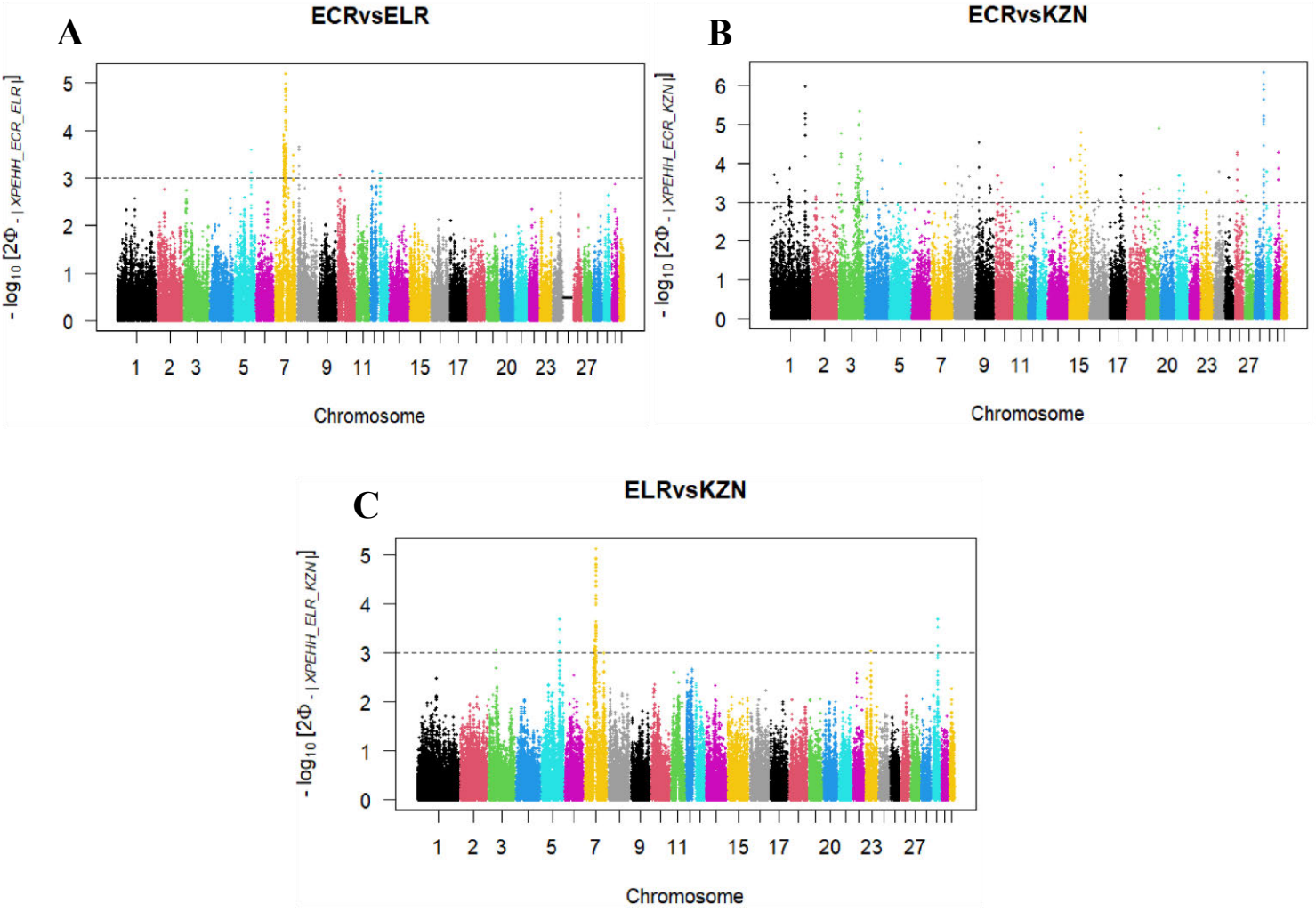


Figure 5.4 Genome-wide scan for distribution of XP-EHH scores across rural breed populations

[XP-EHH scores of significant SNPs were detected above threshold = 3. The three rural horse breed populations were compared as follows: **A**. Eastern Cape Rural (ECR) vs East London Rural (ELR), **B.** Eastern Cape Rural (ECR) vs KwaZulu-Natal (KZN) rural and **C.** East London Rural (ELR) vs KwaZulu-Natal (KZN) rural.]

The XP-EHH score threshold shown in the Manhattan plots in **Figure 5.3** above, detected 408, 227 and 227 significant SNPs between BSP vs ECR (Figure 2 (**A**)), BSP vs ELR (Figure 2 (**B**)) and BSP vs KZN (Figure 2 (**C**)) - respectively. One particular strong selective sweep was consistent across the three different XP-EHH analyses between the indigenous and three rural breed populations, detected on chromosome 7. These consistent selective sweeps contained dense clusters of SNPs, whereby BSP vs ELR and BSP vs KZN both contained 223 SNPs on chromosome 7. Another interesting finding from the XP-EHH analysis was that of the comparison between BSP and ECR breed populations that detected the largest accumulation of significant SNPs across the genome, that also contained smaller significant sweeps that were detected on chromosomes 2, 4, 5 and 19. The XP-EHH score threshold in the above plots (**Figure 5.4**) indicated cumulative significant selection signatures totaling 186, 168 and 75 SNPs within the cross-population comparison between ECR vs

ELR (Figure 3 (**A**), ECR vs KZN (Figure 3 (**B**)) and ELR vs KZN (Figure 3 (**C**) – respectively. The very same observation of consistent selective sweep clusters were seen within **A** and **C** on chromosome 7 only, where the rural populations in **A** clustered 176 significant SNPs and the rural populations in **B** clustered 64 significant SNPs – KZN rural breed population in common amongst the analyses. The XP-EHH analysis that differed from this observation was the comparison of the ECR and KZN populations (**B**), whereby multiple strong selection sweeps were detected on chromosomes 1, 3, 15 and 28 – spanning over a wider range of 25 chromosomes in which contained SNPs above the XP-EHH score threshold.

5.4.3 Candidate gene identification, functional annotation and enrichment analysis

Population-specific selection signatures obtained from the iHS analyses, associated with traits that were adaptive in nature, were identified within each breed populations in **Tables 5.1** – **5.6** below. Functional annotations were documented as well as enrichment analyses, to determine pathways that provide insight into the extent of positive selective pressures, undergone in each breed population. Cross-population selection signatures obtained from the XP-EHH analyses between the indigenous and rural breed populations were functionally annotated and enriched within **Tables 5.7** – **5.12**. Intersected regions yielded between the iHS and XP-EHH analyses in **Table 5.13**, demonstrated adaptive selection signatures fixed within the breed populations.

Table 5.1 Gene identification, functional annotation and enrichment of iHS selection signatures in BSP breed

iHS Selection Regions						Gene Ontology Enrichment			
SNP Name	**Chr**	**Start SNP**	**End SNP**	**iHS Score**	**Gene Name**	**GO Term BP**	**GO Term CC**	**GO Term MF**	**KEGG Pathway**
BIEC2_11874	1	26069383	26069383	5.3883346	*CFAP58*	Sperm axoneme assembly	Sperm Midpiece		
BIEC2_11887	1	26106108	26106108	3.5856224	*CFAP58*	Sperm axoneme assembly	Sperm Midpiece		
BIEC2_27092	1	65491273	65491273	4.8311874	*DLG5*	Epithelial to mesenchymal transition	Plasma membrane	Protein binding	Hippo signalling pathway
BIEC2_32642	1	76385429	76385429	6.8246241	*NID1*	Cell-matrix adhesion	Extracellular region	Calcium ion binding	Cytoskeleton in muscle cells
CUHSNP00055856	1	125458890	125458890	3.5925591	*CORO2B*	Actin filament organization	Actin filament binding		
BIEC2_60393	1	137657362	137657362	3.2003424	*UNC13C*	Neuromuscular junction development	Plasma membrane	Calcium ion binding	Synaptic vesicle cycle
BIEC2_62289	1	142316661	142316661	3.537044	*FBN1*	Skeletal muscle development	Microfibril	Integrin binding	TGF-beta signalling pathway
BIEC2_68032	1	156462763	156462763	4.1185867	*OR4F14S*	G protein-coupled receptor signalling pathway	Plasma membrane	Olfactory receptor activity	Olfactory transduction
BIEC2_477984	2	52282449	52282449	3.4720522	*LOXL2*	Negative regulation of transcription by RNA polymerase II	Chromatin	Copper ion binding	
BIEC2_896620	4	20745394	20745394	4.2389126	*COBL*	Embryonic axis specification	Ruffle	Actin monomer binding	
BIEC2_852714	4	20818822	20818822	3.44936	*COBL*	Embryonic axis specification	Ruffle	Actin monomer binding	
BIEC2_901872	4	28190874	28190874	4.028417	*SEMA3E*	Branching involved in blood vessel morphogenesis	Extracellular space	Semaphorin receptor binding	Axon guidance
BIEC2_857492	4	28333817	28333817	4.560639	*SEMA3E*	Branching involved in blood vessel morphogenesis	Extracellular space	Semaphorin receptor binding	Axon guidance
TBIEC2_902071	4	28769871	28769871	3.079245	*SEMA3E*	Branching involved in blood vessel morphogenesis	Extracellular space	Semaphorin receptor binding	Axon guidance
BIEC2_857781	4	28890573	28890573	5.6694705	*SEMA3E*	Branching involved in blood vessel morphogenesis	Extracellular space	Semaphorin receptor binding	Axon guidance
UKUL935	4	35965073	35965073	3.4106022	*LRRD1*	Signal transduction	Cytoplasm	Protein serine/threonine phosphatase activity	
BIEC2_859105	4	38694617	38694617	3.1388576	*PPP1R9A*	Actin filament organization	Cytoplasm	Actin filament binding	
UKUL947	4	38974096	38974096	3.0449316	*PPP1R9A*	Actin filament organization	Cytoplasm	Actin filament binding	
BIEC2_859436	4	39259177	39259177	3.1413489	*DYNC1I1*	Transport along microtubule	Kinetochore	Spectrin binding	Phagosome
BIEC2-881610	4	103250250	103250250	3.5278979	*PRKAG2*	Glycogen metabolic process	Nucleus	ATP binding	FoxO signalling pathway
BIEC2_907147	5	42039852	42039852	3.1367348	*THEM4*	Lipid metabolic process	Cytoplasm	Hydrolase activity	Fatty acid elongation

BIEC2_956279	5	53084961	53084961	3.1187994	*KCND3*		Voltage-gated potassium channel complex	A-type potassium channel activity	Spinocerebellar ataxia
BIEC2_909018	5	53253842	53253842	3.478282	*INKA2*		Nucleus	Protein kinase binding	
BIEC2_961288	5	63826042	63826042	3.2602496	*SNX7*	Protein transport	Phagophore assembly site	Phosphatidylinositol binding	
BIEC2_946245	6	30490216	30490216	3.3667537	*ERC1*	Regulation of DNA-templated transcripton	Centrosome	Structural constituent of presynaptic active zone	NF-kappa B signalling pathway
BIEC2_992105	7	32759944	32759944	5.8513785	*VWA5A*		Nucleoplasm	G protein-coupled receptor activity	
TBIEC2_1049808	7	36032880	36032880	3.5347207	*CDON*	Cell fate specification	Plasma membrane	Protein binding	Hedgehog signalling pathway
BIEC2_996295	7	38457973	38457973	3.0706505	*ETS1*	Regulation of transcription by RNA polymerase II	Nucleus	DNA-binding transcription factor activity	RAS signalling pathway
BIEC2_1000179	7	53064035	53064035	4.5707981	*OR7D20*	Signal transduction	Plasma membrane	Olfactory receptor activity	Olfactory transduction
BIEC2_1055236	7	54042978	54042978	3.3685068	*OR1M1*	Signal transduction	Plasma membrane	Olfactory receptor activity	Olfactory transduction
BIEC2_1008338	7	81454348	81454348	3.6259576	*SBF2*		Cytoplasm	Guanyl-nucleotide exchange factor activity	
BIEC2_1030845	8	12557502	12557502	3.8432447	*MYO18B*	Actomyosin structure organization	Myosin complex	Cytoskeletal motor activity	Motor proteins
BIEC2_1112825	8	58635226	58635226	3.8707791	*DTNA*	Synaptic signalling	Nucleoplasm	Zinc ion binding	Cytoskeleton in muscle cells
BIEC2_1063705	8	79564639	79564639	3.6243093	*ZNF532*	Zinc Finger Protein	Nucleus		
TBIEC2_1144245	9	30276863	30276863	5.060099	*TCEA1*	DNA-template transcription	Nucleus	Nucleic-acid binding	
BIEC2_1084670	9	30450015	30450015	3.8189758	*ATP6V1H*	Endocytosis	Vacuolar proton-transporting V-type ATPase	Proton-transporting ATPase activity	Oxidative phosphorylation
BIEC2_1092034	9	47665041	47665041	3.328238	*SPAG1*	DNA-template transcription	RNA polymerase II core complex	RNA polymerase I activity	
BIEC2_110815	10	17331188	17331188	4.7484115	*ARHGAP35*	Neural tube closure	Nucleus	GTPase activity	Focal adhesion
TBIEC2_112310	10	19035346	19035346	3.2516764	*FUT1*	Carbohydrate metabolic process	Membrane	Galactoside 2-alpha-L-fucosyltransferase activity	Glycosphingolipid biosynthesis
BIEC2_144777	11	24375884	24375884	3.5164253	*SKAP1*	Positive regulation of cell-matrix adhesion	Immunological synapse	Protein binding	Rap1 signaling pathway
BIEC2_145162	11	25633158	25633158	3.2501299	*FAM117A*	Transmembrane transport	Membrane		
BIEC2_160534	11	39973565	39973565	3.5107178	*SPACA3*	Fusion of sperm to egg plasma membrane involved in single fertilization	Acrosomal vesicle	Lysozyme activity	
BIEC2_195284	12	25031305	25031305	6.2888655	*SAX04*				
BIEC2_259320	14	53417443	53417443	3.0454694	*DTWD2*	tRNA processing		transferase activity	
BIEC2_259320	14	53417443	53417443	3.0454694	*DMXL1*	Vacuolar acidification	RAVE complex	Protein binding	Lysosome
BIEC2_304186	15	38948763	38948763	3.2857925	*WDPCP*	Nervous system development	Plasma membrane	Protein binding	Cilium biogenesis/degradation
TBIEC2_364361	16	49012268	49012268	3.6131489	*ITGA9*	Cell-matrix adhesion	Integrin complex	Integrin binding	PI3K-Akt signalling pathway
BIEC2-350166	16	52986670	52986670	3.3332026	*TRIM71*	G1/S transition of mitotic cell cycle	P-body	RNA binding	MicroRNAs in cancer

BIEC2_357454	16	69405335	69405335	3.8098216	*CPNE4*	Cellular response to calcium ion	Plasma membrane	Calcium-dependent phospholipid binding	
BIEC2-362677	16	77015990	77015990	3.1448178	*PXYLP1*	Glycosaminoglycan biosynthetic process	Golgi membrane	Phosphate activity	
BIEC2_373856	17	22444969	22444969	3.2901866	*RCBTB2*	Regulator of chromosome condensation	Cytoplasm	Protein binding	
BIEC2_374436	17	26406722	26406722	3.165762	*LACC1*	Positive regulation of cytokine production involved in immune response	Nucleus	Adenosine deminase activity	Purine metabolism
BIEC2_523829	20	20080039	20080039	3.1293129	*CDKAL1*	tRNA methylthiolation	Endoplasmic reticulum	Metal ion binding	
BIEC2_550941	21	10002112	10002112	3.1979329	*HTR1A*	Behavioural fear response	Dendrite; synapse	Dopamine neurotransmitter receptor activity	cAMP signaling pathway
BIEC2_550941	21	10002112	10002112	3.1979329	*RNF180*	protein polyubiquitination	Endoplasmic reticulum	protein polyubiquitination	
BIEC2_581885	21	14276360	14276360	3.0447757	*PDE4D*	cAMP catabolic process	Nucleus	cAMP binding	Purine metabolism
BIEC2_561043	21	31107944	31107944	3.6216592	*AGXT2*	Glyoxylate catabolic process	Mitochondrion	Alanine-glyoxylate transaminase activity	Alanine
BIEC2_565510	21	40723491	40723491	3.4073018	*CDH12*	Cell morphogenesis	Plasma membrane	Calcium ion binding	
BIEC2_668387	25	29938227	29938227	4.4300309	*MAPKAP1*	Positive regulation of peptidyl-serine phosphorylation	Nucleoplasm	Phophatidylinositol-4,5-biphosphate binding	mTOR signaling pathway
BIEC2_668423	25	29994739	29994739	3.4145139	*MAPKAP1*	Positive regulation of peptidyl-serine phosphorylation	Nucleoplasm	Phophatidylinositol-4,5-biphosphate binding	mTOR signaling pathway
BIEC2_668497	25	30101253	30101253	3.0194577	*PBX3*	Eye development	Nucleus	DNA-binding transcription factor activity	Transcriptional misregulation in cancer
BIEC2_668555	25	30211964	30211964	3.4239786	*PBX3*	Eye development	Nucleus	DNA-binding transcription factor activity	Transcriptional misregulation in cancer
BIEC2-702094	27	7724890	7724890	4.4397751	*ADRB3*	Regulation of systemic arterial blood pressure	Plasma membrane	Beta3-adrenergic receptor activity	Calcium signaling pathway
BIEC2_718847	27	34493492	34493492	3.1231499	*MCPH1*	Negative regulation of transcription by RNA polymerase II	Centrosome	Identical protein binding	

[Contents of selection signatures for Basotho pony (BSP) include: **SNP name**; Chromosome number (**Chr**); SNP start position (**Start SNP**); SNP end position (**End SNP**); **iHS Score**; Official gene name (**Gene Name**); Gene Ontology Term Biological Process (**GO Term BP**); Gene Ontology Term Cellular Component (**GO Term CC**); Gene Ontology Term Molecular Function (**GO Term MF**) and **KEGG Pathway** (Kyoto Encyclopedia of Genes and Genomes pathway) – provided through DAVID analysis platform (Sherman et al., 2022).]

Table 5.2 Gene identification, functional annotation and enrichment of iHS selection signatures in ECR breed population

iHS Selection Regions						Gene Ontology Enrichment			
SNP Name	**Chr**	**Start SNP**	**End SNP**	**iHS Score**	**Gene Name**	**GO Term BP**	**GO Term CC**	**GO Term MF**	**KEGG Pathway**
BIEC2-35722	1	81894020	81894020	3.0997921	*OGDHL*	Tricarboxylic acid cycle	Mitochondrion	Oxoglutarate dehydrogenase activity	Citrate cycle
CUHSNP00055856	1	125458890	125458890	4.7969015	*CORO2B*	Actin filament organization		Actin filament binding	
TBIEC2_63426	1	136755421	136755421	4.8776772	*RAB27A*	Blood coagulation	Photoreceptor outer segment	GTPase activity	
BIEC2_60393	1	137657362	137657362	4.3896932	*UNC13C*	Neuromuscular	Plasma membrane	Calcium ion binding	
BIEC2_62289	1	142316661	142316661	3.8519779	*FBN1*	Skeletal system development	Microfibril	Integrin binding	Cytoskeleton in muscle cells
BIEC2_476151	2	44728809	44728809	3.9400026	*KCNAB2*	Potassium ion transport	Cytoplasm	NADPH activity	
BIEC2_477984	2	52282449	52282449	6.3773544	*LOXL2*	Negative regulation of transcription by RNA polymerase II	Chromatin	Copper ion binding	
BIEC2_478902	2	54795464	54795464	4.2202392	*EBF2*	Cell fate determination	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
TBIEC2_504258	2	56292401	56292401	3.8954398	*PTK2B*	Positive regulation of cell-matrix adhesion	Nucleus	Protein tyrosine kinase activity	Calcium signaling pathway
BIEC2_483107	2	62780628	62780628	4.1368334	*GALNT7*	Glycosylation	Golgi membrane	Carbohydrate binding	Mucin type O-glycan biosynthesis
BIEC2_485358	2	67390833	67390833	3.9547473	*ANXA10*		Cytoplasm	Phosphatidylserine binding	
BIEC2_488689	2	73061787	73061787	3.9973144	*FSTL5*	Cell differentiation	Extracellular region	Calcium ion binding	
BIEC2_896620	4	20745394	20745394	3.7049164	*COBL*	Embryonic axis specification	Ruffle	Actin monomer binding	
BIEC2_852714	4	20818822	20818822	3.0847991	*COBL*	Embryonic axis specification	Ruffle	Actin monomer binding	
BIEC2_857781	4	28890573	28890573	3.1583366	*SEMA3A*	Neural crest cell migration	Extracellular space	Semaphorin receptor binding	Axon guidance
BIEC2_865322	4	50841739	50841739	3.4460499	*HDAC9*	Response to insulin stimulus	Histone deacetylase complex	Histone deactylase activity	Neutrophil extracellular trap formation
BIEC2_949848	6	39700007	39700007	5.5318611	*STYK1*	Protein phosphorylation	Plasma membrane	Protein tyrosine kinase activity	
BIEC2-1179258	6	63750000	63750000	3.6568006	*NELL2*	Commissural neuron axon guidance	Extracellular space	Calcium ion binding	
BIEC2_1015233	6	69311499	69311499	3.1667828	*SLC11A2*	Response to hypoxia	Nucleus	Copper ion transmembrane transporter activity	Lysosome
BIEC2_984449	7	19767028	19767028	3.1024806	*ARHGAP20*	Signal transduction		GTPase activator activity	
BIEC2_1030306	8	12170083	12170083	4.8517647	*SEZ6L*	Synapse maturation	Endoplasmic reticulum		
BIEC2_1030845	8	12557502	12557502	6.2282417	*MYO18B*	Actomyosin structure organization	Cytoplasm	Cytoskeletal motor activity	Motor proteins

BIEC2_1043648	8	34982846	34982846	4.1256769	*ARHGAP28*	Signal transduction	Cytoplasm	GTPase activator activity	
BIEC2_1089591	9	40426753	40426753	5.5870956	*RUNX1T1*	DNA-templated transcription	Nucleus	Transcription corepressor activity	Pathways in cancer
BIEC2_108373	10	21199565	21199565	3.2767395	*IGLON5*		Neuron projection	Signal	
BIEC2_148248	11	30016859	30016859	3.1468769	*STXBP4*	Protein targeting	Synapse	Syntaxin binding	
BIEC2_312541	15	59977955	59977955	3.4594676	*SOS1*	Hair follicle development	Cytoplasm	Guanyl-nucleotide exchange factor activity	Endocrine resistance
BIEC2_312541	15	59977955	59977955	3.4594676	*ARHGEF33*	Guanyl-nucleotide exchange factor activity			
BIEC2_342973	16	40811163	40811163	4.6583185	*SMARCC1*	Nervous system development	Male germ cell nucleus	Chromatin binding	Thermogenesis
BIEC2_616227	23	17725279	17725279	3.8004466	*TMC1*	Detection of mechanical stimulus involved in sensory perception of sound	Plasma membrane	Voltage-gated calcium channel activity	
CUHSNP00119256	23	27246782	27246782	3.508611	*GLDC*	Glycine decarboxylation via glycine cleavage system	Nucleoplasm	Glycine dehydrogenase activity	Glycine, serine and threonine metabolism
BIEC2_668387	25	29938227	29938227	3.0601204	*MAPKAP1*	Positive regulation of peptidyl-serine phosphorylation	Nucleoplasm	Phosphatidylinositol-4,5-bisphosphate binding	mTOR signaling pathway
BIEC2_878783	31	9821709	9821709	3.30534	*TULP4*	Protein ubiquitination	Cytoplasm	Protein binding	
BIEC2_835465	31	9861318	9861318	3.2771958	*GTF2H5*	Maturation of SSU-rRNA from rRNA Transcript	Transcription factor TFIIH core complex		Basal transcription factors

[Contents of selection signatures for Eastern Cape Rural (ECR) include: **SNP name**; Chromosome number (**Chr**); SNP start position (**Start SNP**); SNP end position (**End SNP**); **iHS Score**; Official gene name (**Gene Name**); Gene Ontology Term Biological Process (**GO Term BP**); Gene Ontology Term Cellular Component (**GO Term CC**); Gene Ontology Term Molecular Function (**GO Term MF**) and **KEGG Pathway** (Kyoto Encyclopedia of Genes and Genomes pathway) – provided through DAVID analysis platform (Sherman et al., 2022).]

Table 5.3 Gene identification, functional annotation and enrichment of iHS selection signatures in ELR breed population

iHS Selection Regions						Gene Ontology Enrichment			
SNP NAME	**Chr**	**Start SNP**	**End SNP**	**iHS Score**	**Gene Name**	**GO Term BP**	**GO Term CC**	**GO Term MF**	**KEGG Pathway**
BIEC2_61219	1	139162818	139162818	3.0951363	*MYO5A*	Actin filament organization	Ruffle	Microfilament motor activity	Motor proteins
AX-103516755	1	140632424	140632424	3.0641956	*SPPL2A*	Membrane protein proteolysis	Lysosomal membrane	Aspartic endopeptidase activity	
BIEC2_457411	2	13603406	13603406	3.1560262	*ElF2B3*	Protein synthesis			
AX-103788096	2	52052990	52052990	3.2074763	*RHOBTB2*	Endocytosis	Nucleus	GTPase activity	Ubiquitin mediated proteolysis
BIEC2_477984	2	52282449	52282449	6.0698713	*LOXL2*	Negative regulation of transcription by RNA polymerase II	Chromatin	Copper ion binding	
BIEC2_477985	2	52282521	52282521	6.0698713	*LOXL2*	Negative regulation of transcription by RNA polymerase II	Chromatin	Copper ion binding	
AX-103430264	2	55170630	55170630	3.1431627	*PPP2R2A*	Response to morphine	Protein phosphatase type 2A complex	Protein binding	mRNA surveillance pathway
BIEC2-481775	2	60253200	60253200	4.6475283	*NEIL2*	Base-excision repair	Nucleus	Damaged DNA binding	Base excision repair
BIEC2_494883	2	86298818	86298818	3.1616241	*TTC29*	Cilium movement	Cytoplasm	Protein binding	
BIEC2_521853	2	89437060	89437060	3.7337226	*INPP4B*		Cytoplasm	Phosphatidylinositol-3,4-bisphosphate 4-phosphatase activity	Inositol phosphate metabolism
BIEC2_497450	2	90849497	90849497	3.9023674	*RNF150*	Ubiquitin-dependent protein catabolic process	Cytoplasm	Metal ion binding	
TBIEC2_523156	2	91300136	91300136	3.8559507	*CLGN*	Protein folding	Endoplasmic reticulum	Calcium ion binding	
BIEC2_777818	3	37540091	37540091	3.2513008	*SLC9B2*	Siderophore biosynthetic process	Cytoplasm	3-hydroxybutyrate dehydrogenase activity	
BIEC2_779089	3	44079562	44079562	3.0298568	*UNC5C*	Axon guidance	Plasma membrane	Netrin receptor activity	Axon guidance
BIEC2_779941	3	50474323	50474323	3.2036409	*FAM13A*	Signal transduction			
TBIEC2_901715	4	27759665	27759665	3.5462498	*PCLO*	Cytoskeleton organization	Postsynaptic density	Protein binding	Insulin secretion
TBIEC2_901719	4	27802674	27802674	3.5462498	*PCLO*	Cytoskeleton organization	Postsynaptic density	Protein binding	Insulin secretion
BIEC2-858405	4	32341420	32341420	3.0772357	*ADAM22*	Adult locomotory behaviour	Membrane	Metalloendopeptidase	
BIEC2_858407	4	32365745	32365745	3.0772357	*ADAM22*	Adult locomotory behaviour	Membrane	Metalloendopeptidase	
BIEC2_858415	4	32455442	32455442	3.0772357	*ADAM22*	Adult locomotory behaviour	Membrane	Metalloendopeptidase	
TBIEC2_902898	4	33055005	33055005	3.4751422	*ZNF804B*		Nucleus	Metal ion binding	
AFFX-102364083	4	54792728	54792728	3.3017286	*HYCC1*	Phosphatidylinositol phosphate biosynthetic process	Cytoplasm		

BIEC2_907545	5	45039908	45039908	3.116309	*BCL9*	Positive regualtion of transcription by RNA polymerase II	cis-Golgi network	Transcription coactivator activity	
AX-103724001	6	30433828	30433828	3.2845693	*ERC1*	Neuromuscular synaptic transmission	Centrosome	Structural constituent of presynaptic active zone	NF-kappa B signaling pathway
BIEC2_996661	7	41161595	41161595	3.4604689	*NTM*		Neuron projection	Signal	
BIEC2_996678	7	41340513	41340513	3.6773508	*NTM*		Neuron projection	Signal	
AFFX-101973792	7	41508901	41508901	4.8540051	*NTM*		Neuron projection	Signal	
BIEC2_996708	7	41826365	41826365	4.0486124	*OPCML*		Neuron projection	Signal	
AFFX-102814641	7	42384587	42384587	3.0701637	*OPCML*		Neuron projection	Signal	
AFFX-101244228	9	32668821	32668821	3.4127407	*PXDNL*	Response to oxidative stress	Extracellular space	Peroxidase activity	
BIEC2_1086540	9	33380282	33380282	3.2882469	*SNTG1*		Intracellular anatomical structure	Structural molecule activity	
BIEC2_145162	11	25633158	25633158	4.1329311	*FAM117A*	Transmembrane transport	Membrane		
BIEC2_150360	11	35882662	35882662	3.1894089	*USP32*	Proteolysis	Golgi apparatus	Calcium ion binding	
BIEC2_257184	14	41929389	41929389	3.2925311	*ZCCHC10*			Nucleic acid binding	

[Contents of selection signatures for East London Rural (ELR) include: **SNP name**; Chromosome number (**Chr**); SNP start position (**Start SNP**); SNP end position (**End SNP**); **iHS Score**; Official gene name (**Gene Name**); Gene Ontology Term Biological Process (**GO Term BP**); Gene Ontology Term Cellular Component (**GO Term CC**); Gene Ontology Term Molecular Function (**GO Term MF**) and **KEGG Pathway** (Kyoto Encyclopedia of Genes and Genomes pathway) – provided through DAVID analysis platform (Sherman et al., 2022).]

Table 5.4 Gene identification, functional annotation and enrichment of iHS selection signatures in KZN breed population

iHS Selection Regions						Gene Ontology Enrichment			
SNP Name	**Chr**	**Start SNP**	**End SNP**	**iHS Score**	**Gene Name**	**GO Term BP**	**GO Term CC**	**GO Term MF**	**KEGG Pathway**
BIEC2_12903	1	29093012	29093012	4.5032774	*SEMA4G*	Neural crest cell migration	Extracellular space	Semaphorin receptor binding	Axon guidance
BIEC2_19162	1	41014892	41014892	3.7503779	*ATAD1*	Extraction of mislocalized protein from mitochondrial outer membrane	Mitochondrial outer membrane	ATP binding	
AX-104408022	1	85005714	85005714	3.6872751	*GRID1*	Social behaviour	Plasma membrane	Glutamate-gated receptor activity	Neuroactive ligand-receptor interaction
BIEC2_59216	1	135802169	135802169	3.9513295	*MNS1*	Cilium organization	Nucleus		
AFFX-102868384	1	135848841	135848841	3.2195977	*MNS1*	Cilium organization	Nucleus		
AFFX-102173515	1	136698858	136698858	3.2712489	*PIGB*	GPI anchor biosynthetic process	Endoplasmic reticulum membrane	alpha-1,2-mannosyltransferase activity	Glycosylphosphatidylinositol (GPI)-anchor biosynthesis
TBIEC2_63426	1	136755421	136755421	5.0691326	*RAB27A*	Blood coagulation	Photoreceptor outer segment	GTPase activity	
AFFX-102844313	1	137550423	137550423	3.3039452	*UNC13C*	Neuromuscular junction development	Plasma membrane	Calcium ion binding	Synaptic vesicle cycle
BIEC2_60473	1	137759895	137759895	3.2290741	*UNC13C*	Neuromuscular junction development	Plasma membrane	Calcium ion binding	Synaptic vesicle cycle
AFFX-102945407	1	149158020	149158020	3.4520707	*DLL4*	Cell communication	Plasma membrane	Notch binding	Endocrine resistance
BIEC2_63964	1	149312024	149312024	3.7154165	*RMDN3*		Nucleus	Microtubule binding	
BIEC2_476151	2	44728809	44728809	3.1088416	*KCNAB2*	Potassium ion transport	Cytoplasm	NADPH activity	
BIEC2_850048	4	14168974	14168974	3.2233136	*HECW1*	Protein polyubiguitination	Cytoplasm	ubiquitin-protein transferase activity	
BIEC2_850142	4	14237452	14237452	3.3393052	*HECW1*	Protein polyubiguitination	Cytoplasm	ubiquitin-protein transferase activity	
BIEC2_851323	4	17459988	17459988	3.7342032	*TNS3*	Dephosphorylation	Cytosol	phosphoprotein phosphatase activity	
BIEC2_851337	4	17476563	17476563	5.4969171	*TNS3*	Dephosphorylation	Cytosol	phosphoprotein phosphatase activity	
BIEC2_851361	4	17504389	17504389	4.5589155	*TNS3*	Dephosphorylation	Cytosol	phosphoprotein phosphatase activity	
AFFX-101927377	4	18614948	18614948	5.3287284	*ABCA13*	Lipid transport	Membrane	Lipid transporter activity	ABC transporters

BIEC2-852529	4	20572040	20572040	5.9423577	*GRB10*	Insulin receptor signaling pathway	Protein-containing complex	Insulin receptor binding	mTOR signaling pathway
BIEC2_896620	4	20745394	20745394	5.8844023	*COBL*	Embryonic axis specification	Ruffle	Actin monomer binding	
BIEC2_852714	4	20818822	20818822	3.9788878	*COBL*	Embryonic axis specification	Ruffle	Actin monomer binding	
BIEC2-856985	4	26936480	26936480	5.5708761	*CACNA2D1*	Regulation of ventricular cardiac muscle cell membrane repolarization	Sarcoplasmic reticulum	Voltage-gated calcium channel activity	MAPK siganling pathway
TBIEC2_901419	4	27027219	27027219	3.8317312	*CACNA2D1*	Regulation of ventricular cardiac muscle cell membrane repolarization	Sarcoplasmic reticulum	Voltage-gated calcium channel activity	
BIEC2_857781	4	28890573	28890573	4.462605	*SEMA3A*	Neural crest cell migration	Extracellular space	Semaphorin receptor binding	Axon guidance
BIEC2_902344	4	29496419	29496419	4.8190175	*SEMA3D*	Neural crest cell migration	Extracellular space	Semaphorin receptor binding	Axon guidance
BIEC2_902678	4	31120307	31120307	3.6101333	*GRM3*	Gene expression	Postsynaptic density	Calcium channel regualtor activity	Neuroactive ligand-receptor interaction
AFFX-102811977	4	31212331	31212331	4.2541738	*GRM3*	Gene expression	Postsynaptic density	Calcium channel regualtor activity	Neuroactive ligand-receptor interaction
UKUL935	4	35965073	35965073	3.1851933	*LRRD1*	Signal transduction	Cytoplasm	Protein serine/threonine phosphatase activity	
BIEC2_859105	4	38694617	38694617	4.4293006	*PPP1R9A*	Actin filament organization	Cytoplasm	Actin filament binding	
BIEC2_859436	4	39259177	39259177	4.5493454	*DYNC1I1*	Transport along microtubule	Kinetochore	Sepctrin binding	Phagosome
CUHSNP00044115	4	39418444	39418444	4.1941499	*DYNC1I1*	Transport along microtubule	Kinetochore	Sepctrin binding	Phagosome
BIEC2_860592	4	41057196	41057196	4.1781876	*C1GALT1*	Protein glycosylation	Membrane	glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase activity	Mucin type O-glycan biosynthesis
AFFX-103081784	4	74106301	74106301	3.0137044	*MET*	Liver development	Semaphorin receptor comlplex	Protein tyrosine kinase activity	EGFR tyrosine kinase inhibitor resistance
AFFX-103087226	4	74353786	74353786	3.8482819	*ST7*		Membrane		
BIEC2_870012	4	74469174	74469174	3.0091807	*ST7*		Membrane		
AX-104432924	4	77389058	77389058	3.2601633	*TSPAN12*	Angiogenesis	Plasma membrane		

BIEC2_954044	6	49422515	49422515	4.782015	*C2CD5*	Insulin receptor signaling pathway	Cytosol	Calcium ion binding	
BIEC2_992105	7	32759944	32759944	7.3210551	*VWA5A*		Nucleoplasm	Olfactory receptor activity	
AFFX-102204260	7	32791602	32791602	3.5347874	*VWA5A*		Nucleoplasm	Olfactory receptor activity	
BIEC2_1051037	7	39505485	39505485	7.2638261	*TMEM45B*		Membrane		
BIEC2-996512	7	39521780	39521780	3.7882074	*TMEM45B*		Membrane		
AX-104935820	7	40865357	40865357	3.217279	*NTM*		Neuron projection		
BIEC2_1051200	7	41236063	41236063	5.5928385	*NTM*		Neuron projection		
AX-103156509	7	41531041	41531041	4.5648319	*NTM*		Neuron projection		
AFFX-102814641	7	42384587	42384587	3.5596311	*OPCML*		Plasma membrane		
BIEC2_1084682	9	30463606	30463606	3.3428395	*ATP6V1H*	Endocytosis	Vacuolar proton-transporting V-type ATPase	Proton-transporting ATPase activity	Oxidative phosphorylation
BIEC2_1085924	9	32676318	32676318	4.537677	*PXDNL*	Response to oxidative stress	Extracellular space	Peroxidase activity	
TBIEC2_122198	10	38005968	38005968	3.0690979	*ME1*	Pyruvate metabolic process	Mitochondrian	Magnesium ion binding	Pyruvate metabolism
BIEC2_141274	11	16407409	16407409	3.4891675	*MARCHF10*			Zinc ion binding	
BIEC2_214653	13	18527832	18527832	3.3115125	*TPST1*	peptidyl-tyrosine sulfation	Golgi membrane	protein-tyrosine sulfotransferase activity	
AFFX-101608216	14	61850125	61850125	3.3042059	*PJA2*	Inflammatory response	Cytoplasm	protein kinase A catalytic subunit binding	
BIEC2_301463	15	32218561	32218561	3.087179	*DYSF*	Plasma membrane repair	Plasma membrane	Calcium ion binding	
BIEC2_374768	17	28436592	28436592	3.7178305	*VWA8*		Cytoplasm	ATP binding	
BIEC2_410567	18	34492299	34492299	3.263952	*FMNL2*	Regulation of cell shape	Cytosol	Small GTPase binding	Cytoskeleton in muscle cells
BIEC2_411547	18	41014312	41014312	3.1489257	*RBMS1*		Nucleus	mRNA binding	
BIEC2_432909	18	41799430	41799430	4.4347772	*TANK*	Canonical NF-kappaB signal transduction		Metal ion binding	Autophagy - animal
AX-103317211	18	43190971	43190971	3.1521217	*KCNH7*	Regulation of membrane potential	Plasma membrane	Inward rectifier potassium channel activity	
BIEC2_433499	18	47166805	47166805	3.4234176	*XIRP2*	Ventricular septum development	Stress fiber	Metal ion binding	Cytoskeleton in muscle cells
AX-103807599	20	38964965	38964965	4.1810555	*BTBD9*	Long-term memory	Cytoplasm	Protein binding	
AX-104188551	24	18621062	18621062	3.6053047	*ZFYVE1*	Lipid droplet formation	Lipid droplet	1-phosphatidylinositol binding	

AFFX-102024681	24	18707699	18707699	3.2429219	*RBM25*	Regulation of alternative mRNA splicing via spliceosome	Spliceosomal complex	mRNA binding	Autophagy - animal
BIEC2_702093	27	7724881	7724881	4.8936092	*ADRB3*	Norepinephrine-epinephrine-mediated vasodilation involved in regulation of systemic arterial blood pressure	Plasma membrane	beta3-adrenergic receptor activity	Calcium signaling pathway
BIEC2_823413	30	16972599	16972599	3.1639615	*USH2A*	Animal organ morphogenesis	Basement membrane	Protein binding	

[Contents of selection signatures for KwaZulu-Natal (KZN) Rural include: **SNP name**; Chromosome number (**Chr**); SNP start position (**Start SNP**); SNP end position (**End SNP**); **iHS Score**; Official gene name (**Gene Name**); Gene Ontology Term Biological Process (**GO Term BP**); Gene Ontology Term Cellular Component (**GO Term CC**); Gene Ontology Term Molecular Function (**GO Term MF**) and **KEGG Pathway** (Kyoto Encyclopedia of Genes and Genomes pathway) – provided through DAVID analysis platform (Sherman et al., 2022).]

Gene identification and enrichment of potentially positively selected regions within the genomes of the indigenous, rural and commercial breed populations detected strong adaptive traits and associated pathways within selective sweeps and regions of population-specific traits. The Gene Ontology Enrichment analysis releaved the biological processes (**GO Term BP**) the detected genes were involved within, the cellular components (**GO Term CC**) where gene functions were localized, and the molecular functions (**GO Term MF**) that the gene products perform. Additionally, the KEGG pathway analysis (**KEGG Pathway**) identified the functional and metabolic pathways in which these gene products participate. BSP gene identification (**Table 5.1**) of adaptive traits found within selective sweep on chromosome 1 associated with fertility (*CFAP58* – two detection signals), nervous system development (*UNC13*), growth and development (*FBN1*), cellular processes (*DLG5*; *NID1*) and olfactory response system (*OR4F14S*). Selective sweep on chromosome 4 contained adaptive traits associated with fertility (*COBL* – two detection signals), feed efficiency (*PRKAG2*) and blood vessel morphogenesis (*SEMA3E* – three detection signals). Chromosome 7 selective sweep contained adaptive traits associated with cellular processes (*VWA5A*; *CDON*; *ETS1*; *SBF2*) and olfactory receptors (*OR7D20*; *OR1M1*). Apart from the selective sweeps and associated adaptive traits, an interesting trait for eye development (PBX3 – two detection signals) was found on chromosome 25. The KZN rural breed population also had selective sweeps detected on chromosomes 1, 4 and 7 (**Table 5.4**) that had very similar adaptive traits and biological pathways of cellular processes (*SEMA4G*; *MNS1*; *ATAD1*; *PIGB*; *DLL4*) found within the indigenous BSP population. However, on chromosome 1, adaptive traits associated with blood coagulation (*RAB27A*) and social behaviour (*GRID3*) were annotated that differed from BSP breed population. On chromosome 4, rural KZN breed population had two signals of detection for fertility (*COBL*), similarly to BSP breed population. Selection sweep on chromosome 7 in the KZN rural breed population similarly contained adaptive traits associated biological pathways pertaining to cellular processes (*TMEM45B* – two detection signals; *NTM* – three detection signals; *OPCML*) and olfactory receptors (*VWA5A* – two detection signals). Distinct selection signatures pertaining to liver development (*MET*) on chromosome 4 (within the selection peak), response to oxidative stress (*PXDNL*) on chromosome 9 and long-term memory (*BTBD9*) on chromosome 14 were detected within the KZN breed population.

The rural ECR population exhibited adaptive traits and biological pathways (**Table 5.2**) in the selective sweep on chromosome 1, associated with cellular processes (*OGDHL*; *CORO2B*), blood coagulation (*RAB27A*), neurological development (*UNC13C*) and skeletal system development (FBN1). The selective sweep on chromosome 2 contained traits associated specifically with cellular processes (*KCNAB2*; *LOXL2*; *EBF2*; *PTK2B*; *GALNT7*; *ANXA10*; *FSTL5*). Unique selection signatures were detected within this population, for traits associated with hair follicle development (*SOS1*) on chromosome 15 and response to mechanical input involved in auditory perception (*TMC1*) on chromosome 23. The rural ELR breed population exhibited adaptive selection signatures (**Table 5.3**) in

the selection sweep accumulated on chromosome 2 associated with cellular processes (*EIF2B3*; *RHOBTB2*; *LOXL2* – two detection signals; *NEIL2*; *TTC29*; *INPP4B*; *RNF150*; *CLGN*) as well as cellular process- related traits on the selection sweep within chromosome 4 (*PCLO* – two detection signals; *ZNF804B*; *HYCC1*). Unique selection signatures detected within the ELR population pertained to traits associated with adult locomotory behaviour (*ADAM22*) on chromosome 4 and pathways (mRNA surveillance pathway) associated with response to morphine in the system *(PPP2R2A*) on chromosome 9.

Both commercial breed populations did not present any selection sweeps in **Figure 5.2**, therefore resulting in only minor adaptive selection signatures associated with cellular processes (*MYOF*; *PALD1*; *GRIA2*; *GLRB*; *RP59*; *CH8*; *RPS9*) in the Saddlebred (SAD) breed population and cellular processes (*HECTD2*) in the Thoroughbred (THR) breed population.

Cross-population selection signatures obtained from the XP-EHH analyses, associated with traits that were adaptive in nature, were identified between indigenous and rural breed populations in **Tables 5.5 – 5.10** below. Functional annotations were documented as well as enrichment analyses, to determine pathways that provide insight into the extent of positive selective pressures, undergone between breed populations.

Table 5.5 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across indigenous BSP and rural ECR breed populations

XP-EHH Selection Regions					Gene Ontology Enrichment				
SNP Name	**Chr**	**Start SNP**	**End SNP**	**XP-EHH Score**	**Gene Name**	**GO Term BP**	**GO Term CC**	**GO Term MF**	**KEGG Pathway**
BIEC2_480833	2	1E+07	1E+07	3.6306	*EIF2B3*	Cytoplasmic translational initiation	Cytosol	Translation initiation factor activity	Herpes simplex virus 1 infection
BIEC2_476930	2	5E+07	5E+07	4.7913	*MEGF6*	Heterophilic cell-cell adhesion via plasma membrane cell adhesion molecules		Scavenger receptor activity	
BIEC2_476935	2	5E+07	5E+07	5.5698	*MEGF6*	Heterophilic cell-cell adhesion via plasma membrane cell adhesion molecules		Scavenger receptor activity	
AFFX-101627629	2	5E+07	5E+07	4.6329	*MEGF6*	Heterophilic cell-cell adhesion via plasma membrane cell adhesion molecules		Scavenger receptor activity	
AFFX-102672468	2	5E+07	5E+07	4.9317	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
BIEC2_476951	2	5E+07	5E+07	4.9317	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
BIEC2-476920_ILMNDUP1	2	5E+07	5E+07	4.9317	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
BIEC2_476953	2	5E+07	5E+07	7.5402	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
BIEC2_476965	2	5E+07	5E+07	3.562	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
BIEC2_476987	2	5E+07	5E+07	5.0137	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
BIEC2_476988	2	5E+07	5E+07	3.9335	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
BIEC2_476991	2	5E+07	5E+07	3.9335	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
BIEC2_477005	2	5E+07	5E+07	3.9335	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
BIEC2_477015	2	5E+07	5E+07	4.4376	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
BIEC2_477045	2	5E+07	5E+07	4.9748	*PRDM16*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Thermogenesis
TBIEC2_504638	2	6E+07	6E+07	3.1716	*ELP3*	tRNA wobble base 5-methoxycarbonylmethyl-2-thiouridinylation	Nucleus	tRNA binding	
BIEC2_510342	2	1E+08	1E+08	3.67	*NPNT*	Branching involved in ureteric bud morphogenesis	Extracellular region	Integrin binding	ECM-receptor interaction
BIEC2_844359	4	4E+06	4E+06	3.1082	*FAM185A*		Extracellular space		
BIEC2_852714	4	2E+07	2E+07	6.33	*COBL*	Embryonic axis specification	Ruffle	Actin monomer binding	
BIEC2_852756	4	2E+07	2E+07	9.0798	*COBL*	Embryonic axis specification	Ruffle	Actin monomer binding	

BIEC2_896745	4	2E+07	2E+07	6.4021	*COBL*	Embryonic axis specification	Ruffle	Actin monomer binding	
BIEC2_857781	4	3E+07	3E+07	3.5236	*SEMA3A*	Neural crest cell migration	Extracellular space	Semphorin receptor	Axon guidance
BIEC2_858347	4	3E+07	3E+07	3.3496	*ABCB1*	Phospholipid translocation	Cytoplasm	ATP binding	ABC transporters
BIEC2_867942	4	6E+07	6E+07	3.1053	*ZNRF2*	Proteasome-mediated ubiquitin-dependent protein catabolic process	Cytosome	Metal ion binding	
BIEC2_870040	4	7E+07	7E+07	4.5961	*ASZ1*	Male meiotic nuclear division	Chromatid body	Protein binding	
BIEC2_870061	4	7E+07	7E+07	4.7222	*ASZ1*	Male meiotic nuclear division	Chromatid body	Protein binding	
BIEC2_915234	4	7E+07	7E+07	3.7388	*CFTR*	Cholesterol biosynthetic process	Nucleus	Chloride channel activity	ABC transporters
AX-103007922	4	7E+07	7E+07	3.616	*CFTR*	Cholesterol biosynthetic process	Nucleus	Chloride channel activity	ABC transporters
BIEC2_915242	4	7E+07	7E+07	6.3647	*CFTR*	Cholesterol biosynthetic process	Nucleus	Chloride channel activity	ABC transporters
BIEC2-870400	4	8E+07	8E+07	5.1829	*TSPAN12*	Angiogenesis	Plasma membrane		
BIEC2_870414	4	8E+07	8E+07	3.5268	*TSPAN12*	Angiogenesis	Plasma membrane		
BIEC2-870421	4	8E+07	8E+07	3.6131	*TSPAN12*	Angiogenesis	Plasma membrane		
BIEC2_891163	5	9E+06	9E+06	3.0117	*SUCO*	Protein folding in endoplasmic reticulum	Cytoplasm		
BIEC2_891341	5	9E+06	9E+06	3.135	*TNFSF4*	T cell proliferation involved in immune response	Extracellular space	Cytokine activity	Cytokine-cytokine receptor interaction
BIEC2_898058	5	2E+07	2E+07	3.245	*PROX1*	Nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_898063	5	2E+07	2E+07	3.6047	*PROX1*	Nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_945393	5	3E+07	3E+07	3.1198	*PLXNA2*	Somitogensis	Semaphorin receptor complex	Protein binding	Axon guidance
BIEC2_908930	5	5E+07	5E+07	3.0796	*ST7L*	Regulates cell proliferation	Membrane	Protein binding	
BIEC2_908951	5	5E+07	5E+07	4.7108	*ST7L*	Regulates cell proliferation	Membrane	Protein binding	
BIEC2_909000	5	5E+07	5E+07	4.438	*KCND3*	Regulation of heart rate by cardiac conduction	Voltage-gated potassium channel complex	A-type (transient outward) potassium channel activity	Spinocerebellar ataxia
BIEC2_909009	5	5E+07	5E+07	4.5965	*DDX20*	Negative regulation of transcription by RNA polymerase II	Cytosol	RNA binding	
BIEC2_910441	5	6E+07	6E+07	3.7663	*VAV3*	Cell surface receptor signalling pathway	Immunological synapse	Protein serine/threonine phosphatase activity	Rap1 signalling pathway
BIEC2_949848	6	4E+07	4E+07	4.0042	*STYK1*	Protein phosphorylation	Plasma membrane	Protein tyrosine kinase activity	
BIEC2-949849	6	4E+07	4E+07	3.4222	*STYK1*	Protein phosphorylation	Plasma membrane	Protein tyrosine kinase activity	
BIEC2_955602	6	5E+07	5E+07	3.9148	*ITPR2*	Response to hypoxia	Nucleoplasm	Inositol 1,4,5-trisphosphate-gated calcium channel activity	Vascular smooth muscle contraction
BIEC2_955650	6	5E+07	5E+07	3.2581	*ITPR2*	Response to hypoxia	Nucleoplasm	Inositol 1,4,5-trisphosphate-gated calcium channel activity	Vascular smooth muscle contraction
BIEC2_1023722	6	8E+07	8E+07	4.7503	*PPM1H*	Protein dephosphorylation	Nucleus	Protein serine/threonine phosphatase activity	

BIEC2_975002	7	3E+06	3E+06	4.2019	*PIP5K1C*	Phosphorylation	Nucleoplasm	ATP binding	Inositol phosphate metabolism
BIEC2_975013	7	3E+06	3E+06	3.4082	*PIP5K1C*	Phosphorylation	Nucleoplasm	ATP binding	Inositol phosphate metabolism
BIEC2_1051000	7	4E+07	4E+07	5.812	*ARHGAP32*	Small GTPase-mediated signal transduction	Fibrillar centre	GTPase activator activity	
BIEC2_1051006	7	4E+07	4E+07	8.2599	*ARHGAP32*	Small GTPase-mediated signal transduction	Fibrillar centre	GTPase activator activity	
BIEC2_996483	7	4E+07	4E+07	8.7318	*ARHGAP32*	Small GTPase-mediated signal transduction	Fibrillar centre	GTPase activator activity	
BIEC2_996497	7	4E+07	4E+07	4.1682	*BARX2*	Negative regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II transcription regulatory region sequence-specific DNA binding	
BIEC2_996499	7	4E+07	4E+07	6.1303	*BARX2*	Negative regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II transcription regulatory region sequence-specific DNA binding	
BIEC2_996504	7	4E+07	4E+07	5.5316	*BARX2*	Negative regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II transcription regulatory region sequence-specific DNA binding	
BIEC2_1051037	7	4E+07	4E+07	5.7064	*TMEM45B*	Regulates cell proliferation	Membrane		
AX-103101847	7	4E+07	4E+07	5.7064	*TMEM45B*	Regulates cell proliferation	Membrane		
BIEC2-996512	7	4E+07	4E+07	8.6602	*TMEM45B*	Regulates cell proliferation	Membrane		
BIEC2_996515	7	4E+07	4E+07	9.8182	*TMEM45B*	Regulates cell proliferation	Membrane		
BIEC2_996517	7	4E+07	4E+07	10.508	*TMEM453*	Regulates cell proliferation	Membrane		
BIEC2_996517	7	4E+07	4E+07	10.508	*NFRKB*	DNA damage response	In080 complex	Protease binding	ATP-dependent chromatin remodeling
BIEC2_996519	7	4E+07	4E+07	10.992	*NFRKB*	DNA damage response	In080 complex	Protease binding	ATP-dependent chromatin remodeling
BIEC2-996537	7	4E+07	4E+07	7.7666	*PRDM10*	Regulation of gene expression	Nucleus	DNA binding	
BIEC2_996539	7	4E+07	4E+07	7.7666	*PRDM10*	Regulation of gene expression	Nucleus	DNA binding	
BIEC2-996543	7	4E+07	4E+07	9.4591	*APLP2*	Central nervous system development	Endomembrane system	Serine-type endopeptidase inhibitor activity	
BIEC2_996548	7	4E+07	4E+07	9.4591	*APLP2*	Central nervous system development	Endomembrane system	Serine-type endopeptidase inhibitor activity	
BIEC2_996554	7	4E+07	4E+07	10.096	*ST14*	Neural tube closure	Extracellular space	Serine-type endopeptidase activity	MicroRNAs in cancer
BIEC2_1051084	7	4E+07	4E+07	9.7007	*ZBTB44*	DNA binding activity	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104935820	7	4E+07	4E+07	5.32	*NTM*	Neurite outgrowth	Plasma membrane		
BIEC2_996658	7	4E+07	4E+07	6.4908	*NTM*	Neurite outgrowth	Plasma membrane		
BIEC2_996704	7	4E+07	4E+07	3.4225	*OPCML*	Tumour suppressor	Plasma membrane		
BIEC2_998629	7	5E+07	5E+07	3.0535	*ZNF653*	Negative regulation of transcription by RNA polymerase II	Extracellular region	Transcription coregulator activity	

BIEC2_1000176	7	5E+07	5E+07	12.15	*OR7D20*	Signal transduction	Plasma membrane	Olfactory receptor activity	Olfactory transduction
BIEC2_1000179	7	5E+07	5E+07	22.463	*OR7D20*	Signal transduction	Plasma membrane	Olfactory receptor activity	Olfactory transduction
BIEC2_1039371	8	3E+07	3E+07	3.0305	*BRI3BP*		Mitochondrion		
BIEC2_1091132	9	4E+07	4E+07	4.9284	*GDF6*	Metanephrose development	Extracellular space	Growth factor activity	Cytokine-cytokine receptor interaction
BIEC2_1091193	9	4E+07	4E+07	3.1959	*MTERF3*	Regulation of DNA-templated transcription	Mitochondrion	Transcription cis-regulatory region binding	
BIEC2_1091257	9	4E+07	4E+07	3.4034	*SDC2*	Dendrite morphogenesis	Cell surface		Cytoskeleton in muscle cells
BIEC2_1091271	9	4E+07	4E+07	3.0314	*SDC2*	Dendrite morphogenesis	Cell surface		Cytoskeleton in muscle cells
BIEC2_1091275	9	4E+07	4E+07	3.0314	*SDC2*	Dendrite morphogenesis	Cell surface		Cytoskeleton in muscle cells
BIEC2_1091292	9	4E+07	4E+07	3.0314	*SDC2*	Dendrite morphogenesis	Cell surface		Cytoskeleton in muscle cells
TBIEC2_1151328	9	4E+07	4E+07	3.6385	*SDC2*	Dendrite morphogenesis	Cell surface		Cytoskeleton in muscle cells
BIEC2_1100111	9	6E+07	6E+07	3.4782	*EXT1*	Protein glycosylation	Endoplasmic reticulum membrane	Acetylglucosaminyltransferase activity	Glycosaminoglycan biosynthesis - heparan sulfate / heparin
BIEC2_1103113	9	7E+07	7E+07	3.7328	*TBC1D31*	Cell projection organization	Centrosome	Protein binding	
BIEC2_1103152	9	7E+07	7E+07	3.2256	*TBC1D31*	Cell projection organization	Centrosome	Protein binding	
BIEC2_139647	10	8E+07	8E+07	4.0149	*EYA4*	Cell differentiation	Cytosol	RNA binding	
BIEC2_139653	10	8E+07	8E+07	3.4365	*EYA4*	Cell differentiation	Cytosol	RNA binding	
BIEC2_139663	10	8E+07	8E+07	3.1944	*EYA4*	Cell differentiation	Cytosol	RNA binding	
BIEC2_134482	11	3E+06	3E+06	3.0123	*TBC1D16*	Regulation of receptor recycling	Early endosome	GTPase activator activity	
BIEC2_210846	13	1E+07	1E+07	5.5778	*GTF2IRD1*	Transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AFFX-103122712	13	1E+07	1E+07	4.8684	*GTF2IRD1*	Transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_243044	13	3E+07	3E+07	4.4189	*CLEC16A*	Cellular response to starvation	Late endosome	Autophagy	
BIEC2_232951	13	3E+07	3E+07	4.5835	*CLEC16A*	Cellular response to starvation	Late endosome	Autophagy	
BIEC2_232956	13	3E+07	3E+07	3.0863	*CLEC16A*	Cellular response to starvation	Late endosome	Autophagy	
BIEC2_235535	13	4E+07	4E+07	3.0597	*RBFOX1*	Nervous system development	Nucleus	mRNA binding	
BIEC2_235536	13	4E+07	4E+07	3.0218	*RBFOX1*	Nervous system development	Nucleus	mRNA binding	
BIEC2_246144	13	4E+07	4E+07	5.4927	*SRL*	Response to muscle activity involved in regulation of muscle adaption	Membrane	GTP binding	
BIEC2_251110	14	3E+07	3E+07	3.0014	*PDGFRB*	Angiogenesis	Golgi apparatus	Platelet-derived growth factor beta-receptor activity	EGFR tyrosine kinase inhibitor resistance

BIEC2_251252	14	3E+07	3E+07	3.4798	*PPARGC1B*	Positive regulation of transcription by RNA polymerase II	Nucleus	RNA binding	Insulin resistance
BIEC2-251311	14	3E+07	3E+07	6.3368	*ARHGEF37*	Guanyl-nucleotide exchange factor activity	Cytoplasm	Guanyl-nucleotide exchange factor activity	
BIEC2_251347	14	3E+07	3E+07	3.8637	*ARHGEF37*	Guanyl-nucleotide exchange factor activity	Cytoplasm	Guanyl-nucleotide exchange factor activity	
BIEC2_251484	14	3E+07	3E+07	3.7426	*AFAP1L1*	Metastasis	Podosome	SH3 domain binding	
BIEC2_251566	14	3E+07	3E+07	3.0744	*ABLIM3*	Transcription by RNA polymerase II	Stress fibre	Metal ion binding	Axon guidance
BIEC2_251598	14	3E+07	3E+07	3.9218	*ABLIM3*	Transcription by RNA polymerase II	Stress fibre	Metal ion binding	Axon guidance
BIEC2_252938	14	3E+07	3E+07	3.4615	*PPP2R2B*	Peptidyl-serine dephosphorylation	Protein phosphatase type 2A complex	Protein binding	mRNA surveillance pathway
BIEC2_253025	14	3E+07	3E+07	4.6521	*PPP2R2B*	Peptidyl-serine dephosphorylation	Protein phosphatase type 2A complex	Protein binding	mRNA surveillance pathway
BIEC2_253054	14	3E+07	3E+07	3.5599	*PPP2R2B*	Peptidyl-serine dephosphorylation	Protein phosphatase type 2A complex	Protein binding	mRNA surveillance pathway
BIEC2-255042	14	3E+07	3E+07	3.0159	*PCDHB14*	Cell adhesion	Plasma membrane	Calcium ion binding	
BIEC2_255609	14	4E+07	4E+07	3.1214	*CXXC5*	Inflammatory response	Nucleus	Zinc ion binding	
BIEC2-273088	14	9E+07	9E+07	3.1935	*LHFPL2*	Development of primary female sexual characteristics	Membrane		
BIEC2_273163	14	9E+07	9E+07	3.3006	*LHFPL2*	Development of primary female sexual characteristics	Membrane		
BIEC2_299008	15	1E+07	1E+07	3.7244	*AFF3*	Regulation of gene expression	Nucleus		
BIEC2_300833	15	1E+07	1E+07	3.1234	*KCNIP3*	Negative regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_353654	16	6E+07	6E+07	3.4027	*RARB*	Negative regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Pathways in cancer
TBIEC2_370905	16	6E+07	6E+07	3.6685	*RARB*	Negative regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Pathways in cancer
BIEC2_355390	16	6E+07	6E+07	3.0893	*ZNF385D*		Nucleus	Nucleic acid binding	
BIEC2_373857	17	2E+07	2E+07	3.0501	*RCBTB2*		Cytoplasm	Protein binding	
BIEC2_373864	17	2E+07	2E+07	3.8351	*RB1*	G1/S transition of mitotic cell cycle	Chromatin	RNA polymerase II transcription regulatory region sequence-specific DNA binding	Endocrine resistance
BIEC2_411283	18	4E+07	4E+07	3.5906	*PKP4*	Cell-cell junction assembly	Spindle pole	Cadherin binding	
BIEC2_432555	18	4E+07	4E+07	3.7069	*PKP4*	Cell-cell junction assembly	Spindle pole	Cadherin binding	
BIEC2_412293	18	5E+07	5E+07	3.2626	*CERS6*	Ceramide biosynthetic process	Nucleus	DNA binding	Sphingolipid metabolism
BIEC2_415890	18	6E+07	6E+07	3.4455	*ZNF385B*		Nucleus	Nucleic acid binding	
BIEC2_430698	19	2E+07	2E+07	3.423	*NAALADL2*		Membrane		

BIEC2_430702	19	2E+07	2E+07	4.3431	*NAALADL2*		Membrane		
BIEC2_430706	19	2E+07	2E+07	4.1663	*NAALADL2*		Membrane		
BIEC2_609227	22	9E+06	9E+06	10.145	*MACROD2*	DNA damage response	Nucleoplasm	ADP-ribosylglutamate hydrolase activity	
BIEC2_582622	22	1E+07	1E+07	3.7988	*PAK5*	Protein phosphorylation	Cytoplasm	Protein serine/threonine kinase activity	ErB signalling pathway
BIEC2_607470	23	5E+06	5E+06	5.1701	*NTRK2*	Neuron migration	Early endosome	Protease binding	MAPK signalling pathway
BIEC2_607488	23	5E+06	5E+06	3.9386	*NTRK2*	Neuron migration	Early endosome	Protease binding	MAPK signalling pathway
BIEC2_607492	23	5E+06	5E+06	3.7401	*NTRK2*	Neuron migration	Early endosome	Protease binding	MAPK signalling pathway
TBIEC2_671574	24	2E+07	2E+07	7.4219	*TTLL5*	Microtubule cytoskeleton organization	Cilium	ATP binding	
BIEC2_639981	24	2E+07	2E+07	3.2194	*NRXN3*	Nervous system development	Membrane	Cell adhesion molecules	
BIEC2_639984	24	2E+07	2E+07	3.3933	*NRXN3*	Nervous system development	Membrane	Cell adhesion molecules	
BIEC2_664180	25	2E+07	2E+07	3.0254	*ASTN2*	Neuron migration	Endosome	Calcium ion binding	
BIEC2_668497	25	3E+07	3E+07	3.7071	*PBX3*	Eye development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Transcriptional misregulation in cancer
BIEC2_668509	25	3E+07	3E+07	3.1539	*PBX3*	Eye development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Transcriptional misregulation in cancer
BIEC2_700458	27	4E+06	4E+06	3.6314	*ANK1*	Endoplasmic reticulum to golgi vesicle-mediated transport	Cytoplasm	Protein binding	
BIEC2_737474	28	3E+07	3E+07	3.0721	*STAB2*	Receptor-mediated endocytosis	Cytosol	Low-density lipoprotein particle receptor activity	Efferocytosis

[Contents of cross-population selection signatures for BSP vs ECR: **SNP name**; Chromosome number (**Chr**); SNP start position (**Start SNP**); SNP end position (**End SNP**); **iHS Score**; Official gene name (**Gene Name**); Gene Ontology Term Biological Process (**GO Term BP**); Gene Ontology Term Cellular Component (**GO Term CC**); Gene Ontology Term Molecular Function (**GO Term MF**) and **KEGG Pathway** (Kyoto Encyclopedia of Genes and Genomes pathway) – provided through DAVID analysis platform (Sherman et al., 2022).]

Table 5.6 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across indigenous BSP and rural ELR breed populations

XP-EHH Selection Regions					Gene Ontology Enrichment				
SNP NAME	**Chr**	**Start SNP**	**End SNP**	**XP-EHH Score**	**Gene Names**	**GO Term BP**	**GO Term CC**	**GO Term MF**	**KEGG Pathway**
BIEC2_923800	5	82366776	82366776	3.37228614	*AK5*	Nucleoside triphosphate biosynthetic process	Cytoplasm	Adenylate kinase activity	Purine metabolism
BIEC2_1051200	7	41236063	41236063	3.38707871	*NTM*	Neurite outgrowth	Plasma membrane		Signal
BIEC2_996675	7	41315018	41315018	3.41583008	*NTM*	Neurite outgrowth	Plasma membrane		Signal
BIEC2_996678	7	41340513	41340513	3.90954682	*NTM*	Neurite outgrowth	Plasma membrane		Signal
BIEC2_996679	7	41363409	41363409	4.24442338	*NTM*	Neurite outgrowth	Plasma membrane		Signal
BIEC2_996681	7	41363562	41363562	4.24442338	*NTM*	Neurite outgrowth	Plasma membrane		Signal
BIEC2_996683	7	41367656	41367656	4.24442338	*NTM*	Neurite outgrowth	Plasma membrane		Signal
AX-103668539	7	41389215	41389215	4.4900171	*NTM*	Neurite outgrowth	Plasma membrane		Signal
BIEC2_996687	7	41489196	41489196	4.41476983	*NTM*	Neurite outgrowth	Plasma membrane		Signal
BIEC2-996688	7	41489510	41489510	4.41476983	*NTM*	Neurite outgrowth	Plasma membrane		Signal
BIEC2_996689	7	41492057	41492057	4.45411539	*NTM*	Neurite outgrowth	Plasma membrane		Signal
AFFX-101973792	7	41508901	41508901	4.6863565	*NTM*	Neurite outgrowth	Plasma membrane		Signal
AX-103156509	7	41531041	41531041	4.44304287	*NTM*	Neurite outgrowth	Plasma membrane		Signal
BIEC2_996690	7	41537197	41537197	4.59965352	*NTM*	Neurite outgrowth	Plasma membrane		Signal
BIEC2_1051229	7	41738138	41738138	3.67331382	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2-996703	7	41796700	41796700	3.71757534	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2_996704	7	41799089	41799089	3.6863113	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2-996706	7	41818480	41818480	4.259374	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2_996707	7	41818542	41818542	4.15627421	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2_996708	7	41826365	41826365	4.19001888	*OPCML*	Tumour suppressor	Plasma membrane		Signal
AX-103596269	7	41830516	41830516	3.98476217	*OPCML*	Tumour suppressor	Plasma membrane		Signal
AX-104326428	7	41899014	41899014	4.19251745	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2_996717	7	41916296	41916296	3.66274498	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2-996747	7	42206320	42206320	3.31265621	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2_996748	7	42206386	42206386	3.29252412	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2_996749	7	42206588	42206588	3.29252412	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2_996751	7	42298766	42298766	3.19173791	*OPCML*	Tumour suppressor	Plasma membrane		Signal

BIEC2_996752	7	42298816	42298816	3.19173791	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2_996753	7	42301832	42301832	3.19173791	*OPCML*	Tumour suppressor	Plasma membrane		Signal
AFFX-102814641	7	42384587	42384587	3.3488433	*OPCML*	Tumour suppressor	Plasma membrane		Signal
AX-103691822	7	42464946	42464946	3.27744224	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2_996766	7	42619193	42619193	3.4586671	*OPCML*	Tumour suppressor	Plasma membrane		Signal
BIEC2_996767	7	42695788	42695788	3.40236852	*OPCML*	Tumour suppressor	Plasma membrane		Signal
AX-104192181	7	43315852	43315852	3.06845979	*JAM3*	Angiogenesis	Membrane	Integrin binding	Cell adhesion molecules
AX-104958661	7	43579048	43579048	3.2775	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galactosidase activity	
AX-104730908	7	43589741	43589741	3.30144126	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galactosidase activity	
AX-103005562	7	43591211	43591211	3.27243329	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galactosidase activity	
AX-103135866	7	43665357	43665357	3.30144126	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galactosidase activity	
AX-102965540	7	43672523	43672523	3.37750715	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galactosidase activity	
AX-104383429	7	43709304	43709304	3.38369643	*B3GAT1*	Carbohydrate metabolic process	Golgi membrane	Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase activity	Mannose type O-glycan biosynthesis
AX-103097171	7	45676233	45676233	3.39719479	*ADGRE3*	Cell surface receptor signalling pathway	Plasma membrane	G protein-coupled receptor activity	
AX-103519045	7	45692689	45692689	3.39040069	*ADGRE3*	Cell surface receptor signalling pathway	Plasma membrane	G protein-coupled receptor activity	
AX-104284005	7	45695243	45695243	3.24210859	*ADGRE3*	Cell surface receptor signalling pathway	Plasma membrane	G protein-coupled receptor activity	
AX-104629108	7	45744927	45744927	3.27998194	*ADGRE3*	Cell surface receptor signalling pathway	Plasma membrane	G protein-coupled receptor activity	
AX-103450515	7	46096450	46096450	3.35986559	*IL27RA*	Cytokine-mediated signaling pathway	External side of plasma membrane	Interleukin-12 receptor binding	Cytokine-cytokine receptor interaction
AX-102971567	7	46172911	46172911	3.39903237	*CC2D1A*	Regulation of respiratory gaseous exchange by nervous system process	Fibrillar centre	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104890475	7	46174153	46174153	3.20246165	*CC2D1A*	Regulation of respiratory gaseous exchange by nervous system process	Fibrillar centre	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104113549	7	46237616	46237616	3.20246165	*ZSWIM4*	Intracellular glycine metabolism	Cul2-RING ubiquitin ligase complex	Zinc ion binding	
AX-103577324	7	46267067	46267067	3.30424121	*YJU2B*	RNA splicing	Nucleus		
AX-103861515	7	46269213	46269213	3.30424121	*YJU2B*	RNA splicing	Nucleus		

AX-103109620	7	46270828	46270828	3.28970904	*YJU2B*	RNA splicing	Nucleus		
AX-104261385	7	46357572	46357572	3.23072315	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	MAPK signalling pathway
AX-104349098	7	46413487	46413487	3.19995361	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	MAPK signalling pathway
AX-104036806	7	46425699	46425699	3.32678145	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	MAPK signalling pathway
AX-104930105	7	46538494	46538494	3.15814776	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	MAPK signalling pathway
AX-104462489	7	46557500	46557500	3.1847092	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	MAPK signalling pathway
AX-103052899	7	46569541	46569541	3.19574069	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	MAPK signalling pathway
BIEC2_997342	7	46673819	46673819	3.29967518	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104710278	7	46678814	46678814	3.29967518	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-103301628	7	46694639	46694639	3.29967518	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104297560	7	46707161	46707161	3.31618319	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_997345	7	46718011	46718011	3.38238862	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_997346	7	46718147	46718147	3.31188816	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-103238885	7	46734424	46734424	3.31188816	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104977211	7	46800568	46800568	3.31188816	*FARSA*	Phenylalanyl-tRNA aminoacylation	Cytoplasm	tRNA binding	Aminoacyl-tRNA biosynthesis
AX-103762761	7	46833343	46833343	3.31188816	*MAST1*	Brain development	Neuronal cell body	Magnesium ion binding	
AX-104737444	7	46840136	46840136	3.25613443	*MAST1*	Brain development	Neuronal cell body	Magnesium ion binding	
BIEC2_997364	7	46930776	46930776	3.39462095	*TNPO2*	Protein import into nucleus	Nucleus	Nuclear localization sequence binding	Nucleocytoplasmic transport
BIEC2_997365	7	46930777	46930777	3.39462095	*TNPO2*	Protein import into nucleus	Nucleus	Nuclear localization sequence binding	Nucleocytoplasmic transport
AX-103643647	7	46948035	46948035	3.39437843	*FBXW9*		Preribosome	Protein binding	
AX-103674066	7	46951310	46951310	3.34863888	*GNG14*	G protein-coupled receptor signalling pathway	Heterotrineric G-protein complex	G-protein beta-subunit	MAPK signalling pathway
AX-103761530	7	46966284	46966284	3.39210461	*MAN2B1*	Learning or memory	Nucleoplasm	Carbohydrate binding	Other glycan degradation
BIEC2_1011868	7	87211576	87211576	3.11215489	*SOX6*	Central nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_1011887	7	87230398	87230398	3.11552947	*SOX6*	Central nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_1011913	7	87279747	87279747	3.29317652	*SOX6*	Central nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_195273	12	25020452	25020452	3.00960538	*SAXO4*	Phosphate binding activity	Sperm flagella	Phosphatase binding	

[Contents of cross-population selection signatures for BSP vs ELR: **SNP name**; Chromosome number (**Chr**); SNP start position (**Start SNP**); SNP end position (**End SNP**); **iHS Score**; Official gene name (**Gene Name**); Gene Ontology Term Biological Process (**GO Term BP**); Gene Ontology Term Cellular Component (**GO Term CC**); Gene Ontology Term Molecular Function (**GO Term MF**) and **KEGG Pathway** (Kyoto Encyclopedia of Genes and Genomes pathway) – provided through DAVID analysis platform (Sherman et al., 2022).]

Table 5.7 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across indigenous BSP and rural KZN breed populations

XP-EHH Selection Regions					Gene Ontology Enrichment				
SNP NAME	**Chr**	**Start SNP**	**End SNP**	**XP-EHH Score**	**Gene Name**	**GO Term BP**	**GO Term CC**	**GO Term MF**	**KEGG Pathway**
BIEC2_923800	5	82366776	82366776	3.3722861	*AK5*	Nucleoside triphosphate biosynthetic process	Cytoplasm	Adenylate kinase activity	Purine metabolism
BIEC2_1051200	7	41236063	41236063	3.3870787	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
BIEC2_996675	7	41315018	41315018	3.4158301	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
BIEC2_996678	7	41340513	41340513	3.9095468	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
BIEC2_996679	7	41363409	41363409	4.2444234	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
BIEC2_996681	7	41363562	41363562	4.2444234	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
BIEC2_996683	7	41367656	41367656	4.2444234	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
AX-103668539	7	41389215	41389215	4.4900171	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
BIEC2_996687	7	41489196	41489196	4.4147698	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
BIEC2-996688	7	41489510	41489510	4.4147698	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
BIEC2_996689	7	41492057	41492057	4.4541154	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
AFFX-101973792	7	41508901	41508901	4.6863565	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
AX-103156509	7	41531041	41531041	4.4430429	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
BIEC2_996690	7	41537197	41537197	4.5996535	*NTM*	Neurite outgrowth	Plasma membrane	Signal	
BIEC2_1051229	7	41738138	41738138	3.6733138	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2-996703	7	41796700	41796700	3.7175753	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996704	7	41799089	41799089	3.6863113	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2-996706	7	41818480	41818480	4.259374	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996707	7	41818542	41818542	4.1562742	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996708	7	41826365	41826365	4.1900189	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
AX-103596269	7	41830516	41830516	3.9847622	*OPCML*	Tumour suppressor	Plasma membrane	Signal	

AX-104326428	7	41899014	41899014	4.1925174	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996717	7	41916296	41916296	3.662745	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2-996747	7	42206320	42206320	3.3126562	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996748	7	42206386	42206386	3.2925241	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996749	7	42206588	42206588	3.2925241	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996751	7	42298766	42298766	3.1917379	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996752	7	42298816	42298816	3.1917379	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996753	7	42301832	42301832	3.1917379	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
AFFX-102814641	7	42384587	42384587	3.3488433	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
AX-103691822	7	42464946	42464946	3.2774422	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996766	7	42619193	42619193	3.4586671	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
BIEC2_996767	7	42695788	42695788	3.4023685	*OPCML*	Tumour suppressor	Plasma membrane	Signal	
AX-104192181	7	43315852	43315852	3.0684598	*JAM3*	Angiogenesis	Membrane	Integrin binding	Cell adhesion molecules
AX-104958661	7	43579048	43579048	3.2775	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galasctoside activity	
AX-104730908	7	43589741	43589741	3.3014413	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galasctoside activity	
AX-103005562	7	43591211	43591211	3.2724333	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galasctoside activity	
AX-103135866	7	43665357	43665357	3.3014413	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galasctoside activity	
AX-102965540	7	43672523	43672523	3.3775071	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galasctoside activity	
AX-104383429	7	43709304	43709304	3.3836964	*B3GAT1*	Carbohydrate metabolic process	Golgi membrane	Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase activity	
AX-103097171	7	45676233	45676233	3.3971948	*ADGRE3*	Cell surface receptor signalling pathway	Plasma membrane	G protein-coupled receptor activity	
AX-103519045	7	45692689	45692689	3.3904007	*ADGRE3*	Cell surface receptor signalling pathway	Plasma membrane	G protein-coupled receptor activity	
AX-104284005	7	45695243	45695243	3.2421086	*ADGRE3*	Cell surface receptor signalling pathway	Plasma membrane	G protein-coupled receptor activity	
AX-104629108	7	45744927	45744927	3.2799819	*ADGRE3*	Cell surface receptor signalling pathway	Plasma membrane	G protein-coupled receptor activity	
AX-103450515	7	46096450	46096450	3.3598656	*IL27RA*	Cytokine-mediated signaling pathway	External side of plasma membrane	Interleukin-12 receptor binding	Cytokine-cytokine receptor interaction
AX-102971567	7	46172911	46172911	3.3990324	*CC2D1A*	Regulation of respiratory gaseous exchange by nervous system process	Fibrillar centre	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104890475	7	46174153	46174153	3.2024616	*CC2D1A*	Regulation of respiratory gaseous exchange by nervous system process	Fibrillar centre	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104113549	7	46237616	46237616	3.2024616	*ZSWIM4*	Intracellular glycine metabolism	Cul2-RING ubiquitin ligase complex	Zinc ion binding	
AX-103577324	7	46267067	46267067	3.3042412	*YJU2B*	mRNA splicing	Nucleus		

AX-103861515	7	46269213	46269213	3.3042412	*YJU2B*	mRNA splicing	Nucleus		
AX-103109620	7	46270828	46270828	3.289709	*YJU2B*	mRNA splicing	Nucleus		
AX-104261385	7	46357572	46357572	3.2307231	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
AX-104349098	7	46413487	46413487	3.1999536	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
AX-104036806	7	46425699	46425699	3.3267814	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
AX-104930105	7	46538494	46538494	3.1581478	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
AX-104462489	7	46557500	46557500	3.1847092	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
AX-103052899	7	46569541	46569541	3.1957407	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
BIEC2_997342	7	46673819	46673819	3.2996752	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104710278	7	46678814	46678814	3.2996752	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-103301628	7	46694639	46694639	3.2996752	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104297560	7	46707161	46707161	3.3161832	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_997345	7	46718011	46718011	3.3823886	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_997346	7	46718147	46718147	3.3118882	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-103238885	7	46734424	46734424	3.3118882	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104977211	7	46800568	46800568	3.3118882	*FARSA*	Phenylalanyl-tRNA aminoacylation	Cytoplasm	tRNA binding	Aminoacyl-tRNA biosynthesis
AX-103762761	7	46833343	46833343	3.3118882	*MAST1*	Brain development	Neuronal cell body	Magnesium ion binding	
AX-104737444	7	46840136	46840136	3.2561344	*MAST1*	Brain development	Neuronal cell body	Magnesium ion binding	
BIEC2_997364	7	46930776	46930776	3.3946209	*TNPO2*	Protein import into nucleus	Nucleus	Nuclear localization sequence binding	Nucleocytoplasmic transport
BIEC2_997365	7	46930777	46930777	3.3946209	*TNPO2*	Protein import into nucleus	Nucleus	Nuclear localization sequence binding	Nucleocytoplasmic transport
AX-103643647	7	46948035	46948035	3.3943784	*FBXW9*		Peribosome	Protein binding	
AX-103674066	7	46951310	46951310	3.3486389	*GNG14*	G protein-coupled receptor signalling pathway	Heterotrimeric G-protein complex	G-protein beta-subunit binding	MAPK signalling pathway
AX-103761530	7	46966284	46966284	3.3921046	*MAN2B1*	Learning or memory	Golgi membrane	Alpha-mannosidase activity	Other glycan degradation
BIEC2_1011868	7	87211576	87211576	3.1121549	*SOX6*	Central nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_1011887	7	87230398	87230398	3.1155295	*SOX6*	Central nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_1011913	7	87279747	87279747	3.2931765	*SOX6*	Central nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_195273	12	25020452	25020452	3.0096054	*SAXO4*			Phosphatase binding	

[Contents of cross-population selection signatures for BSP vs KZN: **SNP name**; Chromosome number (**Chr**); SNP start position (**Start SNP**); SNP end position (**End SNP**); **iHS Score**; Official gene name (**Gene Name**); Gene Ontology Term Biological Process (**GO Term BP**); Gene Ontology Term Cellular Component (**GO Term CC**); Gene Ontology Term Molecular Function (**GO Term MF**) and **KEGG Pathway** (Kyoto Encyclopedia of Genes and Genomes pathway) – provided through DAVID analysis platform (Sherman et al., 2022).]

Table 5.8 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across rural ECR and rural ELR breed populations

XP-EHH Selection Regions					**Gene Ontology Enrichment**				
SNP NAME	**Chr**	**Start SNP**	**End SNP**	**XP-EHH Score**	**Gene Name**	**GO Term BP**	**GO Term CC**	**GO Term MF**	**KEGG Pathway**
BIEC2_923800	5	82366776	82366776	3.5911225	*AK5*	Nucleoside triphosphate biosynthetic process	Cytoplasm	Adenylate kinase activity	Purine metabolism
BIEC2_923827	5	82400937	82400937	3.1295158	*AK5*	Nucleoside triphosphate biosynthetic process	Cytoplasm	Adenylate kinase activity	Purine metabolism
BIEC2_996678	7	41340513	41340513	3.0037129	*NTM*	Neurite outgrowth	Plasma membrane		
BIEC2_996679	7	41363409	41363409	3.4739724	*NTM*	Neurite outgrowth	Plasma membrane		
BIEC2_996681	7	41363562	41363562	3.4739724	*NTM*	Neurite outgrowth	Plasma membrane		
BIEC2_996683	7	41367656	41367656	3.4739724	*NTM*	Neurite outgrowth	Plasma membrane		
AX-103668539	7	41389215	41389215	3.6985334	*NTM*	Neurite outgrowth	Plasma membrane		
BIEC2_996687	7	41489196	41489196	3.6296541	*NTM*	Neurite outgrowth	Plasma membrane		
BIEC2-996688	7	41489510	41489510	3.6296541	*NTM*	Neurite outgrowth	Plasma membrane		
BIEC2_996689	7	41492057	41492057	3.6656618	*NTM*	Neurite outgrowth	Plasma membrane		
AFFX-101973792	7	41508901	41508901	3.8785602	*NTM*	Neurite outgrowth	Plasma membrane		
AX-103156509	7	41531041	41531041	3.6555268	*NTM*	Neurite outgrowth	Plasma membrane		
BIEC2_996690	7	41537197	41537197	3.7990079	*NTM*	Neurite outgrowth	Plasma membrane		
BIEC2-996706	7	41818480	41818480	3.616826	*OPCML*	Tumour supressor	Membrane	Signal	
BIEC2_996707	7	41818542	41818542	3.5209666	*OPCML*	Tumour supressor	Membrane	Signal	
BIEC2_996708	7	41826365	41826365	3.552329	*OPCML*	Tumour supressor	Membrane	Signal	
AX-103596269	7	41830516	41830516	3.3617573	*OPCML*	Tumour supressor	Membrane	Signal	
AX-104326428	7	41899014	41899014	3.5546516	*OPCML*	Tumour supressor	Membrane	Signal	
BIEC2_996717	7	41916296	41916296	3.0637705	*OPCML*	Tumour supressor	Membrane	Signal	
BIEC2_996766	7	42619193	42619193	3.1238368	*OPCML*	Tumour supressor	Membrane	Signal	
BIEC2_996767	7	42695788	42695788	3.0697725	*OPCML*	Tumour supressor	Membrane	Signal	

AX-102965540	7	43672523	43672523	3.0459067	*GLB1L2*	Carbohydrate metabolic process	Vacuole	Beta-galactosidase activity	
AX-104383429	7	43709304	43709304	3.0518476	*B3GAT1*	Carbohydrate metabolic process	Golgi membrane	Galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase activity	Mannose type O-glycan biosynthesis
AX-104629108	7	45744927	45744927	3.182353	*ADGRE3*	Cell surface receptor signalling pathway	Plasma membrane	G protein-coupled receptor activity	
AX-103450515	7	46096450	46096450	3.2619764	*IL27RA*	Positive regulation of activated T cell proliferation	Interleukin-12 receptor complex	Interleukin-12 receptor binding	Cytokine-cytokine receptor interaction
AX-102971567	7	46172911	46172911	3.3010249	*CC2D1A*	Regulation of respiratory gaseous exchange by nervous system process	Fibrillar center	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104890475	7	46174153	46174153	3.1051104	*CC2D1A*	Regulation of respiratory gaseous exchange by nervous system process	Fibrillar center	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104113549	7	46237616	46237616	3.1051104	*ZSWIM4*	Intracellular glycine metabolism	Cul2-RING ubiquitin ligase complex	Zinc ion binding	
AX-103577324	7	46267067	46267067	3.2065305	*YJU2B*	mRNA splicing	Nucleus	S-methyl-5-thioribose-1-phosphate isomerase activity	
AX-103861515	7	46269213	46269213	3.2065305	*YJU2B*	mRNA splicing	Nucleus	S-methyl-5-thioribose-1-phosphate isomerase activity	
AX-103109620	7	46270828	46270828	3.192047	*YJU2B*	mRNA splicing	Nucleus	S-methyl-5-thioribose-1-phosphate isomerase activity	
AX-104261385	7	46357572	46357572	3.1332678	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
AX-104349098	7	46413487	46413487	3.1026118	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
AX-104036806	7	46425699	46425699	3.2289969	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
AX-104930105	7	46538494	46538494	3.0609668	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
AX-104462489	7	46557500	46557500	3.0874252	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
AX-103052899	7	46569541	46569541	3.0984148	*CACNA1A*	Chemical synaptic transmission	Nucleus	Syntaxin binding	Synaptic vesicle cycle
BIEC2_997342	7	46673819	46673819	3.2019797	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104710278	7	46678814	46678814	3.2019797	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-103301628	7	46694639	46694639	3.2019797	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104297560	7	46707161	46707161	3.2184331	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	

BIEC2_997345	7	46718011	46718011	3.2844307	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_997346	7	46718147	46718147	3.2141522	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-103238885	7	46734424	46734424	3.2141522	*NFIX*	DNA replication	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AX-104977211	7	46800568	46800568	3.2141522	*FARSA*	Phenylalanyl-tRNA aminoacylation	Cytoplasm	tRNA binding	Aminoacyl-tRNA biosynthesis
AX-103762761	7	46833343	46833343	3.2141522	*MAST1*	Brain development	Neuronal cell body	Magnesium ion binding	
AX-104737444	7	46840136	46840136	3.1585883	*MAST1*	Brain development	Neuronal cell body	Magnesium ion binding	
BIEC2_997364	7	46930776	46930776	3.2966265	*TNPO2*	Protein import into nucleus	Nucleus	Nuclear localization sequence binding	Nucleocytoplasmic transport
BIEC2_997365	7	46930777	46930777	3.2966265	*TNPO2*	Protein import into nucleus	Nucleus	Nuclear localization sequence binding	Nucleocytoplasmic transport
AX-103643647	7	46948035	46948035	3.2963847	*FBXW9*	Protein ubiquitnation and kinase activities	Preribosome	Protein binding	
AX-103674066	7	46951310	46951310	3.2507847	*GNG14*	G protein-coupled receptor signalling pathway	Heterotrimeric G-protein binding	G-protein beta-subunit binding	MAPK signalling pathway
AX-103761530	7	46966284	46966284	3.2941176	*MAN2B1*	Learning or memory	Golgi membrane	Alpha-mannosidase activity	Other glycan degradation
BIEC2_1011868	7	87211576	87211576	3.1445776	*SOX6*	Central nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_1011887	7	87230398	87230398	3.2470648	*SOX6*	Central nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_1011913	7	87279747	87279747	3.4881678	*SOX6*	Central nervous system development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
UKUL1776	8	10181798	10181798	3.0904636	*CHEK2*	G2/M transition of mitotic cell cycle	Chromosome	Protein serine/threonine kinase activity	Cell cycle
BIEC2_1028365	8	10213477	10213477	3.5987705	*CHEK2*	G2/M transition of mitotic cell cycle	Chromosome	Protein serine/threonine kinase activity	Cell cycle
BIEC2_1028365	8	10213477	10213477	3.5987705	*TTC28*	Regulation of mitotic cell cycle	Midbody	Protein binding	
BIEC2_1028387	8	10225109	10225109	3.6470174	*TTC28*	Regulation of mitotic cell cycle	Midbody	Protein binding	
BIEC2_1028408	8	10248408	10248408	3.2536091	*TTC28*	Regulation of mitotic cell cycle	Midbody	Protein binding	
BIEC2_98548	10	9270482	9270482	3.0666608	*SIPA1L3*	Epithelial cell morphogenesis	Stress fiber	GTPase activator activity	Rap1 signalling pathway
BIEC2_98548	10	9270482	9270482	3.0666608	*WDR87*	Sperm tail assembly	Membrane		
BIEC2_206614	13	5778195	5778195	3.1029815	*BAIAP2L1*	Plasma membrane organization	Nucleoplasm	Protein binding	

[Contents of cross-population selection signatures for ECR vs ELR: **SNP name**; Chromosome number (**Chr**); SNP start position (**Start SNP**); SNP end position (**End SNP**); **iHS Score**; Official gene name (**Gene Name**); Gene Ontology Term Biological Process (**GO Term BP**); Gene Ontology Term Cellular Component (**GO Term CC**); Gene Ontology Term Molecular Function (**GO Term MF**) and **KEGG Pathway** (Kyoto Encyclopedia of Genes and Genomes pathway) – provided through DAVID analysis platform (Sherman et al., 2022).]

Table 5.9 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across rural ECR and rural KZN breed populations

XP-EHH Selection Regions					Gene Ontology Enrichment				
SNP NAME	**Chr**	**Start SNP**	**End SNP**	**XP-EHH Score**	**Gene Name**	**GO Term BP**	**GO Term CC**	**GO Term MF**	**KEGG Pathway**
AX-104838147	1	87564265	87564265	3.0674428	*NRG3*	Nervous system development	Extracellular space	Growth factor activity	ErbB signalling pathway
BIEC2_39403	1	87568867	87568867	3.1537703	*NRG3*	Nervous system development	Extracellular space	Growth factor activity	ErbB signalling pathway
BIEC2_77434	1	163274628	163274628	5.1421537	*MYH6*	In utero embryonic development	Stress fiber	Microfilament motor activity	cGMP-PKG signalling pathway
BIEC2-77469	1	163310160	163310160	5.9676688	*MYH7*	In utero embryonic development	Stress fiber	Microfilament motor activity	cGMP-PKG signalling pathway
BIEC2_77519	1	163376442	163376442	4.7186794	*THTPA*	Thiamine metabolic process	Cytoplasm	Magnesium ion binding	Thiamine metabolism
BIEC2_81656	1	163396175	163396175	5.2799016	*THTPA*	Thiamine metabolic process	Cytoplasm	Magnesium ion binding	Thiamine metabolism
BIEC2_81656	1	163396175	163396175	5.2799016	*ZFHX2*	Brain development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_510173	2	119238687	119238687	3.2020524	*TBCK*	Protein phosphorylation	Kinetochore	Protein kinase activity	
AX-104497232	3	1904114	1904114	3.473856	*N4BP1*	Negative regulation of cytokine production	Nucleus	mRNA binding	
BIEC2_772256	3	10847497	10847497	4.1521278	*MMP15*	Proteolysis	Extracellular space	Metalloendopeptidase activity	Parathyroid hormone synthesis
TBIEC2_835258	3	79024465	79024465	3.0340438	*CLOCK*	DNA damage checkpoint signalling	Nucleus	DNA-binding transcription factor activity	Circadian rhythms
TBIEC2_835258	3	79024465	79024465	3.0340438	*TMEM165*	Protein N-linked glycosylation	Lysosomal membrane	Manganese ion transmembrane transporter activity	
BIEC2_794825	3	79193362	79193362	3.0619322	*KDR*	Ovarian follicle development	Nucleus	Vascular endothelial growth factor receptor activity	EGFR tyrosine kinase inhibitor resistance
BIEC2_835334	3	79219475	79219475	3.4926593	*KDR*	Ovarian follicle development	Nucleus	Vascular endothelial growth factor receptor activity	EGFR tyrosine kinase inhibitor resistance
BIEC2_794876	3	79278101	79278101	3.3651441	*KDR*	Ovarian follicle development	Nucleus	Vascular endothelial growth factor receptor activity	EGFR tyrosine kinase inhibitor resistance

AFFX-102918356	3	88654153	88654153	3.6292362	*APBB2*	Neuron migration	Nucleus	Amyloid-beta binding	
AX-103015312	3	88690583	88690583	3.1583232	*APBB2*	Neuron migration	Nucleus	Amyloid-beta binding	
AX-104663653	3	88771818	88771818	3.7220295	*APBB2*	Neuron migration	Nucleus	Amyloid-beta binding	
AFFX-102070146_ILMNDUP1	3	88878710	88878710	3.3709211	*APBB2*	Neuron migration	Nucleus	Amyloid-beta binding	
AX-103973889	3	89188941	89188941	3.8473297	*RBM47*	Regulation of alternative mRNA splicing	Nucleus	mRNA binding	
AFFX-102544558	3	90069924	90069924	3.9281731	*RFC1*	DNA replication	Nucleus	DNA binding	DNA replication
BIEC2-799664_ILMNDUP1	3	90095202	90095202	3.4828212	*RFC1*	DNA replication	Nucleus	DNA binding	DNA replication
AX-104633567	3	96858251	96858251	3.5961045	*PCDH7*	Cell adhesion	Plasma membrane	Calcium ion binding	DNA replication
BIEC2_804978	3	96865083	96865083	3.0380543	*PCDH7*	Cell adhesion	Plasma membrane	Calcium ion binding	DNA replication
BIEC2-804981	3	96869620	96869620	3.0737013	*PCDH7*	Cell adhesion	Plasma membrane	Calcium ion binding	DNA replication
AFFX-102385430	3	96887359	96887359	3.2867153	*PCDH7*	Cell adhesion	Plasma membrane	Calcium ion binding	DNA replication
AFFX-102783315	3	104364621	104364621	3.0268419	*KCNIP4*	Potassium channels	Cytosol	Potassium ion transport	
BIEC2_845339	4	7256105	7256105	3.0515985	*COG5*	Intra-golgi vesicle-mediated transport	Membrane		
TBIEC2_888991	4	7342482	7342482	3.2045659	*COG5*	Intra-golgi vesicle-mediated transport	Membrane		
BIEC2_888997	4	7394218	7394218	3.2797319	*COG5*	Intra-golgi vesicle-mediated transport	Membrane		
AX-103565705	5	50510979	50510979	3.9954845	*SYCP1*	Meiotic DNA repair synthesis	Synaptonemal complex	Double-stranded DNA binding	
BIEC2_1004430	7	61400387	61400387	3.4806462	*ME3*	Pyruvate metabolic process	Mitochondrion	NAD+ activity	Pyruvate metabolism
BIEC2_1033163	8	16214049	16214049	3.4733428	*SIRT4*	Regulation of glutamine family amino acid metabolic process	Nucleus	NAD+ ADP-ribosyltransferase activity	Nicotine and nicotinamide metabolism
BIEC2_1034815	8	18611435	18611435	3.9270234	*NOS1*	Response to hypoxia	Nucleus	Nitric-oxide synthase activity	Arginine biosynthesis
BIEC2_1043683	8	35025281	35025281	3.4318067	*ARHGAP28*	Signal transduction	Cytoplasm	GTPase activator activity	
AX-104411223	8	43627752	43627752	3.1472798	*ROCK1*	Mitotic cytokinesis	Ruffle	ATP binding	cAMP signalling pathway
BIEC2-1061702	8	69524910	69524910	3.6469026	*SKOR2*	Negative regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_1076655	9	16488952	16488952	4.5418027	*SULF1*	Cartilage condensation	Extracellular space	Arylsulfatase activity	
BIEC2-1076656	9	16489180	16489180	3.8819661	*SULF1*	Cartilage condensation	Extracellular space	Arylsulfatase activity	
TBIEC2_1135999	9	16500582	16500582	3.371025	*SULF1*	Cartilage condensation	Extracellular space	Arylsulfatase activity	

AX-104408612	10	30666628	30666628	3.5011773	*FILIP1*	Brain development	Actin cytoskeleton		
AX-104802102	13	21928560	21928560	3.4560596	*NSMCE1*	Post replication repair	Chromosome	Ubiquitin-protein transferase activity	
AFFX-102199562	13	21953792	21953792	3.1360946	*NSMCE1*	Post replication repair	Chromosome	Ubiquitin-protein transferase activity	
BIEC2_285537	15	8761961	8761961	4.0862726	*NPAS2*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	Circadian rhythms
BIEC2_285546	15	8778532	8778532	4.0654288	*NPAS2*	Regulation of transcription by RNA polymerase II	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_299008	15	9563433	9563433	3.1379311	*AFF3*	Regulation of gene expression	Nucleus		
BIEC2_311792	15	54419913	54419913	3.2320793	*SRBD1*	Nucleobase-containing compound metabolic process	Cytosol	mRNA binding	
BIEC2_326302_ILMNDUP1	15	54773734	54773734	3.1161359	*SIX2*	Kidney development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_311864	15	54785493	54785493	4.7823587	*SIX2*	Kidney development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_311864	15	54785493	54785493	4.7823587	*SIX3*	Eye development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_311866	15	54788791	54788791	4.4659891	*SIX3*	Eye development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_311876	15	54834053	54834053	3.7769323	*SIX3*	Eye development	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
AFFX-101936454_ILMNDUP1	15	85780577	85780577	3.2836455	*KIDINS220*	Nerve growth factor signalling pathway	Membrane	Protein binding	Neurotrophic signalling pathway
AFFX-102273647	16	42730087	42730087	3.0365515	*LARS2*	Leucyl-tRNA	Mitochondrion	Aminoacyl-tRNA editing activity	Aminoacyl-tRNA biosynthesis
AFFX-101680795	17	61876703	61876703	3.0481704	*GPC5*	Cell migration	Extracellular region		
AFFX-102451154	19	56151732	56151732	3.3504181	*ZPLD1*	Vestibular reflex	Extracellular space		
AX-104865181	19	56198826	56198826	4.9026211	*ZPLD1*	Vestibular reflex	Extracellular space		
TBIEC2_582461	21	17151876	17151876	3.6906982	*IL31RA*	Cytokine-mediated signalling pathway	External side of plasma membrane	Cytokine receptor activity	Cytokine-cytokine receptor interaction
TBIEC2_582487	21	17174362	17174362	3.1053912	*IL31RA*	Cytokine-mediated signalling pathway	External side of plasma membrane	Cytokine receptor activity	Cytokine-cytokine receptor interaction
BIEC2_554325	21	17383032	17383032	3.3088494	*SLC38A9*	Amino acid transmembrane transport	Lysosomal membrane	Guanyl-nucleotide exchange factor activity	mTOR signalling pathway
BIEC2_621341	23	25606959	25606959	3.2438895	*CDC37L1*	Protein folding	Cytoplasm	Heat shock protein binding	
BIEC2_639744	24	23468703	23468703	3.7792454	*NRXN3*	Synapse development	Membrane		Cell adhesion molecules
BIEC2_639754	24	23503899	23503899	3.0493608	*NRXN3*	Synapse development	Membrane		

AX-103750649	25	17568016	17568016	3.6324838	*HSDL2*	Siderophore biosynthetic process	Mitochondrion	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase activity	
AX-102984008	26	12268798	12268798	3.299748	*ROBO2*	Cell-cell adhesion	Membrane	Protein binding	Axon guidance
AFFX-102357556	26	38747393	38747393	3.0061924	*C2CD2*	Insulin secretion	Membrane		
BIEC2-744912	28	43233840	43233840	6.026071	*PPARA*	Fatty acid metabolic process	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	PPAR signalling pathway
BIEC2_744947	28	43329623	43329623	6.3404053	*GTSE1*	Cell cycle regulation	Cytoplasmic microtubule	Microtubule binding	p53 signalling pathway
BIEC2_744951	28	43345662	43345662	5.0817532	*TRMU*	tRNA wobble position uridine thiolation	Mitochondrion	tRNA binding	tRNA processing
BIEC2_782985	28	43405672	43405672	5.1155811	*CELSR1*	Hair follicle development	Nucleoplasm	G protein-coupled receptor activity	
BIEC2_744990	28	43452543	43452543	3.7069471	*CELSR1*	Hair follicle development	Nucleoplasm	G protein-coupled receptor activity	
BIEC2_744994_ILMNDUP1	28	43460324	43460324	3.2400765	*CELSR1*	Hair follicle development	Nucleoplasm	G protein-coupled receptor activity	
BIEC2_783014	28	43579157	43579157	3.4179109	*GRAMD4*	Apoptotic process	Mitochondrion		
BIEC2_745020	28	43606571	43606571	3.6594934	*GRAMD4*	Apoptotic process	Mitochondrion		
BIEC2_745042	28	43670383	43670383	3.4851498	*CERK*	Ceramide metabolic process	Plasma membrane	Magnesium ion binding	Sphingolipid metabolism

[Contents of cross-population selection signatures for ECR vs KZN: **SNP name**; Chromosome number (**Chr**); SNP start position (**Start SNP**); SNP end position (**End SNP**); **iHS Score**; Official gene name (**Gene Name**); Gene Ontology Term Biological Process (**GO Term BP**); Gene Ontology Term Cellular Component (**GO Term CC**); Gene Ontology Term Molecular Function (**GO Term MF**) and **KEGG Pathway** (Kyoto Encyclopedia of Genes and Genomes pathway) – provided through DAVID analysis platform (Sherman et al., 2022).]

Table 5.10 Gene identification, functional annotation and enrichment of XP-EHH selection signatures across rural ELR and rural KZN breed populations

XP-EHH Selection Regions					**Gene Ontology Enrichment**				
SNP NAME	**Chr**	**Start SNP**	**End SNP**	**XP-EHH Score**	**Gene Name**	**GO Term BP**	**GO Term CC**	**GO Term MF**	**KEGG Pathway**
BIEC2_776758	3	35316001	35316001	3.0698427	*BANP*	Chromatin organization	Nucleus	DNA binding	
BIEC2_923462	5	81990779	81990779	3.0479996	*USP33*	Proteolysis	Nucleoplasm	G protein-coupled receptor binding	
BIEC2_923800	5	82366776	82366776	3.4883532	*AK5*	Nucleoside triphosphate biosynthetic process	Cytoplasm	Adenylate kinase activity	Purine metabolism
BIEC2_923827	5	82400937	82400937	3.2293488	*AK5*	Nucleoside triphosphate biosynthetic process	Cytoplasm	Adenylate kinase activity	Purine metabolism
AX-103600422	5	82430991	82430991	3.6931616	*AK5*	Nucleoside triphosphate biosynthetic process	Cytoplasm	Adenylate kinase activity	Purine metabolism

AX-103600422	5	82430991	82430991	3.6931616	*PIGK*	Proteolysis	GPI-anchor transamidase complex	GPI-anchor transamidase activity	Glycosylphosphatidylinositol (GPI)-anchor biosynthesis
BIEC2_923860	5	82432955	82432955	3.2133363	*PIGK*	Proteolysis	GPI-anchor transamidase complex	GPI-anchor transamidase activity	Glycosylphosphatidylinositol (GPI)-anchor biosynthesis
BIEC2_1011913	7	87279747	87279747	3.0089918	*SOX6*	Central nervous system	Nucleus	RNA polymerase II cis-regulatory region sequence-specific DNA binding	
BIEC2_621333	23	25493403	25493403	3.0423123	*SLC1A1*	Behavioural fear response	Plasma membrane	Monoatomic anion channel activity	Synaptic vesicle cycle
BIEC2_757868	29	21100468	21100468	3.142649	*ACBD7*	Fatty acid metabolic process	Cytoplasm	Fatty-acyl-CoA binding	
BIEC2_757880	29	21103773	21103773	3.6922407	*OLAH*	Lipid biosynthetic process		Fatty acyl-[ACP] hydrolase activity	Fatty acid biosynthesis
BIEC2_757887	29	21107279	21107279	3.5206597	*OLAH*	Lipid biosynthetic process		Fatty acyl-[ACP] hydrolase activity	Fatty acid biosynthesis
BIEC2_757912	29	21142027	21142027	3.5118488	*DCLRE1C*	Telomere maintenance	Nucleoplasm	Single-stranded DNA endodeoxyribonuclease activity	Non-homologous end-joining
BIEC2_757912	29	21142027	21142027	3.5118488	*MEIG1*	Sperm axoneme assembly	Manchette		

[Contents of cross-population selection signatures for ELR vs KZN: **SNP name**; Chromosome number (**Chr**); SNP start position (**Start SNP**); SNP end position (**End SNP**); **iHS Score**; Official gene name (**Gene Name**); Gene Ontology Term Biological Process (**GO Term BP**); Gene Ontology Term Cellular Component (**GO Term CC**); Gene Ontology Term Molecular Function (**GO Term MF**) and **KEGG Pathway** (Kyoto Encyclopedia of Genes and Genomes pathway) – provided through DAVID analysis platform (Sherman et al., 2022).]

BSP gene identification and biological pathways from the above XP-EHH analyses revealed beneficial selection signatures detected within selection sweeps, and unique selection signatures outside selection sweep regions, that were not found within the population-specific iHS analyses. The BSP vs ECR XP-EHH analyses (**Figure 5.3 A**) that detected selection sweeps on chromosomes 2, 4, 5, 7 and 19 found very similar gene functional annotations that were detected within the iHS gene identification and enrichment in **Tables 5.1 – 5.4**. Functional annotations of cellular processes (*EIF2B3*; *MEGF6* – three detection signals; *PRDM16* – eleven detection signals; *ELP3*) and fertility traits (*NPNT*) were detected on chromosome 2. Selection sweep detected on chromosome 4 contained selection signatures associated with fertility (*COBL* – three detection signals; *ASZ1* – two detection signals), feed efficiency (*ABCB1*; *CFTR* – three detection signals), blood vessel morphogenesis (*TSPAN12* – three detection signals) and cellular processes (*FAM185A*; *SEMA3A*; *ZNRF2*). Chromosome 5 selection sweep contained selection signatures associated with immune response (*TNFSF4*), nervous system development (*PROX1* – two detection signals), cellular processes (*SUCO*; *PLXNA2*; *ST7l* – two detection signatures; *DDX20*; *VAV3*) and cardiac cycle development (*KCND3*). On chromosome 7, selection signatures within the sweep were functionally associated with cellular processes (*PIP5K1C* – two detection signals; *ARHGAP32* – three detection signals; *BARX2* – three detection signals; *TMEM45B* – five detection signals; *NFRKB* – two detection signals; *PRDM10* – two detection signals; *ZBTB44*; *OPCML*; *ZNF653*) and selection sweep on chromosome 19 only contained selection signatures associated with cellular processes (*NAALADL2* – three detection signals). Unique selection signatures outside of selection sweeps were found on chromosome 13 in which functional annotation of cellular response to starvation (*CLEC16A* – three detection signals) and muscle activity response in the regulation of muscle adaption (SRL). Another unique selection signature was detected on chromosome 6 that was associated with response to hypoxia (*ITPR2* – two detection signals).

Selection sweeps were detected in four cross-population XP-EHH analyses, specifically between BSP vs ECR (**Table 5.5**), BSP vs KZN (**Table 5.7**), ECR vs ELR (**Table 5.8**) and ELR vs KZN (**Table 5.10**), on chromosome 7. Indigenous BSP and ELR cross-population analyses functionally annotated selection signatures associated with nervous system development (*NTM* – thirteen detection signals; *MAST1* – two detection signals; *SOX6* – three detection signals; *CC2DIA* – two detection signals; *CACNA1A* – six detection signals), cellular processes (*OPCML* – nineteen detection signals; *ADGRE3* – four detection signals; *IL27RA*; *ZSWIM4*; *YJU2B* – three detection signals; *NFIX* – seven detection signals; *FARSA*; *TNPO2* – two detection signals; *FBXW9*; *GNG14*). There were also selection signatures associated with behaviour (MAN2B1), blood vessel morphogenesis (*JAM3*) and feed efficiency (*GLB1L2* – five detection signals; *B3GAT1*). BSP vs KZN cross-population analyses detected the exact same accumulation of each selection signatures within the BSP vs ECR XP-EHH analyses and associated functional annotations. This trend was also observed within the ECR vs ELR cross-population analyses, whereby a decrease in the number of detection signals of highly selected genes,

including *NTM* (eleven detection signals), *OPCML* (eight detection signals), *ADGRE3* (one detection signal), *YJUB2B* (one detection signal) and *GLB1L2* (one detection signal). It was noted that selection signals associated with blood vessel morphogenesis (*JAM3*) was not detected within the ECR vs ELR cross-population analyses. The ELR vs KZN functional annotation of XP-EHH analysis (**Table 5.10**) yielded selection signatures associated with nervous system development (*SOX6*) within selection sweep as well as a unique selection signature associated with behavioural fear response (*SLC1A1*) detected on chromosome 23.

The last cross-population, XP-EHH analysis detected selection signatures (**Table 5.9**) found on selection sweeps (**Figure 5.4 B**) on chromosomes 1, 3, 15 and 28. Selection sweep on chromosome 1 contained selection signatures associated with nervous system development (*NRG3* – two detection signals; *ZFH* – two detection signals), fertility traits (*MYH6*; *MYH7*) and cellular processes (*THTPA* – two detection signals). Chromosome 3 selection sweep contained selection signatures associated with nervous system development (*APBB2* – four detection signals), fertility traits (*KDR* – three detection signals), cellular processes (*N4BP1*; *CLOCK*; *TMEM165*; *RBM47*; *FC1* – two detection signals; *PCDH7* – four detection signals; *KCNIP4*). Selection sweep on chromosome 15 contained selection signatures associated with nervous system development (*KIDINS220*), cellular processes (*NPA2* – two detection signals; *AFF3*; *SRBD1*), kidney development (*SIX2* – two detection signals) and eye development (*SIX3* – three detection signals). Lastly, the selection sweep on chromosome 28 contained selection signatures associated with cellular processes (*GTSE1*; *TRMU*; *GRAMD4* – two detection signals; *CERK*), feed efficiency (*PPARA*) and hair follicle development (*CELSR1* – three detection signals). Unique selection signatures associated with vestibular reflex (*ZPLD1* – two detection signals) was detected on chromosome 19, response to hypoxia (*NOS1*) on chromosome 8 and cartilage condensation (*SULF1* – three detection signals) on chromosome 9.

To critically assess shared adaptive targets across populations and extent of population specific selection signatures within each population, an intersection comparison of functionally annotated genes was determined within **Table 5.11** below, between iHS and XP-EHH analyses.

Table 5.11 Intersected selection signatures detected within population-specific iHS analyses and cross-population XP-EHH analyses

	Chr	Analysis	Breed	Gene
1	**1**	iHS	BSP; ECR	*CORO2B*
		iHS	ECR; KZN	*RAB27A*
		iHS	BSP; ECR; KZN	*UNC13C*
		iHS	BSP; ECR	*FBN1*
2	**2**	iHS	ECR; KZN	*KCNAB2*
		iHS	BSP; ECR; ELR	*LOXL2*
3	**4**	iHS; XP-EHH	KZN; BSP; ECR; BSPvsECR	*COBL*
		iHS; XP-EHH	KZN; ECR; BSPvsECR	*SEMA3A*
		iHS; XP-EHH	KZN; BSPvsECR	*TSPAN12*
		iHS	BSP; KZN	*LRRD1*
		iHS	BSP; KZN	*PPP1R9A*
		iHS	BSP; KZN	*DYNC1I1*
4	**5**	iHS; XP-EHH	BSP; BSPvsECR	*KCND3*
		XP-EHH	BSPvsELR; ECRvsELR; ELRvsKZN; BSPvsKZN	*AK5*
5	**6**	iHS; XP-EHH	ECR; BSPvsECR	*STYK1*
		iHS	BSP; ELR	*ERC1*
6	**7**	iHS; XP-EHH	KZN; BSPvsECR	*TMEM45B*
		iHS; XP-EHH	ELR; KZN; BSPvsECR; BSPvsELR; BSPvsKZN; ECRvsELR	*NTM*
		iHS; XP-EHH	ELR; KZN; BSPvsECR; BSPvsELR; BSPvsKZN; ECRvsELR	*OPCML*
		iHS; XP-EHH	BSP; BSPvsECR	*OR7D20*
		XP-EHH	BSPvsELR; BSPvsKZN	*JAM3*
		XP-EHH	BSPvsELR; BSPvsKZN; ECRvsELR	*GLB1L2*
		XP-EHH	ECRvsELR; BSPvsELR; BSPvsKZN	*B3GAT1*
		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*ADGRE3*
		XP-EHH	BSPvsELR; BSPvsKZN; ECRvsELR	*IL27RA*
		XP-EHH	BSPvsELR; BSPvsKZN	*CC2D1A*
		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*ZSWIM4*
		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*YJU2B*

		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*CACNA1A*
		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*NFIX*
		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*FARSA*
		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*MAST1*
		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*TNPO2*
		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*FBXW9*
		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*GNG14*
		XP-EHH	BSPvsELR; ECRvsELR; BSPvsKZN	*MAN2B1*
		XP-EHH	BSPvsELR; ECRvsELR; ELRvsKZN; BSPvsKZN	*SOX6*
		iHS	BSP; KZN	*VWA5A*
7	**8**	iHS; XP-EHH	ECR; ECRvsKZN	*ARHGAP28*
		iHS	BSP; ECR	*MYO18B*
8	**9**	iHS	ELR; KZN	*PXDNL*
		iHS	BSP; KZN	*ATP6V1H*
9	**11**	iHS	BSP; ELR	*FAM117A*
10	**12**	XP-EHH	BSPvsELR; BSPvsKZN	*SAXO4*
11	**15**	XP-EHH	BSPvsECR; ECRvsKZN	*AFF3*
12	**17**	iHS; XP-EHH	BSP; BSPvsECR	*RCBTB2*
13	**24**	XP-EHH	BSPvsECR; ECRvsKZN	*NRXN3*
14	**25**	iHS; XP-EHH	BSP; BSPvsECR	*PBX3*
		iHS	BSP; ECR	*MAPKAP1*
15	**27**	iHS	BSP; KZN	*ADRB3*

[The above table accounts for shared selection signatures of specific genes that intersect between populations detected from iHS and XP-EHH analyses. Contents within the table include: Chromosome number (**Chr**); Population analyses (**Analysis**); Breed Populations that were examined (**Breed**); Genes in common (**Gene**).]

Intersected selection signatures determined between the population-specific iHS analyses and cross-population XP-EHH (**Table 5.11**) detected 50 genes that were shared amongst the indigenous BSP breed population and three rural breed populations – across 15 chromosomes. The selection sweep on chromosome 7 stayed consistent within intersected regions of selection signatures as containing the most shared genes across all breed populations, detected within both population analyses. Intersected selection signatures on chromosomes 1, 2 and 9 were specific to the within population iHS, shared between the four breed populations. Intersected selection signatures on chromosomes 12, 15 and 24 were specific to genes detected within the cross-population XP-EHH analyses of BSPvsELR, BSPvsKZN, BSPvsECR and ECRvsKZN – providing insight into the selection mechanisms that underwent across populations.

5.5 Discussion

Adaptive selection signatures in rural horse populations are crucial for understanding how these animals have evolved to thrive under specific environmental, management, and cultural conditions (Saravanan et al., 2020). Selection signatures detected within the rural and indigenous populations within this study had associated positive iHS and XP-EHH scores, which typically indicates that recent selection events favour ancestral alleles, rather than the derived alleles that have arisen in the populations (Klassman and Gautier, 2022). Vadenplas et al. (2016) depicts that crossbred horse populations tend to favour selection of ancestral alleles, due to breed-specific effects inherited from composite precursor breeds, in which those ancestral alleles were predominant. The preference for ancestral alleles is commonly observed in purebred horse populations, such as the Thoroughbred breed (Todd et al., 2018), which serves as a defined precursor in South African indigenous and rural horse populations. The environments inhabited by these crossbred populations may further reinforce the selection of these ancestral alleles. In rural production systems, horse populations face unique selective pressures that include extreme climates, limited dietary resources, exposure to endemic diseases, and distinct labour demands (Guyo et al., 2024). Through the identification of selection signatures, genetic adaptations that enable horses to survive and perform in these challenging environments can be determined. The application of iHS (integrated haplotype score) and XP-EHH (cross-population extended haplotype homozygosity) analyses within this study, was fundamental for identifying selection sweeps in these locally-adapted horse populations, providing insight into the beneficial mechanism of the 'hitchhiking' effect of highly selected signatures accumulating neighbouring signatures to be incorporated within the genomic regions (Kim and Maruki, 2011).

Selective sweeps in the BSP population (**Figure 5.2 A** and **Table 5.1**) on chromosome 1 showcased genes involved in fertility (*CFAP58*), nervous system development (*UNC13*), and growth and development (*FBN1*), whereby these beneficial signatures correlate to studies emphasizing the role of selective sweeps in adaptive traits among horse populations (Petersen et al., 2013; Schubert et al., 2014). Positive selection of genes associated with nervous system development plays a crucial role in shaping the behavioural responses and learning abilities of horses, which are essential for adapting to diverse functional roles within their environments (Mendonca et al., 2019). Similarly, adaptive traits associated with cellular processes (*DLG5*; *NID1*) and olfactory receptors (*OR4F14S*) support findings that olfactory-related genes are critical in the

survival and adaptation of horses to detect food resources safe for consumption, social recognition of their surroundings and avoiding danger (Hughes et al., 2018; Bini de Lima et al., 2023). Chromosome 4 in BSP populations also showed enrichment of traits related to blood vessel morphogenesis (*SEMA3E*), aligning with research indicating vascular development as a response to altitude or climate-related selection pressures (Librado et al., 2016) – characteristic of the environment within the Eastern Cape province. Additionally, on chromosome 7, genes linked to cellular processes (*VWA5A*; *ETS1*) further highlight metabolic adaptations that are crucial in varying environmental contexts (Grossman et al., 2013).

The KZN rural breed population (**Figure 5.2 D** and **Table 5.4**) exhibited selective sweeps with overlapping traits in cellular processes and unique adaptations. Notably, *RAB27A* was associated with blood coagulation pathways (Brooks et al., 2010), distinct from the BSP population was detected. The presence of selection signals in olfactory receptor genes (*OR7D20*; *OR1M1*) within the KZN and BSP populations also supports the adaptive role of olfactory systems in behavioural responses (Gutiérrez-Gil et al., 2015). Among rural ECR (**Figure 5.2 B** and **Table 5.2**) and ELR (**Figure 5.2 C** and **Table 5.3**), unique adaptive traits associated with hair follicle development (*SOS1*) in ECR and with locomotory behaviour (*ADAM22*) in ELR populations. This could indicate specialization for environmental stressors such a fluctuation in temperature and movement demands, akin to findings in livestock breed adaptations (Randhawa et al., 2016). Such population-specific traits highlight adaptive differentiation, as found associated studies based on horse populations whereby regional genetic divergence was driven by local environmental factors (Gurgul et al., 2019). Commercial populations, such as the Saddlebred (SAD) (**Figure 5.2 E**) and Thoroughbred (THR) (**Figure 5.2 F**), lacked major selective sweeps, reflecting different selective pressures compared to indigenous and rural populations. Both of these commercial purebred undergo intense selection pressures for athletic performance drives artificial selective mechanisms within these populations and limits their ability to adapt to new or changing environmental challenges (Hanot et al., 2018).

Cross-population XP-EHH analyses provided additional insights into shared adaptive strategies across the breed populations. For instance, the comparison between BSP and ECR (**Figure 5.3 A** and **Table 5.5**) populations revealed strong evidence for selection signatures in genes related to nervous system development and behaviour (*NTM*; *MAST1*; *SOX6*), with several functional annotations linked to blood vessel morphogenesis (*JAM3*) and feed efficiency (*GLB1L2*). These genes are likely important for both metabolic efficiency and neural plasticity in these populations (Han et al., 2023; McBride et al., 2017). The comparison between BSP and KZN (**Figure 5.3 C** and **Table 5.7**) populations showed similar patterns of selection signatures, which further supports the robustness of these adaptations. However, the ECR vs ELR analysis (**Figure 5.4 A** and **Table 5.8**) revealed a decrease in the number of detection signals for genes such as *NTM* and *OPCML*, suggesting that the ELR population may have undergone different evolutionary pressures, potentially due to environmental or historical factors (Fariello et al., 2013). This may also indicate variation in the contribution of purebred precursor breeds within the ELR population, with the selection of dominant ancestral alleles potentially being favoured in that specific regional context (Vandenplas et al., 2016). The

unique selection signatures observed in the ELR vs KZN (**Figure 5.4 C** and **Table 5.10**) comparison, such as those associated with the fear response (*SLC1A1*), whereby fear response genes play a role in stress resistance and behavioural adaptations to environmental changes, particularly in response to predators (Janczarek et al., 2020) or harsh climates (Kang et al., 2023). Rogers and Bell (2022) report that equine behaviourists suggest fear responses in horses are often triggered by stressors, which, in the context of rural horse populations, may include limited access to veterinary care, as well as inadequate food and water availability in low-income households. Similarly, the discovery of selection signatures for nervous system development (*SOX6*) (Hamada-Kanazawa et al., 2004) in the ELR vs KZN analysis suggests that this population may have adapted to unique ecological niches, with specific neural and behavioural adaptations that confer survival advantages.

Combining iHS and XP-EHH analyses (**Table 5.11**) enables the ability to distinguish between within-population selection sweeps and unique selection signatures that differentiate populations (Mohd-Assaad et al., 2018). While iHS reveals selection pressures shared within a population, XP-EHH identifies alleles under selection that may be near fixation in one population but variable in others, often revealing adaptive traits that differ by population (Grossman et al., 2013). This was the case for shared selection on genes involved in neurological development (*OPCML*; *NTM*) within ELR; KZN; BSPvsECR; BSPvsELR; BSPvsKZN and ECRvsELR analyses - suggesting that these traits are universally beneficial. Another universally beneficial gene identified across the KZN, BSP, ECR, and BSPvsECR analyses was *COBL*, which is associated with fertility and reproduction. In rural populations, where structured breeding programs are often lacking, natural selection for fertility and reproductive traits is vital for ensuring the survival and continuity of South Africa's rural and indigenous horse breeds (Hobbs and Hinds, 2018). Therefore, this study underscores the value of combining XP-EHH and iHS analyses to reveal both sweep and unique selection signatures in these indigenous and rural South African horse breed populations, providing deeper insights into their adaptive evolution and functional diversity (O'Connell et al., 2014).

5.6 Conclusion

The combination of iHS and XP-EHH analyses have provided valuable insights into the genetic basis of adaptation in the indigenous, rural and commercial horse breed populations contained within this study. The identification of selection signatures across multiple chromosomes and biological pathways emphasizes the role of genes involved in nervous system development, fertility, immune response, feed efficiency and cellular processes in population-specific adaptations as well as shared adaptions amongst the breed populations. These findings are consistent with previous studies in livestock and human populations, which have highlighted the importance of these pathways in maintaining survival and reproductive success in varying environmental conditions. Selection signatures detected within the rural and indigenous horse populations in this study were characterized by positive iHS and XP-EHH scores, suggesting that recent selection events predominantly favour ancestral alleles over newly arisen derived alleles. This pattern may reflect the influence of composite precursor breeds in crossbred populations, where ancestral alleles were historically dominant and continue to

be favoured under current environmental and functional pressures. The presence of fear repsonses in rural and indigenous populations underscore the urgent need for improved management practices in rural areas, where targeted interventions such as community-based welfare programs, regular health monitoring, and structured human-animal interaction can ultimately enhance the well-being and utility of these horses in their working environments. Future studies incorporating additional populations and environmental factors will help to further elucidate the fundamental basis of adaptation in South African horse breed populations.

5.7 References

Ardestani, S.S., Aminafshar, M., Maryam, M.B.Z.B., Banabazi, M.H., Sargolzaei, M., Miar, Y. (2020). Whole-genome signatures of selection in sport horses revealed selection footprints related to musculoskeletal system development processes. Animals 10(1). Available: https://doi.org/10.3390/ani10010053 [accessed July 30, 2024].

Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57(1), 289-300.

Bini de Lima, A.C., Sebastiao de Fe, V.C., Hernandes, M.S.P., Oliveira dos Santos, V.M. (2023). Olfactory stimulation as environmental enrichment for domestic horses – A review. Animals 13(20). Available: https://doi.org/10.3390/ani13203180 [accessed November 8, 2024].

Brooks, S.A., Gabreski, N., Miller, D., Brisbin, A., Brown, H.E., Streeter, C., Mezey, J., Cook, D., Antczak, D.F. (2010). Whole-genome SNP association in the horse: Identification of a deletion in Myosin Va responsible for lavender foal syndrome. PLOS Genetics 6(4). Available: https://doi.org/10.1371/journal.pgen.1000909 [accessed November 8, 2024].

Browning, B.L., Zhou, Y., Browning, S.R. (2018). A one-penny imputed genome from next-generation reference panels. The American Journal of Human Genetics 103(3), 338-348.

Cook, D., Gallagher, P., Bailey, E. (2010). Genetics of swayback in American Saddlebred horses. Animal Genetics 41, 64–71.

De Simoni Gouveia, J.J., Barbosa da Silva, M.V.G., Paiva, S.R., Pinheiro de Oliveria, S.M. (2014). Identification of selection signatures in livestock species. Genetics and Molecular Biology 37(2), 330-342.

Fariello, M.I., Boitard, S., Naya, H., SanCristobal, M., Servin, B. (2013). Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. Genetics 193(3), 929-941.

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

Grossman, S.R., Andersen, K.G., Shlyakhter, I., Tabrizi, S., Winnicki, S., Yen, A., Park, D.J., Griesemer, D., Karlsson, E.K., Wong, S.H., Cabili, M., Adegbola, R.A., Bamezai, R.N.K., Hill, A.V.S., Vannberg, F.O., Rinn, J.L., Lander, E.S., Schaffner, S.F., and Sabeti, P.C. (2013). Identifying recent adaptions in large-scale genomic data. Cell 152(4), 703-713.

Gurgul, A., Jasielczuk, I., Semik-Gurgul, E., Pawlina-Tyszko, K., Stefaniuk-Szmukier, M., Szmatoła, T., Polak, G., Tomczyk-Wrona, I., Bugno-Poniewierska, M. (2019). A genome-wide scan for diversifying selection signatures in selected horse breeds. PLOS One 14. Available: https://doi.org/10.1371/journal.pone.0210751 [accessed January 28, 2024].

Gutierrez-Gil, B., Arranz, J.J., Wiener, P. (2015). An interpretive review of selective sweep studies in Bos Taurus cattle populations: Identification of unique and shared selection signal across breeds. Frontiers in Genetics 13(6). Available: https://doi.org/10.3389/fgene.2015.00167 [accessed November 8, 2024].

Guyo, M., Tareke, M., Tonamo, A., Bediye, D., Defar, G. (2024). Multivariate analysis of morphometric traits of the horse ecotypes reared in highlands of Bale Zone, Ethiopia. Veterinary and Animal Science 25. Available: https://doi.org/10.1016/j.vas.2024.100391 [accessed November 7, 2024].

Hamada-Kanazawa, M., Ishikawa, K., Nomoto, K., Uozumi, T., Kawai, Y., Narahara, M., Miyake, M. (2004). Sox6 overexpression causes cellular aggregation and the neuronal differentiation of P19 embryonic carcinoma cells in the absence of retinoic acid. FEBS Letters 560(3), 192-198.

Harrison, P. W., Amode, M. R., Austine-Orimoloye, O., Azov, A. G., Barba, M., Barnes, I., Becker, A., Bennett, R., Berry, A., Bhai, J., Bhurji, S. K., Boddu, S., Branco Lins, P. R., Brooks, L., Budhanuru Ramaraju, S., Campbell, L. I., Carbajo Martinez, M., Charkhchi, M., Chougule, K., Cockburn, A., Davidson, C., De Silva, N. H., Dodiya, K., Donaldson, S., El Houdaigui, B., El Naboulsi, T., Fatima, R., Garcia Giron, C., Genez, T., Grigoriadis, D., Ghattaoraya, G. S., Gonzalez Martinez, J., Gurbich, T. A., Hardy, M., Hollis, Z., Hourlier, T., Hunt, T., Kay, M., Kaykala, V., Le, T., Lemos, D., Lodha, D., Marques-Coelho, D., Maslen, G., Merino, G. A., Mirabueno, L. P., Mushtaq, A., Hossain, S. N., Ogeh, D. N., Sakthivel, M. P., Parker, A., Perry, M., Piližota, I., Poppleton, D., Prosovetskaia, I., Raj, S., Pérez-Silva, J. G., Abdul Salam, A. I., Saraf, S., Saraiva-Agostinho, N., Sheppard, D., Sinha, S., Sipos, B., Sitnik, V., Stark, W., Steed, E., Suner, M.-M., Surapaneni, L., Sutinen, K., Tricomi, F. F., Urbina-Gómez, D., Veidenberg, A., Walsh, T. A., Ware, D., Wass, E., Willhoft, N. L., Allen, J., Alvarez-Jarreta, J., Chakiachvili, M., Flint, B., Giorgetti, S., Haggerty, L., Ilsley, G. R., Keatley, J., Loveland, J. E., Moore, B., Mudge, J. M., Naamati, G., Tate, J., Trevanion, S. J., Winterbottom, A., Frankish, A., Hunt, S. E., Cunningham, F., Dyer, S., Finn, R. D., Martin, F. J., Yates, A. D. (2024). Ensembl 2024. Nucleic Acids Research 52, 891-899.

Han, H., Randhawa, I.A.S., MacHugh, D.E., McGivney, B.A., Katz, L.M., Dugarjaviin, M., Hill, E.W. (2023). Selection signatures for local and regional adaption in Chinese Mongolian horse breeds reveal candidate genes for hoof health. BMS Genomics 24. Available: https://doi.org/10.1186/s12864-023-09116-8 [accessed November 8, 2024].

Hanot, P., Herrel, A., Guintard, C., Cornette, R. (2018). The impact of artificial selection on morphological integration in the appendicular skeleton of domestic horses. Journal of Anatomy 232(4), 657-673.

Hobbs, R.J., Hinds, L.A. (2018). Could current fertility control methods be effective for landscape-scale management of populations of wild horses (Equus caballus) in Australia? Wildlife Research 48, 195-207.

Hughes, G.M., Boston, E.S.M., Finarelli, J.A., Murphy, W.J., Higgins, D.G., Teeling, E.C. (2018). The birth and death of olfactory receptor gene families in mammalian niche adaption. Molecular Biology and Evolution 35(6), 1390-1406.

Janczarek, I., Stachurska, A., Kedzierski, W., Wisniewska, A., Ryzak, M., Koziol, A. (2020). The intensity of physiological and behavioural responses of horses to predator vocalizations. BMC Veterinary Research 16(431). Available: https://doi.org/10.1186/s12917-020-02643-6 [accessed November 8, 2024].

Jose de Simoni Gouveia, J., Barbosa da Silva, M.V.G., Paiva, S.R., Pinheiro de Oliveria, S.M. (2008). Identification of selection signatures in livestock species. Genetics and Molecular Biology 37(2), 330-342.

Kalbfleisch, T., Rice, Edward E.S., DePriest, M.S., Walenz, B.P., Hestand, M.S., Vermeesch, J., O'Connell, B., Fiddes, I.T., Vershinina, A., Petersen, J.L., Finno, C.J., Bellone, R., McCue, M.E., Brooks, S.A., Bailey, E., Orlando, L., Green, R.E., Miller, D., Antczak, D., MacLeod, J.N. (2018). EquCab3, an Updated Reference Genome for the Domestic Horse. BioRxiv. Available: http://dx.doi.org/10.1101/306928 [accessed October 16, 2023].

Kang, H., Zsoldos, R.R., Sole-Guitart, A., Narayan, E., Cawdell-Smith, A.J., Gaughan, J.B. (2023). Heat stress in horses: A literature review. International Journal of Biometeorology 67, 957-973.

Kim, Y., Maruki, T. (2011). Hitchhiking effect of a beneficial mutation spreading in a subdivided population. Genetics 189(1), 213-226.

Klassmann, A., Gautier, M. (2022). Detecting selection using extended haplotype homozygosity (EHH)-based statistics in unphased or unpolarized data. PLOS One 17(1). Available: https://doi.org/10.1371/journal.pone.0262024 [accessed May 18, 2025].

Koropoulis, A., Alachiotis, N., Pavlidis, P. (2020) in "Detecting positive selection in populations using genetic data". Statistical Population Genomics. Methods in Molecular Biology, 2090 (eds. Dutheil, J.Y.). Humana, New York, NY. Available: https://doi.org/10.1007/978-1-0716-0199-0_5 [accessed January 28, 2024].

Lee, J.H., Lee, T., Lee, H.K., Cho, B.W., Shin, D.H., Do, K.T., Sung, S., Kwak, W., Kim, H.J., Kim, H., Cho, S., Park, K.D. (2014). Thoroughbred horse single nucleotide polymorphism and expression database: HSDB. Asian-Australian Journal of Animal Sciences 27(9), 1236-1243.

Librado, P., Fages, A., Gaunitz, C., Leonardi, M., Wagner, S., Khan, N., Hanghoj, K., Alquraishi, S.A., Alfarhan, A. H., Al-Rasheid, K.A., Der Sarkissian, C., Schubert, M., Orlando, L. (2016). The evolutionary origin and genetic makeup of domestic horses. Genetics 204(2), 423-434.

Librado, P., Gamba, C., Gaunitz, C., Der Sarkissian, C., Pruvost, M., Albrechtsen, A., Fages, A., Khan, N., Schubert, M., Jagannathan, V., Serres-Armero, A., Kuderna, L.F.K., Povolotskaya, I.S., Seguin-Orlando, A., Lepetz, S., Neuditschko, M., Thèves, C., Alquraishi, S., Alfarhan, A.H., Al-Rasheid, K., Rieder, S., Samashev, Z., Francfort, H.-P., Benecke, N., Hofreiter, M., Ludwig, A., Keyser, C., Marques-Bonet, T., Ludes, B., Crubézy, E., Leeb, T., Willerslev, E., Orlando, L. (2017). Ancient genomic changes associated with domestication of the horse. Science 356, 442-445.

Lukanova, N., Stefanova, K., Stoykova-Grigorova, R. (2018). DMRT3 gene mutation and highly linked SNP in gaitedness horse breeds. Bulgarian Journal of Agricultural Science 24(4), 679-685.

Machmoum, M., Boujenane, I., Azelhak, R., Badaoui, B., Petit, D., Piro, M. (2020). Genetic diversity and population structure of Arabian horse populations using microsatellite markers. Journal of Equine Veterinary Science 93. Available: https://doi.org/10.1016/j.jevs.2020.103200 [accessed July 30, 2024].

McBride, S.D., Parker, M.O., Roberts, K., Hemmings, A. (2017). Applied neurophysiology of the horse; Implications for training, husbandry and welfare. Applied Animal Behaviour Science 190, 90-101.

McGivney, B.A., Han, H., Corduff, L.R., Katz, L.M., Tozaki, T., MacHugh, D.E., Hill, E.W. (2020). Genomic inbreeding trends, influential sire lines and selection in the global Thoroughbred horse population. Scientific Reports 10. Available: https://doi.org/10.1038/s41598-019-57389-5 [accessed January 28, 2024].

Mendonca, T., Bienboire-Frosini, C., Kowalczyk, I., Leclercq, J., Arroub, S., Pageat, P. (2019). Equine activities influence horses' responses to different stimuli: Could this have an impact of equine welfare? Animals (Basel) 9(6). Available: https://doi.org/10.3390/ani9060290 [accessed May 18, 2025].

Metzger, J., Karwath, M., Tonda, R., Beltran, S., Agueda, L., Gut, M., Gut, I.G., Distl, O. (2015). Runs of homozygosity reveal signatures of positive selection for reproduction traits in breed and non-breed horses. BMC Genomics, 16(1). Available: https://doi.org/10.1186/s12864-015-1977-3 [accessed February 1, 2024].

Millin, R. (2011). Saddlebred horse breeding in South Africa. SA Horseman. Available: https://journals.co.za/doi/pdf/10.10520/EJC14633 [accessed January 29, 2024].

Mohd-Assaad, N., McDonald, B.A., Croll, D. (2018). Genome-wide detection of genes under positive selection in worldwide populations of the barley scald pathogen. Genome Biology and Evolution 10(5), 1315-1332.

O'Connell, J., Gurdasani, D., Delaneau, O., Pirastu, N., Ulivi, S., Cocca, M., Traglia, M., Huang, J., Huffman, J. E., Rudan, I., McQuillan, R., Fraser, R. M., Campbell, H., Polasek, O., Asiki, G., Ekoru, K., Hayward, C., Wright, A. F., Vitart, V., Navarro, P., Zagury, J.-F., Wilson, J. F., Toniolo, D., Gasparini, P., Soranzo, N., Sandhu, M. S., Marchini, J. (2014). A general approach for haplotype phasing across the full spectrum of relatedness. PLOS Genetics 10(4). Available: https://doi.org/10.1371/journal.pgen.1004234 [accessed November 8, 2024].

Paul, K., Restoux, G., Phocas, F. (2024). Genome-wide detection of positive and balancing signatures of selection shared by four domesticated rainbow trout populations (Oncorhynchus mykiss). Genetics Selection Evolution 56. Available: https://doi.org/10.1186/s12711-024-00884-9 [accessed November 7, 2024].

Pavlidis, P., Alachiotis, N. (2017). A survey of methods and tools to detect recent and strong positive selection. Journal of Biological Research-Thessaloniki 24. Available: https://doi.org/10.1186/s40709-017-0064-0 [accessed January 28, 2024].

Petersen, J.L., Mickelson, J.R., Cothran, E.G., Andersson, L.S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M.M., Borges, A.S., Brama, P., da Câmara Machado, A., Distl, O., Felicetti, M., Fox-Clipsham, L., Graves, K.T., Guérin, G., Haase, B., Hasegawa, T., Hemmann, K., Hill, E.W., Leeb, T., Lindgren, G., Lohi, H., Lopes, M.S., McGivney, B.A., Mikko, S., Orr, N., Penedo, M.C.T., Piercy, R.J., Raekallio, M., Rieder, S., Røed, K.H., Silvestrelli, M., Swinburne, J., Tozaki, T., Vaudin, M., Wade, C.M., and McCue, M.E. (2013). Genetic diversity in the modern horse illustrated from genome-wide SNP data. PLOS One 8(1). Available: https://doi.org/10.1371/journal.pone.0054997 [accessed November 8, 2024].

Randhawa, I.A.S., Khatkar, M.S., Thomson, P.C., Raadsma, H.W. (2016). A meta-assembly of selection signatures in cattle. PLOS One 11(4). Available: https://doi.org/10.1371/journal.pone.0153013 [accessed November 8, 2024].

Rogers, S., Bell, C. (2022). Perceptions of fear and anxiety in horses as reported in interviews with equine behaviourists. Animals (Basel) 12(21). Available: https://doi.org/10.3390/ani12212904 [accessed May 18, 2025].

RStudio Team. (2020) RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. Available: http://www.rstudio.com/ [accessed September 23, 2024].

Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z.P., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J., Ackerman, H.C., Campbell, S.J., Altshulder, D., Cooper, R., Kwiatkowski, D., Ward, R., Lander, E.S. (2002). Detecting recent positive selection in the human genome from haplotype structure. Nature 419, 832-837.

Saravanan, K.A., Panigrahi, M., Kumar, H., Bhushan, B., Dutt, T., Mishra, B.P. (2020). Selection signatures in livestock genome: A review of concepts, approaches and applications. Livestock Science 241. Available: https://doi.org/10.1016/j.livsci.2020.104257 [accessed November 7, 2024].

Schubert, M., Jónsson, H., Chang, D., Der Sarkissian, C., Ermini, L., Ginolhac, A., Albrechtsen, A., Dupanloup, I., Foucal, A., Petersen, B., Fumagalli, M., Raghavan, M., Seguin-Orlando, A., Korneliussen, T.S., Velazquez, A.M.V., Stenderup, J., Hoover, C.A., Rubin, C.-J., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., MacHugh, D.E., Kalbfleisch, T., MacLeod, J.N., Rubin, E.M., Sicheritz-Ponten, T., Andersson, L., Hofreiter, M., Marques-Bonet, T., Gilbert, M.T.P., Nielsen, R., Excoffier, L., Willerslev, E., Shapiro, B., and Orlando, L. (2014). Prehistoric genomes reveal the genetic foundation and cost of horse domestication. Proceedings of the National Academy of Sciences 11(52), E5661-E5669.

Sherman, B.T., Hao, M., Qiu, J., Jiao, X., Baseler, M.W., Lane, H.C., Imamichi, T., Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Research 50. Available: https://doi.org/10.1093/nar/gkac194 [accessed November 7, 2024].

Swart, S. (2010). “The world the horses made”: A South African case study of writing animals into social history. International Review of Social History 55(2), 241-263.

Szmatola, T., Gurgul, A., Jaseielczuk, Oclon, E., Ropka-Molik, K., Stefanuik-Szukier, M., Polak, G., Wrona-Tomczyk, I., Bugno-Poniewierska, M. (2022). Assessment and distribution of runs of homozygosity in horse breeds representing different utility types. Animals 12(23). Available: https://doi.org/10.3390/ani12233293 [accessed July 30, 2024].

Tiwari, M., Gujar, G., Shashank, C.G., Ponsuksili, S. (2024). Selection signatures for high altitude adaption in livestock: a review. Gene 927. Available: https://doi.org/10.1016/j.gene.2024.148757 [accessed December 18, 2024].

Todd, E.T., Ho, S.Y.W., Thomson, P.C., Ang, R.A., Velie, B.D., Hamilton, N.A. (2018). Founder-specific inbreeding despression affects racing performance in Thoroughbred horses. Scientific Reports 8. Available: https://doi.org/10.1038/s41598-018-24663-x [accessed May 18, 2025].

Vadenplas, J., Calus, M.P.L., Sevillano, C.A., Windig, J.J., Bastiaansen, J.W.M. (2016). Assigning breed origin to alleles in crossbred animals. Genetics Selection Evolution 48(16). Available: https://doi.org/10.1371/journal.pone.0262024 [accessed May 18, 2025].

Voight, B.F., Kudaravalli, S., Wen, X., Pritchard, J.K. (2006). A map of recent positive selection in the human genome. PLOS Biology 4(3). Available: https://doi.org/10.1371/journal.pbio.0040072 [accessed October 30, 2024].

Chapter 6

6.1 General Discussion

6.1.1 Evaluation of economic, socio-economic and cultural roles of rural, indigenous and commercial horse populations

The diverse functional roles of horses in South Africa highlight their economic, socio-economic, and cultural importance across different communities. Historically, horses have served as multi-functional livestock, adapting to evolving societal needs (Skinner et al., 2022). These roles range from agricultural labour and transportation in rural areas to commercial equine sports, leisure activities, and equine-assisted therapy for mental and physical conditions (De Klerk et al., 2020; McShane & Swart, 2011; Letsoalo et al., 2024). In rural areas, horses such as the Basotho-type (BSP), Eastern Cape Rural (ECR), KwaZulu-Natal (KZN), and East London Rural (ELR) are critical for farming and transportation. For agricultural purposes, the KZN rural population exhibited the highest statistical significance (45 %) in supporting small-scale farming, where horse power offers an affordable alternative to machinery (Rydberg & Jansen, 2002). Horses are also vital for transporting goods and people, particularly in challenging terrains such as the mountainous Alfred Nzo district, where 100% of the Basotho-type population was used for transport across farmlands where motorized vehicles could not drive through (de Klerk et al., 2020). Additionally, the indigenous Basotho-type (BSP) pony, served a unique cultural role as a source of meat, particularly among Sotho and Tswana groups, who consume horse meat only from aged or injured animals, adhering to long-standing traditions (Soal, 1999). This meat production role was distinct among rural breeds, as evidence provided by the hierarchical clustering analyses. Rural racing also played a significant role, especially in the Eastern Cape, where events major rural racing events such as the Berlin November Horse Race contribute to socio-economic upliftment, particularly for women-led SMMEs (Ngumbela, 2023).

Commercial breeds such as the Thoroughbred (THR) and Saddlebred (SAD) dominate equine sports, including racing and showing, which significantly boost South Africa's equine industry. The equine sporting industry drives revenue streams in breeding (Scholtz et al., 2022), private equine veterinary practices (Marlow, 2010) and animal feed industry (Pienaar et al., 2021). Clustering analyses revealed the phenotypic specialization of these breeds, such as the high-arched necks of Saddlebreds suited for gaited flares within the arena and the musculature of Thoroughbreds optimized for racing performance (Staiger et al., 2016; Zsoldos and Licka, 2015). Phenotypic-trait analyses from the CATPCA revealed that rural breeds retain short, stocky frames suited for draught power and rugged terrain, while some specimens lacked musculature and other traits that were influenced by environmental factors like limited nutrition (Getchell et al., 2002). Ordinal regression analysis showed that rural horse populations do not significantly impact household income directly, as most rural households operate on subsistence farming with low-income levels (Mamabolo et al., 2024). However, rural horses contribute indirectly by enhancing livelihoods through agricultural productivity and mobility (De Klerk

et al., 2020). Commercial breeds, in contrast, significantly impact South Africa's multi-million-rand equine industry, reflecting their larger economic and socio-economic footprint (Swart, 2010).

6.1.2 Determination of the genetic diversity of South African horse breed populations

Investigation of the genetic diversity and population structure of indigenous, rural, and purebred commercial horse breed populations, yielded detailed insight into the population structure, differentiation and diversity of the six South African breed populations used within this study. South Africa PCA analysis identified four genetic clusters showing distinct groupings for Thoroughbreds (THR), SA Saddlebreds (SAD), Basotho ponies (BSP), and rural breeds (ECR, KZN and ELR) - highlighting varying degrees of relatedness to one another. Cluster 1 exclusively contained Thoroughbred (THR) breed samples, which clustered separately from the rest of the dataset. This outcome was expected, as South African Thoroughbred racehorses have been a closed breeding stock since 1906, when the first volume of the SA Studbook was published and the Jockey Club of South Africa was established to prevent crossbreeding with other horse breeds for flat racing (Herringer, 2004). This closed breeding system has intensified selection pressures and limited the gene pool (Todd et al., 2020), as shown by the uniform ancestral proportions in the sNMF analysis. The purebred, commercial SA Saddlebred (SAD) population clustered exclusively within Cluster 4, which was spatially positioned adjacent to the indigenous and rural population Cluster 2. This proximity suggests that, despite being recognized as a distinct purebred breed in South Africa (Millin, 2011), the results presented indicate a potential degree of genetic relatedness between these clusters. Shared ancestry and gene flow among indigenous and rural breeds were shaped by geographic proximity and historical migration (Mokoena & Balkaran, 2018). The global PCA further revealed the collective genetic contributions from Asian, European, and American horse breeds to South African populations, influenced by their use in warhorse breeding improvement (Pozharskiy et al., 2023; Falkingham, 2018). The intra-population parameter (F_{IS}) from the AMOVA analysis showed a high genetic variation of F_{IS} = 94.390 % within the six representative breed populations. This indicates that while there was ongoing gene flow among the populations, individuals within each population retained a high degree of genetic diversity (Pongratz et al., 2002).

Runs of homozygosity (ROH) analyses revealed adaptive and selective pressures across populations, which provided a level of the determination of diversity within each breed population. Indigenous and rural breeds exhibited shorter ROH stretches due to ancient contiguous genomic regions that have been continuously shaped by environmental adaptations (McQuillan et al., 2008), while commercial breeds showed longer ROH due to intensive selective breeding (Chen et al., 2023). ROH islands identified genes related to adaptive traits within the rural and indigenous breed populations that included immunity (*FIZ1*; *SSC5D*; *ZFP90*), fertility and reproduction (*NLRP13*; *NLRP5*) and growth and development traits (*EPN1*; *SNAI1*). Both commercial, purebred populations contained breed-specific selection signatures associated with the breed's sporting discipline, however, two deleterious genes (*MAGI1*; *ADGRL3*) (Gmel et al., 2023; Posbergh et al., 2018) were detected within two islands carried within the SAD population, were detected. McQuillan et al. (2008) suggested using total genomic ROH length to assess inbreeding levels (F_{ROH}), offering a more accurate

measure of genetic diversity than traditional pedigree estimates (Forutan et al., 2018). In the East London Rural (ELR) population, F_{ROH} levels varied from 0.102 to 0.326, reflecting variation in inbreeding patterns amongst the rural specimens. Szwaczkowski et al. (2016) noted that uncontrolled mating in rural populations can lead to higher inbreeding and increased heterosis, explaining the variation seen in the ELR population. The KZN population showed moderate inbreeding (F_{ROH} = 0.169–0.239) with ROH islands carrying adaptive traits critical for reproduction and survival (de Simoni Gouveia et al., 2014). The BSP population demonstrated low inbreeding (F_{ROH} < 0.100), indicating maintained genetic diversity (Chen et al., 2023). The commercial SAD population had the highest inbreeding (F_{ROH} = 0.133–0.242), with specific autosomes showing high inbreeding and deleterious genes that could reduce fitness and performance (Bosse et al., 2019). The findings emphasize the interconnectedness of South African horse breeds due to shared ancestry and selective breeding, underscoring the need for sustainable management of genetic diversity. This aligns with previous studies documenting the evolutionary and ecological factors influencing equine populations (Shen et al., 2019; Schaefer & McCue, 2020).

6.1.3 Identification of selection signatures found within and between South African horse breed populations

This study emphasizes the importance of identifying adaptive selection signatures in rural horse populations to understand how these animals have evolved to adapt to environmental and management challenges (Saravanan et al., 2020). Horses in rural production systems face unique pressures, such as harsh climates, limited resources, endemic diseases, and specific labour requirements, necessitating adaptive traits for survival and performance (Guyo et al., 2024). The use of iHS (integrated haplotype score) and XP-EHH (cross-population extended haplotype homozygosity) analyses revealed significant selection sweeps and their associated traits within these populations, which may be due to the 'hitchhiking' effect, where selected alleles accumulate neighbouring alleles (Kim and Maruki, 2011). In the BSP population, selection sweeps on chromosome 1 indicated genes related to fertility (*CFAP58*), nervous system development (*UNC13*), and growth (*FBN1*), aligning with prior studies on adaptive traits that are necessary in horse populations (Petersen et al., 2013; Schubert et al., 2014). Traits associated with olfactory functions (OR4F14S) and blood vessel morphogenesis (*SEMA3E*) were detected, the latter being relevant for adaptation to climate or altitude (Librado et al., 2016) that is characteristic of the environmental layout of the Eastern Cape. Genes on chromosome 7 linked to cellular processes (*VWA5A*; *ETS1*) highlighted metabolic adaptations (Grossman et al., 2013). The KZN rural breed displayed distinct selection signals, including *RAB27A* related to blood coagulation (Brooks et al., 2010), and overlapping adaptive traits in olfactory receptor genes (*OR7D20*; *OR1M1*) similar to BSP, confirming the importance of olfactory adaptations for survival and behaviour (Gutiérrez-Gil et al., 2015). Unique adaptive traits were observed in ECR (*SOS1* - related to hair follicle development) and ELR (*ADAM22* linked to locomotion), suggesting specialization in temperature and movement responses (Randhawa et al., 2016). These population-specific traits underscore adaptive differentiation, as supported by studies on horse populations in which regional genetic divergence was driven by local environmental factors (Gurgul et al.,

2019). Commercial SAD and THR breed populations displayed fewer selection sweeps, reflecting their intense selection for athletic traits, which limits adaptation to new environmental challenges (Hanot et al., 2018).

XP-EHH analyses compared across populations showed shared adaptive traits, such as genes for nervous system development (*NTM*; *MAST1*; *SOX6*) and metabolic efficiency (*GLB1L2*) in BSP vs ECR populations, linking these to neural plasticity and survival strategies (Han et al., 2023; McBride et al., 2017). The ECR vs ELR comparison highlighted different evolutionary pressures, particularly genes related to fear response (*SLC1A1*), which play roles in stress resistance (Janczarek et al., 2020; Kang et al., 2023). The ELR vs KZN analysis noted genes such as *SOX6*, suggesting unique neural adaptations (Hamada-Kanazawa et al., 2004). Combining iHS and XP-EHH analyses allowed for distinguishing population-wide sweeps from unique selection signatures (Mohd-Assaad et al., 2018). These results illustrate the value of these combined methods in exploring functional diversity and adaptive evolution in indigenous and rural South African horse populations (O'Connell et al., 2014).

6.2 References

Bosse, M., Megens, H.J., Derks, M.F.L., de Cara, A.M.R., Groenen, M.A.M. (2019). Deleterious alleles in the context of domestication, inbreeding, and selection. Evolutionary Applications 12(1), 6-17.

Brooks, S.A., Gabreski, N., Miller, D., Brisbin, A., Brown, H.E., Streeter, C., Mezey, J., Cook, D., Antczak, D.F. (2010). Whole-genome SNP association in the horse: Identification of a deletion in Myosin Va responsible for lavender foal syndrome. PLOS Genetics 6(4). Available: https://doi.org/10.1371/journal.pgen.1000909 [accessed November 8, 2024].

Chen, C., Zhu, B., Tang, X., Chen, B., Liu, M., Gao, N., Li, S., Gu, J. (2023). Genome-wide assessment of runs of homozygosity by whole-genome sequencing in diverse horse breeds worldwide. Genes 14(6). Available: https://doi.org/10.3390%2Fgenes14061211 [accessed September 30, 2024].

De Klerk, J.N., Quan, M., Grewar, J.D. (2020). Socio-economic impacts of working horses in urban and peri-urban areas of the Cape Flats, South Africa. Journal of the South African Veterinary Association 91. Available: https://doi.org/10.4102/ jsava.v91i0.2009 [accessed November 25, 2024].

De Simoni Gouveia, J.J., Barbosa da Silva, M.V.G., Paiva, S.R., Pinheiro de Oliveria, S.M. (2014). Identification of selection signatures in livestock species. Genetics and Molecular Biology 37(2), 330-342.

Falkingham, L. (2018). 12 New Forest pony facts every native fan needs to know. Horse and Hound. Available: https://www.horseandhound.co.uk/features/new-forest-pony-facts-642356 [accessed September 30, 2024].

Forutan, M., Mahyari, S.A., Baes, C., Melzer, N., Schenkel, F.S., Sargolzaei, M. (2018). Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. BMC Genetics 19(98). Available: https://doi.org/10.1186/s12864-018-4453-z [accessed November 16, 2023].

Getchell, J.K., Vatta, A.F., Motswatswe, P.W., Krecek, R.C., Moerane, R., Pell, A.N., Tucker, T.W., Leshomo, S. (2002). Raising livestock in resource-poor communities of the North West Province of South Africa – a participatory rural appraisal study. Journal of the South African Veterinary Association 73(4), 177-184.

Gmel, A.I., Brem, G., Neuditschko, M. (2023). New genomic insights into the conformation of Lipizzan horses. Scientific Reports 13. Available: https://doi.org/10.1038/s41598-023-36272-4 [accessed September 22, 2024].

Grossman, S.R., Andersen, K.G., Shlyakhter, I., Tabrizi, S., Winnicki, S., Yen, A., Park, D.J., Griesemer, D., Karlsson, E.K., Wong, S.H., Cabili, M., Adegbola, R.A., Bamezai, R.N.K., Hill, A.V.S., Vannberg, F.O., Rinn, J.L., Lander, E.S., Schaffner, S.F., and Sabeti, P.C. (2013). Identifying recent adaptions in large-scale genomic data. Cell 152(4), 703-713.

Gurgul, A., Jasielczuk, I., Semik-Gurgul, E., Pawlina-Tyszko, K., Stefaniuk-Szmukier, M., Szmatoła, T., Polak, G., Tomczyk-Wrona, I., Bugno-Poniewierska, M. (2019). A genome-wide scan for diversifying selection signatures in selected horse breeds. PLOS One 14. Available: https://doi.org/10.1371/journal.pone.0210751 [accessed January 28, 2024].

Gutierrez-Gil, B., Arranz, J.J., Wiener, P. (2015). An interpretive review of selective sweep studies in Bos Taurus cattle populations: Identification of unique and shared selection signal across breeds. Frontiers in Genetics 13(6). Available: https://doi.org/10.3389/fgene.2015.00167 [accessed November 8, 2024].

Guyo, M., Tareke, M., Tonamo, A., Bediye, D., Defar, G. (2024). Multivariate analysis of morphometric traits of the horse ecotypes reared in highlands of Bale Zone, Ethiopia. Veterinary and Animal Science 25. Available: https://doi.org/10.1016/j.vas.2024.100391 [accessed November 7, 2024].

Hamada-Kanazawa, M., Ishikawa, K., Nomoto, K., Uozumi, T., Kawai, Y., Narahara, M., Miyake, M. (2004). Sox6 overexpression causes cellular aggregation and the neuronal differentiation of P19 embryonic carcinoma cells in the absence of retinoic acid. FEBS Letters 560(3), 192-198.

Han, H., Randhawa, I.A.S., MacHugh, D.E., McGivney, B.A., Katz, L.M., Dugarjaviin, M., Hill, E.W. (2023). Selection signatures for local and regional adaption in Chinese Mongolian horse breeds reveal candidate genes for hoof health. BMS Genomics 24. Available: https://doi.org/10.1186/s12864-023-09116-8 [accessed November 8, 2024].

Hanot, P., Herrel, A., Guintard, C., Cornette, R. (2018). The impact of artificial selection on morphological integration in the appendicular skeleton of domestic horses. Journal of Anatomy 232(4), 657-673.

Herringer, P. (2004). Thoroughbred horse racing and breeding in South Africa. Thoroughbred Heritage. Available: https://www.tbheritage.com/Breeders/SABreeding/SABreeders.html [accessed September 13, 2022].

Janczarek, I., Stachurska, A., Kedzierski, W., Wisniewska, A., Ryzak, M., Koziol, A. (2020). The intensity of physiological and behavioural responses of horses to predator vocalizations. BMC Veterinary Research 16(431). Available: https://doi.org/10.1186/s12917-020-02643-6 [accessed November 8, 2024].

Kang, H., Zsoldos, R.R., Sole-Guitart, A., Narayan, E., Cawdell-Smith, A.J., Gaughan, J.B. (2023). Heat stress in horses: A literature review. International Journal of Biometeorology 67, 957-973.

Kim, Y., Maruki, T. (2011). Hitchhiking effect of a beneficial mutation spreading in a subdivided population. Genetics 189(1), 213-226.

Letsoalo, D.L., Menoe, M.B., Semenya, M.V. (2024). Equine-assisted therapy in South Africa: A scoping review of applications and effectiveness. E-Journal of Humanities, Arts and Social Sciences 5(13), 2328-2343.

Librado, P., Fages, A., Gaunitz, C., Leonardi, M., Wagner, S., Khan, N., Hanghoi, K., Alquraishi, S.A., Alfarhan, A.H., Al-Rasheid, K., Der Sarkissian, C., Schubert, M., Orlando, L. (2016). The evolutionary origin and genetic makeup of the domestic horse. Genetics 204(2), 423-434.

Mamabolo, M., Chaminuka, P., Machethe, C. (2024). District analysis of poverty dynamics in rural South Africa. African Journal of Science, Technology, Innovation and Development 15(6). Available: https://hdl.handle.net/10520/ejc-aa_ajstid_v15_n6_a690 [accessed December 05, 2024].

Marlow, C.H.B. (2010). A brief history of equine private practice in South Africa. Journal of the South African Veterinary Association 81(4), 190-200.

McBride, S.D., Parker, M.O., Roberts, K., Hemmings, A. (2017). Applied neurophysiology of the horse; Implications for training, husbandry and welfare. Applied Animal Behaviour Science 190, 90-101.

McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., MacLeod, A.K., Farrington, S.M., Rudan, P., Hayward, C., Vitart, V., Rudan, I., Wild, S.H., Dunlop, M.G., Wilson, J.F. (2008). Runs of homozygosity in European populations. AJHG 83(3), 359-372.

McShane, C., Swart, S. (2011). Designing equids in South Africa and North America. Safundi 12(2), 203-227.

Millin, P. (2007). The renowned Basotho pony. SA Horseman. Available: https://journals.co.za/doi/pdf/10.10520/EJC14081 [accessed September 08, 2022].

Millin, R. (2011). Saddle horse breeding in South Africa. SA Horseman, 10-13.

Mohd-Assaad, N., McDonald, B.A., Croll, D. (2018). Genome-wide detection of genes under positive selection in worldwide populations of the barley scald pathogen. Genome Biology and Evolution 10(5), 1315-1332.

Mokoena, S.K., Balkaran, S. (2018). An exploration of constraints for free movement of people in Africa – A case of Lesotho and South Africa. African Journal of Public Affairs 10(2), 109-126.

Ngumbela, X.G. (2023). Empowering rural women through tourism entrepreneurship: The case of the Annual Berlin November horse racing in the Eastern Cape Province. International Journal of Social Science Research and Review 6(11), 195-207.

O'Connell, J., Gurdasani, D., Delaneau, O., Pirastu, N., Ulivi, S., Cocca, M., Traglia, M., Huang, J., Huffman, J. E., Rudan, I., McQuillan, R., Fraser, R. M., Campbell, H., Polasek, O., Asiki, G., Ekoru, K., Hayward, C., Wright, A. F., Vitart, V., Navarro, P., Zagury, J.-F., Wilson, J. F., Toniolo, D., Gasparini, P., Soranzo, N., Sandhu, M. S., Marchini, J. (2014). A general approach for haplotype phasing across the full spectrum of relatedness. PLOS Genetics 10(4). Available: https://doi.org/10.1371/journal.pgen.1004234 [accessed November 8, 2024].

Petersen, J.L., Mickelson, J.R., Cothran, E.G., Andersson, L.S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M.M., Borges, A.S., Brama, P., da Câmara Machado, A., Distl, O., Felicetti, M., Fox-Clipsham, L., Graves, K.T., Guérin, G., Haase, B., Hasegawa, T., Hemmann, K., Hill, E.W., Leeb, T., Lindgren, G., Lohi, H., Lopes, M.S., McGivney, B.A., Mikko, S., Orr, N., Penedo, M.C.T., Piercy, R.J., Raekallio, M., Rieder, S., Røed, K.H., Silvestrelli, M., Swinburne, J., Tozaki, T., Vaudin, M., Wade, C.M., McCue, M.E. (2013). Genetic diversity in the modern horse illustrated from genome-wide SNP data. PLOS One 8(1). Available: https://doi.org/10.1371/journal.pone.0054997 [accessed November 16, 2023].

Pienaar, L., Meyer, F., Otterman, H., Davids, T. (2021). Deep-dive of the animal feed industry in South Africa: Strategic analysis. Report commissioned by the animal feed manufacturers association of South Africa. Pretoria: Bureau for Food and Agricultural Policy. Available: https://www.researchgate.net/publication/363670158_The_South_African_Feed_Industry_A_Strategic_Perspective [accessed December 04, 2024].

Pongratz, N., Gerace, L., Michiels, N.K. (2002). Genetic differentiation within and between populations of a hermaphroditic freshwater planarian. Hereditary 89, 64-69.

Posbergh, C.J., Pollott, G.E., Southard, T.L., Divers, T.J., Brooks, S.A. (2018). A nonsynonymous change in adhesion G protein-coupled receptor L3 associated with risk for equine degenerative myeloencephalopathy in the Caspian horse. Journal of Equine Veterinary Science 70, 96-100.

Pozharskiy, A., Abdrakhmanova, A., Beishova, I., Shamshindin, A., Nametov, A., Ulyanova, T., Bekova, G., Kikebayev, N., Kovalchuk, A., Ulyanov, V., Turabayev, A., Khusnitdinova, M., Zhambakin, K., SapaKhova, Z., Shamekova, Gritsenko, D. (2023). Genetic structure and genome-wide association study of the traditional Kazakh horses. Animal 17(9). Available: https://doi.org/10.1016/j.animal.2023.100926 [accessed September 21, 2024].

Randhawa, I.A.S., Khatkar, M.S., Thomson, P.C., Raadsma, H.W. (2016). A meta-assembly of selection signatures in cattle. PLOS One 11(4). Available: https://doi.org/10.1371/journal.pone.0153013 [accessed November 8, 2024].

Rydberg, T., Jansen, J. (2002). Comparison of horse and tractor traction using emergency analysis. Ecological Engineering 19(1), 13-28.

Saravanan, K.A., Panigrahi, M., Kumar, H., Bhushan, B., Dutt, T., Mishra, B.P. (2020). Selection signatures in livestock genome: A review of concepts, approaches and applications. Livestock Science 241. Available: https://doi.org/10.1016/j.livsci.2020.104257 [accessed November 7, 2024].

Schubert, M., Jónsson, H., Chang, D., Der Sarkissian, C., Ermini, L., Ginolhac, A., Albrechtsen, A., Dupanloup, I., Foucal, A., Petersen, B., Fumagalli, M., Raghavan, M., Seguin-Orlando, A., Korneliussen, T.S., Velazquez, A.M.V., Stenderup, J., Hoover, C.A., Rubin, C.-J., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., MacHugh, D.E., Kalbfleisch, T., MacLeod, J.N., Rubin, E.M., Sicheritz-Ponten, T., Andersson, L., Hofreiter, M., Marques-Bonet, T., Gilbert, M.T.P., Nielsen, R., Excoffier, L., Willerslev, E., Shapiro, B., and Orlando, L. (2014). Prehistoric genomes reveal the genetic foundation and cost of horse domestication. Proceedings of the National Academy of Sciences 11(52), E5661-E5669.

Schaefer, R.J., McCue, M.E. (2020). Equine genotyping arrays. Veterinary Clinics of North America: Equine Practice 36(2), 183-193.

Scholtz, M., Guthrie, A.J., Fosgate, G.T., Schulman, M.L. (2022). Predictors of foaling outcomes in barren and maiden Thoroughbred mares in South Africa. Journal of the South African Veterinary Association 93(2), 168-175.

Shen, Y., Wang, L., Fu, J., Xu, X., Yue, G.H., Li., J. (2019). Population structure, demographic history and local adaption of the grass carp. BMC Genomics 20 (467). Available: https://doi.org/10.1186/s12864-019-5872-1 [accessed November 15, 2023].

Skinner, J.E., Hilly, L., Li, X., Cawdell-Smith, A.J., Bryden, W.L. (2022) in "Equine production systems and the changing role of horses in society". Livestock: Production, Management Strategies and Challenges (Editors V.R. Squires and W.L. Bryden). Nova Science Publishers, NY. pp. 389-433.

Soal, J. (1999). Meaty prejudices give food for thought. IOL. Available: https://www.iol.co.za/news/south-africa/meaty-prejudices-give-food-for-thought-20109#:~:text=Horse%20meat%20is%20part%20of,%22Meat%20is%20meat [accessed December 04, 2024].

Staiger, E.A., Bellone, R.R., Sutter, N.B., Brooks, S.A. (2016). Morphological variation in gaited horse breeds. Journal of Equine Veterinary Science 43, 55-65.

Swart, S. (2010). "The world the horses made": A South African case study of writing animals into social history. International Review of Social History 55(2), 241-263.

Szwaczkowski, T., Gregula-Kania, M., Stachruska, A., Borowska, A., Jaworski, Z., Gruszecki, T.M. (2016). Inter-and intra-genetic diversity in the Polish Konik horse: implications for the conservation program. Canadian Journal of Animal Science 96, 570-580.

Todd, E.T., Hamilton, N.A., Velie, B.D., Thomson, P.C. (2020). The effects of inbreeding on covering success, gestation length and foal sex ratio in Australian Thoroughbred horses. BMC Genetics 21 (41). Available: https://doi.org/10.1186/s12863-020-00847-1 [accessed November 16, 2023].

Zsoldos, R.R., Licka, T. (2015). The equine neck and it function during movement and locomotion. Zoology 118(5), 346-376.

Chapter 7

7.1 Conclusions and Recommendations

This study underscores the multi-faceted roles that horses perform in South African society, touching upon economic, socio-economic, and cultural dimensions. Horses are valuable assets, particularly in rural areas, where they are utilized for agricultural work, transportation and traditional recreational activities such as rural racing. The Basotho-type (BSP), Eastern Cape Rural (ECR), KwaZulu-Natal (KZN) Rural, and East London Rural (ELR) populations have shown their functional diversity and essential roles in supporting rural livelihoods. Unique cultural practices, such as the consumption of horse meat within the Sotho and Tswana communities, illustrate how deeply integrated horses are within local traditions. The socio-economic impact of rural racing, particularly in the ELR population, has been notable, promoting community cohesion and empowering small enterprises, including women-led SMMEs. Despite these contributions, the direct impact on household income is relatively limited, reflecting broader challenges faced by low-income rural communities. Commercial horse populations, such as the Saddlebred (SAD) and Thoroughbred (THR) breeds, play a critical role in South Africa's thriving equine sporting industry - a multi-million rand sector that bolsters the economy through high-stakes racing and show events and supports related industries involved within breeding, veterinary services, and animal feed manufacturing. This highlights the adaptability and multi-functionality of horses in meeting the evolving needs of South African communities, from rural farmlands to commercial sectors.

The application of genomic tools, such as the Equine SNP70K and SNP80K BeadChip, has provided crucial insights into the genetic diversity of these horse populations, revealing low inter-population diversity and high intra-population diversity between the South African breeds. While the study identified recent and historical selection events through ROH stretches, it also highlighted concerning inbreeding levels, especially in the ELR and SAD populations, where high inbreeding may affect overall fitness. Conversely, the THR population demonstrated low inbreeding levels, preserving beneficial genetic signatures despite intense selective pressures. Furthermore, the use of iHS and XP-EHH analyses have enhanced the understanding of adaptive traits among indigenous and rural horse breed populations. These analyses pointed to genes associated with nervous system development, fertility, immune response, feed efficiency, and other cellular processes that support survival and adaptation. These findings align with research in other livestock and human populations, reinforcing the critical role of these pathways for resilience and reproductive success under varying environmental conditions. The practical implications are clear; understanding the genetic basis and potential of these horse breeds can aid in breeding and managing more resilient and well-adapted populations, supporting rural livelihoods and sustainable agricultural practices.

Future research should focus on expanding the genetic and functional characterization of rural horse populations across the country. Assessments should include variability studies for all regional populations and investigate how environmental factors contribute to genetic diversity of each locally adapted population. This

approach will deepen insights into population-specific and shared adaptive mechanisms and support strategies for maintaining the health and sustainability of South Africa's diverse equine populations.