Unravelling Genes Associated with Coat Colour and Coat Colour Patterns in indigenous South African Meat-type Goats.

by

Sithembisile Gcabashe

(216018468)

June 2024

A thesis submitted in accordance with the requirements for

Master of science

in Genetics

School of Life Sciences

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg

South Africa

DECLARATIONS

DECLARATION 1 – PLAGIARISM

I, **Sithembisile Lwazi Gcabashe** declare that

1. The research reported in this thesis is my original research, except where otherwise indicated.
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:
   a. their words have been re-written but the general information attributed to them have been referenced, and
   b. 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 section/s.

Signed …[redacted]……

Sithembisile Lwazi Gcabashe Date: 14 June 2024

[redacted]

Signed: Edgar Farai Dzomba Date: 12 June 2024

[redacted]

Signed: Farai Catherine Muchadeyi Date: 12 June 2024

ACKNOWLEDGMENTS

I would like to acknowledge and express my many thanks to the following individuals for assisting me complete this work:

Dr EF Dzomba and Dr FC Muchadeyi, who introduced me to the study of animal genomics, and continuously supported my work, providing me with advice and guidance which improved my skills as a researcher over the years it has taken me to complete this work.

The University of KwaZulu-Natal, and the National Research Foundation, who financially supported this study, as well as the Agricultural Research Council-Biotechnology Platform who provided with the platform, tools and environment to complete this work.

Many thanks to the whole Agricultural Research Council-Biotechnology Platform team (students and staff) for their assistance with key aspects of my research including sampling, genotyping, and data analysis. A special thank you to Dr K Hadebe, Dr N Dlamini, and Dr N Hlongwane for sharing their research expertise which allowed me to complete my work.

I would also like to thank my family: My mother, brothers, nieces (Ntandokazi, Minenhle, and Melokuhle) and my girlfriend (Mpilonhle) for the love, sacrifices and support they provided which have allowed me to complete my studies.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

Abstract

Small ruminants such as goats are a vital part of the economy, especially that of farmers in rural communities, where people have successfully bred these animals for sustenance despite the limiting conditions found in extensive production systems which offer little protection from harsh environmental conditions and limited feed. South African meat-type ecotypes, which are primarily found in communal production systems have managed to strive under such adverse conditions due to their high genetic diversity and adaptability to a wide range of environments. Due to these characteristics, the ecotypes are economically important as a livestock species because they hold high production potential, especially in the face of drastically changing environmental conditions caused by global warming and an increase in meat demand due to the rising human population.

However, the lack of both information about proper breeding strategies, and investment capital within communal goat production systems inhibits the adoption of modern technologies, necessary to implement improved breeding schemes that utilise genetic, and phenotypic information to employ genomic selection. Thus, it is necessary to generate low-cost and easy-adoption selection technologies according to low-input communal production system requirements. A key step towards this goal is the description of the genetic composition of village goat populations. This genetic description involves the study of the genes that influence goat production traits. The detection of genes responsible for economical traits such as coat colour can help open opportunities towards implementing improved selection schemes with improved selection accuracy and intensity allowing for the early selection of reproducers within communal goat production systems.

This study used 51 767 SNPs from 329 indigenous goats with various coat colour and coat colour patterns to investigate genes responsible for goat coat colour. Case/control GWASs carried out for all the coat colours/patterns using the GAPIT revealed several genes associated with various coat colours including white (*CDK5*), red (*CELF5, TLE6*), black (*CACNA2D1*), grey (*GSK3B*), and coat colour patterns such as white-body red head (*CADPS2, SLC13A1*), speckled (*KIT, TYRP1*), patchy (*GNAI3*), belted (*TYRP1*), white-sided (*AHCY*), and blacklegged (*AHCY*). Golden Helix SVS's univariate method from the copy number analysis module (CNAM) detected 2 047 CNVs and 279 CNVRs overlapping 5 222 genes involved in several biological processes, molecular functions, cellular

components, and pathways. Several coat colour genes overlapping these CNVRs included *GSK3B, Notch1, Notch2, CDK5, ADAMTS20, TYRO3, MAP2K1, ITCH, ASIP, AHCY, SLC45A3, EDNRA, ADCY2*, and *TYR*. Candidate coat colour genes were found to be under selection within (iHS) and between (XP-EHH) the goats. These genes were involved in melanogenesis pathways such as the MAPK signalling pathway (*DUSP16, KDR, FGF10, MAP3K7*), the PI3K-Akt pathway (*CDKN1B, COL6A3, CRK, CREB5*), the Wnt signalling pathway (SERPINF1, WNT2), the Notch signalling pathway (*HEY2*), and Melanogenesis (*TYRP1*). In addition, genes associated with various traits such as metabolism (*FXN, FTO, IRX3, EDNRA*), reproduction (*PGR*), growth and development (*SMARCAD1*), immune response (*ACAD8, DLG4, GPS2*), and environmental adaptation (*CNR1*) were detected, suggesting a possible link between coat colour and goat productivity. This study's findings were in line with previous studies which have revealed that coat colour is a polygenic trait whose variation is influenced by the epistatic effects of multiple genes involved in various regulatory pathways that affect melanogenesis. Coat colour is a key factor driving the economic efficacy of indigenous goats. Therefore, knowledge of its genetic mechanisms is vital as it will allow farmers to utilize breeds with specific environmental adaptations in selection strategies for genetic improvement.

Keywords: goats, coat colour, improved breeding schemes, GWAS, selection (iHS and XP-EHH), CNV

Chapter 1: Introduction

Coat colour is the easiest trait to identify on an animal. Hence, the trait often forms part of breed definition and characterisation. It is also of economic importance as it can determine the selling price of a livestock animal. Beyond that, coat colour has also been reported to play a role in heat tolerance in some animals (Peters, 1980). Genetic studies on coat colour have identified some of the loci involved in coat pigmentation in goats, and these include melanocyte-stimulating hormone receptor (*MC1R*) and agouti signalling protein (*ASIP* or Agouti) (Martin et al., 2016; Fontanesi et al., 2009; Barsh, 1996).

These regions have been reported to influence the type of pigment produced by the *MC1R* protein. Under normal circumstances, the *MC1R* protein produces the dark pigment of eumelanin (Martin et al., 2016; Fontanesi et al., 2009; Barsh, 1996). However, the presence of ASIP causes a switch to light pheomelanin production (Martin et al., 2016; Fontanesi et al., 2009; Barsh, 1996). Despite this knowledge, the biological mechanisms underlying coat colour patterns are still poorly understood, and the factors controlling coat colour remain complex.

In South African commercial meat-type goats (i.e., The Boer, the Savanna, and the Kalahari Red), coat colour is an integral part of set breed standards used by breeders' associations for the characterisation of these breeds. These breeds' characterisation is vital to ensure that their unique traits and genetic resources are not lost due to continuous selection and crossbreeding. The Boer goat breed standards were first described by the South African Boer Goat Association, founded in 1959 (Visser and van Marle-Koster, 2018; Pieters et al., 2009; Campbell, 2003). These standards specify a white body and a redhead (Visser and van Marle-Koster, 2018; Pieters et al., 2009; Campbell, 2003).

The Kalahari Red breed standards specify a uniform, solid, red colour (Visser and van Marle-Koster, 2018; Pieters et al., 2009; Campbell, 2003). This dark red coat is believed to provide the goats with heat resistance to endure heat and intense sunlight (Pieters et al., 2009). The Savanna goat's breed standards specify a white coat (Visser and van Marle-Koster, 2018; Pieters et al., 2009; Campbell, 2003). During the winter, the goats develop extra fluffy cashmere hair for protection (Pieters et al., 2009).

Ecotypes made up of goat populations kept in villages across the country, for various breeding goals such as meat production for home consumption and use in cultural ceremonies, have not been well characterised (Ncube et al., 2020; Mdladla et al., 2016b). These populations do not

undergo selection for improvement; therefore, breed standards do not exist. However, selection for coat colour has been reported (Mdladla et al., 2017). This is likely due to the high demand for specific-coloured goats for cultural ceremonies. There is a need for the detailed study of goat coat colour patterns at the genome level due to these reasons. Such a study can provide information that can inform the breed identification and characterisation of these goat breeds.

While most characterisation studies on South African meat-type goats have focused on genetic diversity and population structure (Chokoe et al., 2020a; Chokoe et al., 2020b; Monau et al., 2020; Ncube et al., 2020; Mdladla et al., 2016a; Visser et al., 2016; Bosman et al., 2015a; Bosman et al., 2015b), few genome-wide studies for production and disease-related traits exist. Thus, goat genetic improvement has been slow compared to other livestock species in South Africa, such as cattle and sheep. Goat genomic resources present an excellent opportunity for goat genetic improvement. For instance, the unimproved ecotypes offer unique adaptive traits that can be utilised to improve the broader goat population (Visser and van Marle-Koster, 2018; Erasmus, 2000; Malan, 2000). However, they also pose a danger if the selection strategies used are not well formulated and implemented (Visser and van Marle-Koster, 2018).

Therefore, genomic studies are needed to incorporate genetic information into the design and implementation of improved selection programmes. In goat populations from other countries, such studies have been implemented to identify genes/loci that influence various economically important traits including body morphology (Rahmatalla et al., 2018), reproduction traits (Islam et al., 2020), lactation yield of milk, fat, protein, and somatic cell score (Scholtens et al., 2020). This study aims to use selection signatures, genome-wide association studies, and copy number variations to identify genes/loci that have significant effects on coat colour patterns in the meat goat populations of South Africa.

1.1 Justification

Goats found in South African communal production systems present great potential for genetic improvement as they display variation in production traits, and adaptation to harsh environmental conditions. However, these population are faced with erosion due to their indiscriminate crossbreeding and replacement with highly productive commercial breeds (Mdladla et al., 2016). This endangers the livelihood of goat farmers in communal production systems in which these goats are an important resource for their economy. Due to recent advances in molecular genetics and the reduced costs of SNP genotyping using commercial genotyping platforms that cover thousands of SNPs across animal genomes have allowed the genotyping of several livestock species at low prices and the use of SNP data to estimate breeding values for genomic selection (Dekkers, 2011; VanRaden et al., 2009; König and Swalve, 2009; Schaeffer, 2006).

The investigation of selection signatures, genome-wide association, and copy number variations analysis is a key interest in animal genetics as these methods can be used to investigate trait-linked genes and SNPs responsible for livestock adaptation to different environments (Zhao et al., 2015). For instance, information from such studies can reveal the genetic consequences of adaptive evolution, domestication, and selection that have resulted from the development of diverse livestock breeds adapted to different environments and production systems (Otto, 2000). In addition, such information will be vital in the implementation of genomic selection which will enable the use of genome-wide associations for economically important traits, to inform selection decisions made in breeding programs for the genetic improvement of livestock species and their conservational programmes (Saravanan et al., 2020).

Coat colour has been demonstrated as an indicator trait for livestock adaptation to different climates (Onasanya et al., 2018). This is because goats with different coat colours possess different thermoregulatory abilities, such as heat tolerance (Joy et al., 2020) with studies reporting an increased adaptation to cold weather conditions in black goats, compared to white goats (Ferreira et al., 2021; Chokoe et al., 2020). Such differences in goat adaptation due to coat colour has led to farmers in communal programmes selecting reproducers based on their apparencies. In this regard, the investigation of coat colour variation within the South African meat-type populations offers an opportunity to unravel the genetic mechanisms responsible for indigenous goats' adaptation to adverse climatic and environmental conditions, which will

facilitate efforts to implement improved breeding schemes in the low scaled production systems of these populations. Furthermore, coat colour has also been linked to other economical traits important for goat productivity such as reproductive characteristics (Anzures-Olvera et al., 2019; Choudhury et al., 2013; Berhanu et al., 2012), milk and meat production (Mia and Mondal., 2020; Archana et al., 2018), and growth traits (Hossain et al., 2020; Choudhury et al., 2013; Mabrouk et al., 2010; Daramola and Adeloye et al., 2009).

1.2 Aims and objectives

Aim

The overall aim of this study was to use selection signatures, genome-wide association studies, and copy number variations to identify genes/loci that have significant effects on coat colour and coat colour patterns in indigenous meat-type goat populations of South Africa.

Objectives

- To use SNP marker information to study the genetic population structure of South African meat-type goats and infer its relationship with the diversity of coat colour and coat colour patterns.
- To use Genome-wide association studies and copy number variation analysis to investigate genomic regions associated with coat colours and coat colour patterns in South African indigenous meat-type breeds.
- To detect signatures of selection for genes under positive selection for coat colour and coat colour patterns in South African indigenous meat-type goats.

1.3 Outline of thesis

The dissertation begins with a general introduction and literature covering the subject of goat genetic improvement in South Africa and the application of genomic tools to provide insights into gene-trait relationships for the implementation of genomic selection in livestock breeding. The chapters are organized as follows:

Chapter 1: General introduction, and aims and objectives.

Chapter 2: Literature review.

Chapter 3: The population structure of South African indigenous meat-type goats and inferences on its relationship with the variation of coat colour and coat colour patterns.

Chapter 4: Genome-wide association studies and copy-number variation analysis reveal genes for coat colour and coat colour patterns in South African indigenous meat-types goats.

Chapter 5: Genome-wide detection of signatures of selection for coat colour in South African indigenous meat-type goats with various coat colours and patterns.

Chapter 6: General discussion and conclusion, as well as study limitations and future recommendations.

1.4 References

ANZURES-OLVERA, F., VÉLIZ, F. G., DE SANTIAGO, A., GARCÍA, J. E., MELLADO, J., MACÍAS-CRUZ, U., AVENDAÑO-REYES, L. & MELLADO, M. 2019. The impact of hair coat color on physiological variables, reproductive performance and milk yield of Holstein cows in a hot environment. *Journal of Thermal Biology,* 81**,** 82-88.

ARCHANA, P. R., SEJIAN, V., RUBAN, W., BAGATH, M., KRISHNAN, G., ALEENA, J., MANJUNATHAREDDY, G. B., BEENA, V. & BHATTA, R. 2018. Comparative assessment of heat stress induced changes in carcass traits, plasma leptin profile and skeletal muscle myostatin and HSP70 gene expression patterns between indigenous Osmanabadi and Salem Black goat breeds. *Meat Sci,* 141**,** 66-80.

BARSH, G. S. 1996. The genetics of pigmentation: from fancy genes to complex traits. *Trends Genet,* 12**,** 299-305.

BERHANU, T., THIENGTHAM, J., TUDSRI, S., ABEBE, G., TERA, A. & PRASANPANICH, S. 2012. Purposes of keeping goats, breed preferences and selection criteria in pastoral and agro-pastoral districts of South Omo Zone. *Livestock Research for Rural Development,* 24.

BOSMAN, L., VAN MARLE-KOSTER, E. & VISSER, C. 2015a. Population structure of South African commercial dairy goats. *European Regional Conference on Goats 2014*.

BOSMAN, L., VAN MARLE-KÖSTER, E. & VISSER, C. 2015b. Genetic diversity of South African dairy goats for genetic management and improvement. *Small Ruminant Research,* 123**,** 224-231.

CAMPBELL, Q. 2003. The origin and description of southern Africa's indigenous goats. *S Afr J Anim Sci,* 4**,** 18-22.

CHOKOE, T., MATELELE, T., MAQHASHU, A., FHULUFHELO, R., TUMUDI, M., MPOFU, T., NEPHAWE, K. & MTILENI, B. 2020a. Phenotypic Diversity of South African Indigenous Goat Population in Selected Rural Areas. *American Journal of Animal and Veterinary Sciences,* 15**,** 59-66.

CHOKOE, T. C., MDLADLA-HADEBE, K., MUCHADEYI, F., DZOMBA, E., MATELELE, T., MPHAHLELE, T., MPOFU, T. J., NEPHAWE, K. & MTILENI, B. 2020b. Genetic Diversity of South African Indigenous Goat Population from Four Provinces Using Genome-Wide SNP Data. *Sustainability,* 12**,** 10361.

CHOUDHURY, M., SARKER, S. C., ISLAM, F., ALI, A., BHUIYAN, A. K. F. H., IBRAHIM, M. N. M. & OKEYO, A. 2013. Morphometry and performance of Black Bengal goats at the rural community level in Bangladesh. *Bangladesh Journal of Animal Science,* 41.

DARAMOLA, J. O. & ADELOYE, A. A. 2009. Physiological adaptation to the humid tropics with special reference to the West African Dwarf (WAD) goat. *Trop Anim Health Prod,* 41**,** 1005-16.

DEKKERS, J. C. 2011. Use of high-density marker genotyping for genetic improvement of livestock by genomic selection. *Animal Science Reviews 2010***,** 197.

ERASMUS, J. 2000. Adaptation to various environments and resistance to disease of the Improved Boer goat. *Small Ruminant Research,* 36**,** 179-187.

FERREIRA, J., SILVEIRA, R. M. F., DE SOUSA, J. E. R., DE VASCONCELOS, A. M., GUILHERMINO, M. M. & FAÇANHA, D. A. E. 2021. Evaluation of homeothermy, acid-base and electrolytic balance of black goats and ewes in an equatorial semi-arid environment. *Journal of Thermal Biology,* 100**,** 103027.

FONTANESI, L., BERETTI, F., RIGGIO, V., DALL'OLIO, S., GONZÁLEZ, E. G., FINOCCHIARO, R., DAVOLI, R., RUSSO, V. & PORTOLANO, B. 2009. Missense and nonsense mutations in melanocortin 1 receptor (MC1R) gene of different goat breeds: association with red and black coat colour phenotypes but with unexpected evidences. *BMC Genetics,* 10**,** 47.

HOSSAIN, S., AFROZ, F., HABIB, M., MIRAZ, M. F. & JALIL, M. 2020. Genetic and phenotypic performance of three coat color variants of Black Bengal Goat in a closed nucleus flock at BLRI.

ISLAM, R., LIU, X., GEBRESELASSIE, G., ABIED, A., MA, Q. & MA, Y. 2020. Genome-wide association analysis reveals the genetic locus for high reproduction trait in Chinese Arbas Cashmere goat. *Genes & Genomics,* 42.

JOY, A., SEJIAN, V., KRISHNAN, G., MADIAJAGAN, B., PRATHAP, P. & BHATTA, R. 2020. Heat stress impact on blood biochemical response and plasma aldosterone level in three different indigenous goat breeds. *Journal of Animal Behaviour and Biometeorology,* 8**,** 266-275.

KÖNIG, S. & SWALVE, H. 2009. Application of selection index calculations to determine selection strategies in genomic breeding programs. *Journal of Dairy Science,* 92**,** 5292-5303.

MABROUK, O., SGHAIER, N., COSTA, R., GADDOUR, A., BENAMMAR ELGAAÏED, A. & DELGADO, J. V. 2010. The effect of non-genetic factors on the early body weights of Tunisian local goats. *Revista Brasileira De Zoologia - REV BRAS ZOOL,* 39.

MALAN, S. W. 2000. The improved Boer goat. *Small Rumin Res,* 36**,** 165-170.

MARTIN, P. M., PALHIÈRE, I., RICARD, A., TOSSER-KLOPP, G. & RUPP, R. 2016. Genome Wide Association Study Identifies New Loci Associated with Undesired Coat Color Phenotypes in Saanen Goats. *PloS one,* 11**,** e0152426-e0152426.

MDLADLA, K., DZOMBA, E. & MUCHADEYI, F. 2016a. P5039 A landscape genomic approach to unravel the genomic mechanism of adaptation in indigenous goats of South Africa. *Journal of Animal Science,* 94**,** 134-135.

MDLADLA, K., DZOMBA, E. F., HUSON, H. J. & MUCHADEYI, F. C. 2016b. Population genomic structure and linkage disequilibrium analysis of South African goat breeds using genome-wide SNP data. *Anim Genet,* 47**,** 471-82.

MDLADLA, K., DZOMBA, E. F. & MUCHADEYI, F. C. 2017. Characterization of the village goat production systems in the rural communities of the Eastern Cape, KwaZulu-Natal, Limpopo and North West Provinces of South Africa. *Trop Anim Health Prod,* 49**,** 515-527.

MIA, M. & MONDAL, A. 2020. A STUDY ON PRODUCTIVE AND REPRODUCTIVE PERFORMANCES OF BLACK BENGAL GOAT IN RELATION TO COAT COLORS.

MONAU, P. I., VISSER, C., MUCHADEYI, F. C., OKPEKU, M., NSOSO, S. J. & VAN MARLE-KÖSTER, E. 2020. Population structure of indigenous southern African goats based on the Illumina Goat50K SNP panel. *Trop Anim Health Prod,* 52**,** 1795-1802.

NCUBE, K. T., HADEBE, K., DZOMBA, E. F., SOMA, P., FRYLINCK, L. & MUCHADEYI, F. C. 2020. Relationship between population genomic structure and growth profiles of South African goats under different production systems. *Trop Anim Health Prod,* 52**,** 1277-1286.

ONASANYA, G., DE CAMPOS, J., ANWO, O., IKEOBI, C. & AMUSAN, S. 2018. *Impact of Coat Colour Genes and Coat types on zoometric Traits of African Dwarf Sheep bred under Traditional Extensive System*.

OTTO, S. P. 2000. Detecting the form of selection from DNA sequence data. *Trends Genet,* 16**,** 526-9.

PETERS, K. J. 1980. The importance of coat colour and coat type for productive performance of beef cattle under subtropical conditions. *Annales de Génétique et de Sélection Animale,* 12**,** 131-132.

PIETERS, A., MARLE-KOSTER, V., VISSER, C. & KOTZE, A. 2009. South African developed meat type goats: A forgotten animal genetic resource?

Chapter 2: Literature Review

Abstract

Goats in South Africa are a major source of income and food security for many goat farmers. The goats are bred for various goals, including mohair, milk, and meat production. The success of goats as livestock results from their genetic diversity, which allows them to thrive in different environmental conditions. Goat genetic diversity has been studied extensively using DNA markers and, more recently, with high-density single nucleotide polymorphism (SNP) arrays. However, despite these studies, goat genetic improvement is lagging compared to other similar livestock species such as cows and sheep. This is because there is still a lack of knowledge about the genetic mechanisms that underlie goat production traits, such as coat colour. Coat colour in goats forms a part of breed definition and characterisation. Coat colour has also been reported to play a role in heat tolerance for goats. Furthermore, the trait is of high economic importance as it can determine the selling price of the goats. Knowledge about the genes that influence the trait is therefore important. This study will use SNP array technologies (i.e., Signatures of Selection, GWASs, and Copy number variation) to identify genes/loci that have effects on coat colour patterns in the meat goat populations of South Africa. As such this review discusses the history of goat breeding in South Africa as well as its current landscape in the light of recent advances in Molecular biology which allows the potential implementation of genomic selection in goat production systems. The study of coat colour will provide information that can be used to design and implement improved breeding and conservation programmes for SA goats. This will increase the genetic improvement of SA meat goats and aid in conserving the genetic diversity of South Africa's unimproved village goats, which is vital as these goats have been reported to have unique adaptive characteristics, such as increased resistance against tick-borne diseases compared to the commercial breeds.

Keywords: SNP array, coat colour, signatures of Selection, GWASs, copy number variation

2.1 Introduction

Goats are a big part of the livestock industry in South Africa, with Southern Africa producing a large percentage of Africa's goat population (Visser and van Marle-Koster, 2018). The goats contribute towards the economic and food security of smallholder farmers and communal farmers (Visser and van Marle-Koster, 2018). This is because goats have proven to be resilient livestock, capable of surviving in a wide range of environments. For instance, goat populations found in the extensive and intensive production systems of South Africa, which are the two main goat production systems in the country, are exposed to vastly different selection pressures, environments, and breeding goals (Mohlatlole et al., 2015). Despite these differences, the goat populations thrive. The commercial breeds in the extensive production system dominate commercial farms, while ecotypes in the intensive system dominate small scale production systems (Mdladla et al., 2017).

The resilience of goats is attributed to their genetic diversity. Goat genetic diversity has been vastly studied using DNA markers, with initial studies using microsatellites. The first of these studies was by Visser et al. (2004). Despite the successes of these early studies in identifying genetic diversity, the effects of genetic diversity on the phenotypic variation of economically essential traits were still not well understood (Visser and van Marle-Koster, 2018). However, with the recent rise of high-throughput sequencing technologies and the drop in sequencing costs, this has started to change.

The sequencing of the goat genome with next-generation sequencing technologies and the development of molecular tools which enable the study of the variation of common single nucleotide polymorphisms (SNPs) now catalyses a high-throughput characterisation of goat genetic variation (Visser and van Marle-Koster, 2018). This characterisation will provide information about the genes and genetic mechanisms that underlie goat production traits leading to better goat breeding and conservation programmes similar to those of other important livestock species such as cattle and sheep (Visser and van Marle-Koster, 2018). This is a review on the history of goat breeding and improvement in South Africa and the role of genomic tools in livestock production, in particular the recent availability of high-density SNP genotyping panels.

2.2 Goat domestication and breeding history

Goat (Capra hircus) domestication which occurred around 10,000 years ago has resulted in various goat populations adapted to diverse environments due to differences in their

morphological traits which also impact their productivity (Zeder and Hesse, 2000). This variation was driven by both genetic and environmental factors, including intentional selection for specific traits, genetic drift, isolation, and founder effects (Zeder and Hesse, 2000). The early selective breeding of goats, dates back 7000 years ago to Ancient Egypt, where goats were bred for morphological traits such as twisted horns and long lop ears (Porter et al., 2016). Along with horn and ear shapes, several other traits including the presence of wattles, long hair, and coat colours were developed through domestication (Porter et al., 2016). Following their initial domestication and success as livestock, goats were dispersed around the world through several migration events (Porter et al., 2016).

The dispersal of goats led to a population expansion in other continents outside of their original place of domestication. Nowadays, most of the world's goat population is found in Asia (58.2%) and Africa (36.2%) (Amills et al., 2017). Throughout the ages, goats have been raised to meet many needs, including milk, meat, and fibre production. In recent years specialised breeds with high productivity for these production traits have been developed and spread worldwide (Ajmone-Marsan et al, 2014). Currently, Asia has the largest populations of specialised breeds that produce fresh milk, meat, and fibre globally, with Africa coming in second (Amills et al., 2017). Although the world's goat population has significantly grown over the years, there has been a lack of improved goat production, reproduction, and management programmes (Dubeuf & Boyazoglu, 2009).

Until recently, improved selection schemes have been reserved to only a few commercial breeds, most of which are found in Europe, North America, and Australia despite the low goat census of these countries (Dubeuf & Boyazoglu, 2009). Selection programmes in these countries have had a significant impact on goat production and improvement. For instance, selection programmes implemented in France have resulted in an annual increase of milk per lactation, and protein and fat contents in their French Alpine and Saanen goats (Amills et al., 2017). While in regions without selection programmes, where breeding animals are not chosen based on evaluations of their genetic potential and phenotype recordings and identification systems are rarely available, goat improvement is lagging (Aziz, 2010).

2.3 South African goat breeds

There are seven prominent goat breeds in South Africa. These include the Angora breed which produces mohair, three breeds which produce milk (British Alpine, Toggenburg, and Saanen), and another three breeds that are used in the production of meat (Boer, Savanna, and Kalahari

Red) (Visser and van Marle-Koster, 2018). In addition, there are many unimproved indigenous goats that are reared by smallholders and subsistence farmers across the country for various reasons such as the production of hides, milk, and meat (Visser and van Marle-Koster, 2018). These indigenous veld populations make up around 63% of the goat population in Southern Africa (Mahanjana and Cronje, 2000).

The Angora goat was domesticated in Turkey (Visser et al., 2011b). According to historical evidence, the goats were imported into the South Africa during the nineth century (1838) following an unsuccessful attempt to establish a mohair industry in Europe. South Africa proved to be a better-suited region for mohair production due to its better climate compared to Europe (Visser et al., 2011b). In 2018, the SA mohair industry was approximated at 644,000 Angora goats and it is accounts for over 50% of the mohair produced globally (Visser and van Marle-Koster, 2018) meaning the goat sector is a significant part of the economy in South Africa (Visser and van Marle-Koster, 2018).

The Boer, Kalahar Red, and the Savanna breeds are said to be indigenous to South Africa and were developed from unimproved veld goat reared by people of the Khoisan tribe approximately 500 AD (Maree and Plug, 1993). In this sense, the SA veld goats were vital in the genetic improvement and development of the Boer, Kalahari Red, and the Savanna breeds. In 2018, South Africa's commercial meat goat industry was reported to have approximately 1.3 million goats (Visser and van Marle-Koster, 2018). In contrast, South Africa's prominent dairy goats which include the British Alpine, Toggenburg, and Saanen breeds were imported into the country during the 1990s from the United Kingdom and Switzerland (Visser and van Marle-Koster, 2018; Muller, 2005). The commercial dairy goat industry has around four thousand dairy goats, making it a smaller sector compared to the mohair and meat sectors (Visser and van Marle-Koster, 2018).

2.3.1 The genetic characterisation of South African goat breeds

There are many goat breeds; however, the exact number of these breeds is unknown as several local populations have gone uncharacterised. A contributing factor to this lack of goat characterisation is the absence of proper management and conservation strategies for local populations. These populations are usually not managed via phenotypic standardisation, herd book registration and controlled reproduction and thus, no characterisation data on them are available (Porter et al., 2016; Taberlet et al., 2008).

Due to the lack of genetic conservation programmes for these populations, together with their replacement and uncontrolled crossbreeding with more productive foreign varieties, these local goat populations tend to suffer a decline in their genetic resources (Taberlet et al., 2011). For instance, South African veld goats have been used in crossbreeding programs aimed at improving local commercial breeds such as the Boer, Kalahari Red, and the Savanna. However, these goats have not undergone genetic improvement, and selection programs within the populations have been limited (Campbell, 2003).

Characterisation studies on the Angora breed have revealed that the goats have small bodies covered by white mohair, with drooping ears and horns (Visser and van Marle-Koster, 2018; Campbell, 2003). The Angora Goat Breeders' Society, which has been active since 1892, describes the breed standards for the Angora goat as including a white fleece free from kemp or coloured fibres (Snyman, 2014). In contrast, breed standards of the South African Boar describe a redhead with long ears, and a soft white coat (Visser and van Marle-Koster, 2018; Campbell, 2003). In addition, the breed has a sturdy head with a compressed nose and backward curved horns (Pieters et al., 2009; Malan, 2000; Sambraus, 1992). Furthermore, the breed has demonstrated a higher disease resistance to prussic acid poisoning, bluetongue, and enterotoxaemia (Erasmus, 2000; Malan, 2000). Breed standards for the Kalahari Red describe a dark red coat which provides the goats with a higher tolerance to UV radiation (Pieters et al., 2009). While the Savanna breed has a white coat (Pieters et al., 2009).

In contrast to the commercial breeds, indigenous veld goat populations have only been characterised on the basis of their diverse phenotypes such as coat colour, horn shape, and ear length (Visser and van Marle-Koster, 2018; Campbell, 2003). As such distinct breed IDs do not exist, and populations are differentiated based on the geographical regions in which they are found. For instance, populations found in KwaZulu-Natal, Northern and Eastern Cape are called the Nguni goats, the Northern Cape Speckled goats, and the Xhosa Lop ear goats, respectively (Mohlatlole et al., 2015). These veld populations differ greatly in size but generally have smaller bodies and reduced carcass yields compared to the commercial breeds (Visser and van Marle-Koster, 2018; Campbell, 2003). In addition, the goats a better adapted to diverse environments compared to the commercial populations. For instance, veld goats have been shown to have a higher disease tolerance (Erasmus, 2000; Malan, 2000).

The South African Milch Goat Breeders' Society, which manages the breeding standards for dairy goats, was formed in 1958 (Olivier et al., 2005). South Africa first imported the Saanen

breed in 1898 (Olivier et al., 2005). Today the breed is characterised by its large yields of milk (Olivier et al., 2005). Similarly, the British Alpine and Toggenburg breeds, were imported during the twentieth century, and are known for their dark pigmentation which allowed them to adapt to the diverse climatic conditions of South Africa which include extremely hot temperatures and high UV radiation (Visser and van Marle-Koster, 2018). The Toggenburg breed is also known for producing higher butterfat milk than the Saanen (Olivier et al., 2005).

2.3.2 The genetic diversity and improvement of South African goat populations

Genetic diversity is the genetic variation present in a population, and it has a direct influence on genetic progress, selection strategies and the control of inbreeding levels (Sponenberg, 2020). Genetic diversity also plays a significant role in genetic characterisation, which involves explaining the genetic variation of important livestock production traits by identifying quantitative trait loci (QTL) that underlie the traits (Sponenberg, 2020). Therefore, information about genetic diversity can inform genetic characterisation, which can help preserve the genetic resources of indigenous populations.

Information about genetic diversity can increase the accuracy of estimated breeding values (EBVs) resulting in a faster rate of genetic improvement (Sponenberg, 2020). A comprehensive and thorough description of the genetic diversity of goat breeds assessed with nuclear and mitochondrial markers has been reported (Ajmone-Marsan et al., 2014). Furthermore, several genetic diversity studies on South African breeds have been carried out (Chokoe et al., 2020; Ncube et al., 2020; Mdladla et al., 2016a; Mdladla et al., 2016b; Huson et al., 2014; Pieters et al., 2009; Visser et al., 2004; Kotze et al., 2004). However, research on investigating the relationships between genotypes and phenotypes in goats is still lacking compared to similar livestock species such as sheep and cattle, and this has resulted in slow genetic improvement in goats (Amills, 2014).

Most of the research on genotype-phenotype relationships in SA goat populations has been mainly focused on the Angora due to the high economic value of mohair (Visser and van Marle-Koster, 2018). As a result, the Angora goat has undergone the most genetic improvement compared to the rest of the country's goat populations. Initial studies of the Angora targeted specific QTL to detect genomic regions linked to traits that impact the production and quality of mohair (Visser et al., 2011a) and preweaning growth (Visser et al., 2013). The studies successfully identified Trait-associated segments; however, the identified fragments only

partially explained the phenotypic variation of the traits, and this limited marker-assisted selection (MAS).

The development of the 50K SNP chip, which covers over 52 thousand SNPs across the goat genome (Tosser-Klopp et al., 2016) allowed researchers to overcome the initial limitations of MAS. Studies that have used the Bead-chip to investigate the genetic variation of the Angora breed have shown high levels of genetic diversity within the population to allow for successful selection strategies and genetic improvement (Lashmar et al., 2016). One of these studies also found that the SA Angora goat is genetically distinct from Angora goats found in France and Argentina (Visser et al., 2016). That study concluded that the distinct clustering of the South African population was likely the result of strict selection in mohair production systems (Visser et al., 2016).

The first study investigating the genetic variation of SA meat goats used microsatellite markers and found that the commercial breeds were distinct from each other. In contrast, the indigenous goats were found to have high levels of genetic relatedness (Visser et al., 2004). When further investigated using the Illumina Goat SNP50K genotyping array, it was found that the distinctiveness of the commercial breeds was due to local selection and adaptation (Pieters et al., 2009; Kotze et al., 2004; Hefer et al., 2004), and that the indigenous populations had high levels of genetic diversity and low levels of inbreeding despite also showing some genetic relatedness (Mdladla et al., 2016b; Huson et al. 2014). These studies concluded that the low-inbreeding levels were likely due to the absence of strict selection in the communal production systems the goats are kept (Mdladla et al., 2016b; Huson et al. 2014).

In dairy goats, genetic diversity has previously been studied using microsatellite markers which have revealed high degrees of genetic diversity (Bosman et al., 2015b). In addition, the populations were shown to have low inbreeding levels, and low genetic differentiation between the breeds (Bosman et al., 2015b). This was expected within one production type (Bosman et al., 2015b). However, an admixture group of animals was revealed, indicating that purebred animals had undergone crossbreeding at some point (Bosman et al., 2015b). These findings were later supported by results from the study by Lashmar et al. (2016), which used SNP data to reveal high gene diversity within the breeds and levels of co-ancestry between the breeds.

Although the meat and the dairy goats have been included in genetic diversity studies, several challenges, such as the high costs of using modern technology for measuring traits of economic importance, have hindered their genetic improvement by limiting phenotypic and pedigree

recordings for these goat breeds (Visser and van Marle-Koster, 2018). Overcoming these challenges will enable the evaluation of genetic parameters for traits of economic importance through the estimation of genetic breeding values, which will lead to a faster genetic improvement for the goats.

2.4 Genomic tools in livestock production

Literature shows that prior to the introduction of molecular genetics into livestock breeding programs, selection programs were based on EBVs estimated from phenotype measurements. While this method of EBV estimation was successful, it suffered several limitations related to difficulties in the phenotyping of selection candidates (i.e., traits like longevity can only be recorded late in life, while others like meat quality, require animals to be sacrificed) and the high costs of phenotype recording (Dekkers, 2011). These phenotyping constraints limited the amount of genetic progress that was being made.

The introduction of molecular genetic tools provided solutions to these limitations by enabling the use of DNA markers such as SNPs in the detection of genomic regions which encode genes that influence economically important traits (Dekkers, 2011). These strategies were initially used to investigate the genetic basis of single gene defects and create methods to identify these defects (Dekkers, 2011). For quantitative traits, these tools provided the ability to carry out the marker-assisted selection of candidates at an early age based on QTL information. Marker-assisted selection (MAS) uses information generated from genetic markers associated with QTLs and phenotypic information to inform selection (Smith and Simpson, 2010; Lande and Thompson, 1990; Soller, 1978).

Since its introduction into animal breeding, molecular genetics has facilitated the discovery of several QTLs and genotype-phenotype relationships as well as some causative mutations (Dekkers, 2004; Dekkers and Hospital, 2002). However, until recently, the application of this information in breeding programs, was limited primarily due to the high costs of routinely genotyping selection candidates and because most of the QTL studies tended to explain only a limited amount of genetic variation of the traits of interest (Dekkers, 2004). These limitations were overcome with the recent drop in genotyping costs and the introduction of SNP genotyping arrays.

2.5 The application of SNP genotyping arrays in livestock production

SNP genotyping arrays have become a popular genomic tool in livestock production. In recent years commercial genotyping platforms that cover thousands of SNPs across animal genomes have become available, thus allowing the genotyping of several livestock species at low prices (Dekkers, 2011). For instance, the high-density 50K Bovine Illumina SNP genotyping panel (Matukumalli et al., 2009), has allowed the genotyping of many dairy and beef cattle (Dekkers, 2011). In addition, a 70k Bovine SNP panel has recently been developed. Similarly, in goats, sheep, poultry, and pigs, high-density 40 and 65K SNP panels have been developed (Dekkers, 2011) and higher density panels are also currently under development.

The availability of high-density SNP genotyping panels like these has allowed the estimation of accurate EBVs (estimated breeding values) by providing information about the effects of thousands of SNPs across the genome, enabling the implementation of genomic selection (Dekkers, 2011; VanRaden et al., 2008; Hayes et al., 2009; Meuwissen et al., 2001). Therefore, in contrast to phenotypic prediction methods of breeding values, in genomic selection, the breeding value of potential reproducers is estimated according to both their phenotypes and their genotypes (Coster et al., 2010; Verbyla et al., 2010).

The use of SNP data generated from high-density SNP genotyping panels to estimate breeding values for genomic selection has been carried out in several livestock species. For instance, genomic selection in dairy cattle has allowed the selection of young bulls for breeding through the accurate estimation of breeding values at a young age (VanRaden et al., 2009) without the need for progeny data. The use of genomic selection in dairy cattle breeding programs has resulted in reduced generation intervals and reduced costs because the method doesn't require progeny testing (König and Swalve, 2009). In addition, genomic selection provides an improved selection method for traits such as longevity and fertility which show low heritability (Schaeffer, 2006). As such, the implementation of genomic selection in other livestock species is under development (Dekkers, 2011).

2.5.1 Signatures of selection

Selection signatures have been described as unique genetic changes in the genome left behind by selection whether natural or artificial (Jensen et al., 2016; Nielsen, 2005). These changes often occur in genomic regions linked to important production traits such as body morphology and conformation, reproductive performance, adaptation, and disease resistance (Saravanan et al., 2020; Jensen et al., 2016; Nielsen, 2005). These genetic changes are responsible for

livestock adaptation to adverse environmental conditions (Groeneveld et al., 2010). The process in which selection in livestock produces signatures is called a selective sweep (Smith and Haigh, 2007; Fay and Wu, 2000; Braverman et al., 1995). Selective sweeps occur when the increase of the frequency of a previously rare allele cause a reduction of variability up and downstream that allele (Saravanan et al., 2020; Maynard and Haigh, 2007; Fay and Wu, 2000; Braverman et al., 1995).

The increase in frequency can occur rapidly, in which case, the selected allele becomes fixed in the entire population upon selection and causes a reduction in the population's genetic diversity (Pritchard et al., 2010). In this sense, hard selective sweeps are recognizable by reduced genetic diversity in the population. In addition, hard selective sweeps increase linkage disequilibrium, and homozygosity (Pritchard et al., 2010). On the other hand, soft sweeps do not cause reductions in genetic diversity (Pritchard et al., 2010). However, the signatures of selection of the soft sweeps are harder to detect because they tend to be less pronounced (Pritchard et al., 2010). In addition, depending on the type, frequency, and origin of the selected allele, selective sweeps can be partial or complete. Sweeps are considered partial when the selected alleles do not become fixed, rather they segregate along with neutral sites in the population. In contrast, when the alleles become fixed it is considered a complete sweep (Hermisson and Pennings, 2017).

The identification of selection signatures is a key interest in animal genetics as it can be used to investigate trait-linked genes and SNPs responsible for livestock adaptation to different environments (Zhao et al., 2015). Furthermore, information from a signatures of selection analysis can reveal the genetic consequences of adaptive evolution, domestication, and selection that have resulted from the development of diverse livestock breeds adapted to different environments and production systems (Otto, 2000). In addition, such information will be vital in the implementation of genomic selection which will enable the use of genome-wide associations for economically important traits, to inform selection decisions made in breeding programs for the genetic improvement of livestock species and their conservational programmes (Saravanan et al., 2020).

2.5.1.1 Methods and tools used for the detection of selection signatures

The availability of high-density SNP arrays, together with next-generation sequencing (NGS) technologies have enabled whole genome scanning for selection signatures using genetic markers such as SNPs found in genomic regions of interest and under selection among several livestock species. In this regard, several approaches that use DNA markers to detect signatures of selection have been developed to aid the implementation of genomic selection within livestock breeding programs (Qanbari and Simianer, 2014; Oleksyk et al., 2010).

The currently available methods are based on detecting signatures of selection within-population or between-populations (Saravanan et al., 2020). Methods that detect signatures of selection within-populations can be further classified into three groups, which include methods that detect signatures based on reduced local variability, linkage disequilibrium and site frequency spectrum (Weigand and Leese, 2018). In contrast, methods that detect signals of selection between-populations are either based on single site or haplotype differentiation of locus-specific allele frequencies (Zhao et al., 2015). The detection of selection signatures requires genotype data, generated using either SNP data from SNP genotyping panels or DNA sequence data from high-throughput sequencing technologies. However, the data usually requires pre-processing before it can be used in the statistical analysis as current methods and tools of detecting selection signatures, each require unique input formats (Cadzow et al., 2014). To account for the diverse input and output formats required, to run the analysis, several user-friendly software tools can be used to convert data between the different formats. Examples of these programs include PGDSpider (Lischer and Excoffier, 2012) and fcGENE (Roshyara and Scholz, 2014).

For SNP data quality assurance, which is required before analysis, open-source software tools like PLINK, SNPQC, and R are normally utilised. PLINK (Purcell et al., 2007) and R statistical software packages (R Core Team, 2023) are the most popular for SNP data manipulation, quality assurance, statistical analysis, and graphical visualization due to their speed and reliability. Alternatively, licensed software tools such as JMP Genomics (SAS Institute Inc., Cary, NC, USA) and SNP Variation Suite Golden Helix (SVS, Golden Helix, Bozeman, MT, USA) can also be used. After data quality assurance steps are completed, selection signature detection methods which generate integrated haplotype scores (iHS) based on Linkage Disequilibrium (LD) and those which detect cross-population extended haplotype homozygosity (XP-EHH) based on haplotype differentiation require further data pre-

processing to phase haplotype blocks and account for missing genotypes (Saravanan et al., 2020). In this regard, haplotype phasing and imputation can be performed using several software tools such as fastPHASE, IMPUTE2, and Beagle (Nicolazzi et al., 2015). For SNP data imputation and haplotype phasing in livestock genomics, Beagle (Browning et al., 2018) and IMPUTE2 (Howie et al., 2009) are the most popular (Nicolazzi et al., 2015).

There are several different software programs available for the detection of selection signatures (Saravanan et al., 2020). While the available software tools use different approaches, majority of them can be run on computers from the command line prompt. However, to use the software tools to run the analysis some level of experience with coding is required. Therefore, the decision of which tools to use for data analysis depends on the preferences and experience of the user, data input and output requirements, the operating system of the computer being used as well as hardware limitations such as computing speed. Most of the available software tools work in the Windows, Linux and macOS operating systems (Saravanan et al., 2020).

2.5.1.2 Applications for selection signatures in livestock production

In recent years, there have been numerous studies that have investigated signatures of selection within livestock breeding. Such studies have succeeded in explaining the cause of differentiation between numerous livestock breeds, compared to studies of genetic diversity which have only been able to quantity levels of genetic differentiation. Therefore, understanding signatures of selection is vital for livestock characterisation as it offers a way to identify genes linked to livestock production traits (Cesarani et al., 2018).

Through investigating the causal genes of economically important traits, the study of selection signatures helps generate information about the beneficial mutations responsible for livestock adaptation to diverse environments (Yurchenko et al., 2018). This information presents an opportunity to better understand the historical development of different livestock breeds as well as the genetic mechanisms underlying phenotypic differentiation. The study of signatures of selection can also be used to explain the genetic consequences of artificial selection implemented for the genetic improvement of populations. Knowing the targets of selection and understanding the responses to selection of different livestock populations is necessary for the design and implementation of improved breeding programs aimed at improving production traits (Gurgul et al., 2018).

Studying signatures of selection can be useful in genomic selection, as candidate genes detected to be under selection by the analysis can further be associated with phenotypes of interest using

Genome-wide association studies (GWAS) (Chen et al., 2016). During a signatures of selection analysis, SNP data is statistically analysed using population genetic parameters to pinpoint genomic regions linked to candidate genes under selection (Bomba et al., 2015). In this sense, the analysis is able to link the genotypes of animals to their phenotypes without requiring phenotypes to be measured (Zhao et al., 2015). This contrasts with genome-wide association studies which use SNP data to scan the genome for genetic variants linked to measured variation in the phenotypes (Tam et al., 2019). Therefore, when dealing with traits that are difficult or costly to measure such as disease resistance, and tolerance or adaptation to diverse climates, the analysis of signatures of selection can be a better suited alternative to GWAS (Saravanan et al., 2020). In addition, the study of signatures of selection offers breeders the opportunity to assess the progress of their breeding programs by providing information about the animal responses to different selection methods. As a result, the signatures of selection analysis are sometimes preferred over GWAS to identify SNPs linked to genes that influence the phenotypic variation in traits of interest (Maiorano et al., 2018).

2.5.1.3 Studies of selection signatures in goat populations

After the successful application of the signatures of selection analysis in humans, similar studies were carried out in cattle, and eventually other livestock species such as goats, sheep, and horses (Onzima et al., 2018; Brito et al., 2017; Manunza et al., 2016; Fariello et al., 2014). Examples of such studies in goats include the study by Brito et al. (2017) who investigated signals of selection in diverse goat breeds using different statistically analysis including ROH, FST, and hapFLK. In their study, Brito et al. (2017) revealed selection signals in genomic regions overlapping genes involved with various trait of economic importance such as reproduction and morphology (Brito et al., 2017).

In another study, Onzima et al. (2018) used genotype data generated by the Goat 50K SNP chip to study selection signals in several goat breeds in Uganda. Candidate genes under selection within the different Ugandan breeds were found to be involved with various goat production traits such as thermotolerance and immune response which are vital for goat adaptation to diverse environments and climates (Onzima et al., 2018). In addition, numerous other studies have been carried out in small ruminants such as goats and sheep to identity genes under selection for adaptation to diverse environments (Álvarez et al., 2020; Edea et al., 2019; Bertolini et al., 2018; Guo et al., 2018; Mastrangelo et al., 2017). For instance, Mastrangelo et al. (2017) revealed the selection of genes involved adaptive traits such as thermoregulation,

while Bertolini et al. (2018) identified selection signals linked to coat colour. Furthermore, Guo et al. (2018) detected selection signatures involved in high-altitude adaptation (Guo et al., 2018), while Edea et al. (2019) and Álvarez et al. (2020) found selection signals for cold adaptation and thermotolerance, respectively.

2.5.1.4 Challenges for selection signature studies

The application of signatures of selection studies in livestock breeding has been successful, however, several limitations and drawbacks of have been reported. For instance, the occurrence of false-positive and false-negative results in the identification of signatures in livestock populations can lead to misleading results that mimic selection signatures (Weigand and Leese, 2018). These effects can be brought about by several factors. For instance, the removal of disadvantageous alleles through negative (purifying or background) selection can result in reduced genetic diversity across the genome similar to selection signatures (Vitti et al., 2013). Similarly, events such as population bottlenecks, expansions, and migrations can cause changes in allele frequencies due to genetic drift that reduce genetic diversity (Nielsen, 2005). While genomic regions with low recombination have been shown to mimic the effects of signatures of selection (Haasl and Payseur, 2016). In addition, ascertainment bias during SNP marker discovery can lead to the omission of low frequency alleles and since the signatures of selection analysis depends on genotypic data generated by SNP microarrays the absence of low frequency alleles may cause misleading results (Lachance and Tishkoff, 2013; Vitti et al., 2013).

In this regard, the literation has provided several methods to reduce the occurrence of false positives and negatives, as well as to address problems that may arise from the ascertainment bias (Malomane et al., 2018; Yurchenko et al., 2018; Randhawa et al., 2014; Lachance and Tishkoff, 2013). For instance, combining strategies such as like composite selection signals (CSS) and de-correlated composite of multiple signals (DCMS) to detect signatures of selection has been demonstrated as a suitable alternative that address most of the limitations of traditional methods (Yurchenko et al., 2018; Randhawa et al., 2014). To prevent the effects of ascertainment bias present by the use SNP microarrays in signatures of selection, alternative strategies such as whole-genome sequencing have been proposed. Furthermore, it has been reported that the use of larger sample sizes, and statistical methods such as maximum likelihood, Bayesian and haplotype-based methods can also limit ascertainment bias (Lachance

and Tishkoff, 2013; Malomane et al., 2018). The high costs involved in the whole genome sequencing of larger population sizes still limit the implementation of these interventions.

2.5.2 Genome-wide association studies

Genome-wide association studies (GWASs) were first used to analyse human diseases (Sharma et al., 2015). GWAS was introduced into animal genetics and breeding when information about the genomic sequences and single nucleotide polymorphisms (SNPs) of several domestic species became available through high throughput genome sequencing (Sharma et al., 2015). Before the advancements of high throughput sequencing technologies, small microsatellites found in the genomes of animals were used to map quantitative trait loci that influence traits of interest (Dekkers, 2004). Information about these loci is important because it can be used to understand the genetic mechanisms responsible for the phenotypic variation of livestock production traits and estimate their genetic breeding values.

Although the microsatellite method successfully mapped QTLs, it suffered several limitations, the foremost being low precision and only identifying a few causative genes (Dekkers, 2004). High-density SNP arrays were developed using the genomic sequences and SNP information provided by high throughput sequencing technologies. These SNP arrays enabled the identification of SNPs linked to genes that influence the phenotypic variation of traits of interest at a genomic scale, hence the analysis is called a genome-wide association study (GWAS) (Schmid and Bennewitz, 2017). The logic of GWAS is as follows, if an identified SNP has a statistically significant association with a causative gene for a trait, then that SNP is linked to the gene and thus, it directly or indirectly influences the phenotype of the trait (Schmid and Bennewitz, 2017).

Since their introduction into livestock breeding, GWASs have gained popularity and are now routinely used to map loci that influence various traits considered to be economically important. Such traits include meat quantity and quality, milk yield, percentage of fat and protein, egg production, and reproduction traits. But because these are complex traits, it is commonly found that different loci influence the same trait in different breeds of the same species (Goddard et al., 2016). Therefore, GWASs are not only ideal for identifying genes influencing complex traits, but they are also able to explain gene interactions underlying these traits.

The primary use for mapping causative genes with GWASs in livestock breeding is to provide genomic information that can be used in marker-assisted selection. This form of marker-

assisted selection is called genomic selection (Meuwissen et al., 2001). Genomic selection was introduced to improve upon traditional marker-assisted selection methods based on QTLs mapped using microsatellites. Traditional marker-assisted selection was limited because mapped QTLs explained very little of the variation observed in complex traits. Genomic selection is better suited in that it enables the estimation of genomic breeding values, which can be used do marker-assisted selection on a genome-wide scale and accelerate genetic gain (Meuwissen et al., 2001). GWASs can also be used to gain biological information about trait expression through the post-GWAS annotation of the identified genes' functions (Schmid and Bennewitz, 2017).

2.5.2.1 GWASs in livestock populations

GWASs have been carried out in several livestock species, including cattle, pigs, sheep, chickens, and goats. In cattle, a GWAS study by Bolorma et al. (2011), who used the 50K SNP chip and a 10K SNP, identified SNPs on cow chromosomes 5 and 8, which were associated with residual feed intake (RFI) in beef and dairy cattle. In pigs, a GWAS by Duijvesteijn et al. (2010), who used the Illumina Porcine 60K+SNP Beadchip, managed to identify 37 SNPs on sus scrofa chromosome 1 (SSC1) and SSC6, which affected androstenone levels in fat tissue of intact boars. In chickens, a GWAS using the 60 k SNP panel identified SNPs influencing essential traits such as body weight and growth (Gu et al., 2011; Sharma et al., 2015). Gu et al. (2011) reported nine different SNP markers which affected growth on chicken chromosome 4 (GGA4), and Sharma et al. (2015) reported the same region as containing bodyweight QTL.

In goats, morphological traits are essential for breed identification and classification. However, these traits result from several genetic and environmental factors that differ across different production systems in different geographic regions. Therefore, GWASs are needed to determine the fraction of the observed variation caused solely by genetic factors. Results from a GWAS carried out by Rahmatalla et al. (2018) reported several SNP-linked genes associated with body morphology, including adult body mass (G-protein coupled receptor 61/GPR61), morphological traits (Ring finger protein 157/RNF157), outer ear development (Homeobox protein SIX2/SIX2) and ear morphogenesis (WNT5A).

Other traits that have been focused on in goat GWASs are meat, dairy, and reproduction traits. A GWAS by (Scholtens et al., 2020) used the Caprine 50 K SNP chip and detected genetic regions associated with lactation yields of milk, fat, protein, and somatic cell score in dairy goats. The study identified a total of 43 SNPs. SNPs associated with milk, fat, and protein yield

were found on chromosome 19, while SNPs linked to somatic cell score were located on chromosome 29. A GWAS by Islam et al., (2020) used the Illumina GoatSNP53K Beadchip panel and identified 66 genomic regions in six loci associated with reproduction traits. Candidate genes included genes involved in goat fecundity trait (Islam et al., 2020).

2.5.2.2 Challenges for genome-wide association studies

The current challenges of the use of GWAS in livestock populations revolve around proper study design and execution. Due to the decreasing price of genotyping and the accumulation of freely accessible animal whole-genome data, almost anyone can carry out a GWAS on their personal computer. This, however, presents a challenge because as the number of people who can readily access this information increases, so does the possibility for the detection of false associations. So, proper study design and execution are crucial for the utilisation of GWASs.

A significant part of proper GWA study design and execution is quality control (QC). Overall data quality determines the accuracy of GWAS to detect genuine genotype-phenotype associations. QC is a method to clean genome-wide data by removing samples and SNPs with high error rates (Sharma et al., 2015). These data would otherwise cause the detection of false associations (Sharma et al., 2015). However, executing proper quality control can be a challenge as constantly evolving quality control procedures are often computationally intensive and operationally challenging. Common challenges associated with QC include data file format conversions, the installation and running of software packages for data manipulation and analysis, sample quality and relatedness, population substructure, marker quality, and batch effects (Sharma et al., 2015).

With regards to sample quality, there are several challenges a researcher can face. The first could be sample identity problems which can cause sample mix-ups. Such a problem can be the result of the mishandling of samples (Turner et al., 2011). The simplest way to deal with this issue is through checking discrepancies by comparing the reported and predicted sex of individuals within the data. If differences are found, the study can be revised to correct errors caused by sample mishandling (Turner et al., 2011). Another factor that can cause a potential challenge with sample identity is sample relatedness which can increase the occurrence of type I and type II errors (Turner et al., 2011). Statistically techniques such as mixed model regression can be used to account for sample relatedness (Aulchenko et al., 2007).

Similarly, population sub-structuring within the study data can cause false-positive results due to differences in ancestry rather than true genotype-phenotype associations (Cardon and

Palmer., 2003). GWAS assumes that the population under study is genetically homogeneous. Therefore, there should be no population stratification in the dataset (Sharma et al., 2015). Strategies to limit the effects of population structure include the use of a Mixed-model GWA analysis. Mixed models address population structure in the data by incorporating phenotypic covariance to limit the occurrence of false-positive associations (Sharma et al., 2015). Other strategies to adjust for population stratification include statistical methodologies such as Structured association (Pritchard et al., 2000) implemented in the STRUCTURE software (Hubisz et al., 2009) and Eigenstrat analysis (Patterson et al., 2006) which uses the principal components analysis to adjust for population structure (Price et al., 2006).

Regarding marker quality, challenges may arise when deciding on the threshold to use to remove SNPs with low marker genotyping efficiency, minor allele frequency and those that deviate from the Hardy-Weinberg Equilibrium (HWE). In GWAS, SNPs with low marker genotyping efficiency (i.e., low call rates) are more likely to produce false-positive results (Turner et al., 2011). A recommended range for the threshold to use when removing such SNPs is 98-99%, and the threshold should minimise the number of samples dropped while maximising genotyping efficiency (Turner et al., 2011). However, picking the appropriate threshold can be a challenge as it is affected by several other factors, such as DNA quality, the genotyping platform used, and human and equipment error during genotyping (Turner et al., 2011).

This same challenge is present when picking a threshold for removing SNPs with minor allele frequencies and those that deviate from the HWE. SNPs with minor allele frequencies are not desirable because statistical power is extremely low for rare SNPs (Turner et al., 2011). Therefore, it is recommended to remove such SNPs to lessen the computational and multiple testing correction burdens (Turner et al., 2011). However, studies have also shown that rare SNPs can be beneficial when dealing with larger sample sizes (Gorlov et al., 2008). Therefore, choosing a threshold to filter rare SNPs can depend on the size of the study. SNPs that are not in HWE are removed because they are often the result of genotyping errors or population structure. However, disequilibrium can also result from gene-trait associations (Wittke-Thompson et al., 2005). Therefore, filtering these SNPs can also become a challenge.

2.5.3 Copy number variation

Copy number variation (CNV) refers to a genetic phenomenon in which individuals in a population have varying repeats of genomic regions (1Kb-5Mb). CNVs are a source of genetic

diversity and therefore they play a vital role in evolution (Emerson et al., 2008). CNVs overlapping protein coding regions can alter protein function, while those found in the regulatory regions of genes can affect gene expression levels and therefore impact phenotype variation (Liu et al., 2009). CNVs result from rearrangements in the genome caused by errors in normal DNA processes. For instance, when non-homologous chromosomes recombine during meiosis or mitosis due sharing similar to DNA sequences it can result in the duplication, deletion, or inversion of those chromosomal segments during crossing-over (Hastings et al., 2009; Gu et al., 2008). Furthermore, the deletion of nucleotides resulting from errors in DNA repair processes such as non-homologous end-joining can produce CNVs (Bhanuprakash et al., 2018). Other DNA processes that can produce CNVs include L1-mediated retro transposition (Ostertag and Kazazian Jr, 2001) and fork stalling and template switching (Lee et al., 2007).

2.5.3.1 Methods used for the Identification of CNV

The availability of SNP genotyping arrays and the dropping costs of genotyping individuals have enable the study of CNVs in several livestock including cattle (Wang et al., 2015; Jiang et al., 2012; Hou et al., 2011; Bae et al., 2010), sheep (Yang et al., 2018; Fontanesi et al., 2011), goats (Fontanesi et al., 2010), pigs, (Paudel et al., 2015; Chen et al., 2012; Ramayo-Caldas et al., 2010; Fadista et al., 2008) and chickens (Crooijmans et al., 2013). In addition, the decreased cost of next-generation sequencing which covers the whole genome now offers the opportunity to study CNVs in livestock breeding at a much higher resolution compared to SNP genotype data (Alkan et al., 2011). For the detection of CNVs several algorithms such as PennCNV, CnvPartition, cn.MOPS (Mixture of PoissonS), QuantiSNP, and CNVFinder are available (Bhanuprakash et al., 2018; Yavaş et al., 2009; Korn et al., 2008; Wang et al., 2007; Colella et al., 2007).

PennCNV uses a Hidden Markov Model to detect CNVs (Wang et al., 2007; Colella et al., 2007; Peiffer et al., 2006). The algorithm improves the call rate and overcomes genomic waves by fitting regression models with GC content (Bickhart et al., 2012). In addition, and it uses pedigree information to improve boundary mapping (Hou et al., 2011). The CnvPartition algorithm detects CNVs by processing Log R ratio and B allele frequency through a strategy that uses a sliding window. To ensure quality calls, CnvPartition only detects homozygous deletion events (Cooper et al., 2011). The Mixture of PoissonS (cn.MOPS) algorithm uses a Bayesian statistically analysis on next-generation sequence data to detect overlapping genomic regions and allele copy number variations (Klambauer et al., 2012). QuantiSNP is comparable

to PennCNV, because it also performs CNV detection using Hidden Markov Models. The two algorithms differ in that QuantiSNP independently considers Log R ratios and B allele frequencies whereas PennCNV combines them. CNVFinder uses whole exome sequencing data to identify CNVs in python (Fiegler et al., 2006).

2.5.3.2 Applications for CNV detection in livestock production

The development of next-generation sequencing techniques together with recent drops in sequencing costs, offer researchers the opportunity to investigate the contributions of copy number variations to genetic variations in livestock and its effect on the variation of economical traits (Bhanuprakash et al., 2018). The use of CNV detection to explain genetic variation by identifying copy number variation regions overlapping economical traits can be used to accelerate livestock genetic improvement. In addition, because CNVs explain more of the genetic variation than SNP sites, it can offer insightful information about the genetic mechanisms involved in complex process such as evolution, the domestication livestock populations and their adaptation to diverse environments (Bhanuprakash et al., 2018). The detection of copy number variation has aided in the development of CNV maps that can be used in the detection of copy number variation regions that overlap the regulatory regions of genes that influence livestock production traits (Bhanuprakash et al., 2018). Thus, improvement of CNV detection strategies could provide vital information about the genetic variation of livestock species which can be used to inform improved practices for genomic selection in farm animals, thus increasing genetic improvement (Bhanuprakash et al., 2018).

Examples of CNV detection studies in goat populations include the study by Fontanesi et al. (2010) which studied the CNVs in the goat genome and identified one hundred and twenty copy number variation regions overlapping genes involved in environmental adaptation in addition, gene ontology analysis and functional studies have shown that genes responsible for goat and sheep environmental adaptation are found on copy number variation regions (Ma et al., 2015). Furthermore, Lai et al. (2008) studied CNVs in small ruminants such as goats and sheep and revealed that coat colour variation is influenced by copy number variation at the Agouti locus.

2.6 Conclusion

Selection signatures, GWAS, and CNV all present an opportunity for mapping trait-associated SNPs and genes underlying the genetic variation of quantitative traits in livestock populations. With the decrease of genotyping costs and the increase of whole-genome data available for livestock populations, it can be expected that the number of detected causative SNPs for traits with high economic value will increase.

2.7 References

AJMONE-MARSAN, P., COLLI, L., JIANLIN, H., ACHILLI, A., HOVIRAG, L., JOOST, S., PAOLA, C., PILLA, F., STELLA, A., TABERLET, P., BOETTCHER, P., NEGRINI, R., LENSTRA, J., GOAT, I., CONSORTIA, E. & CONSORTIA, G. 2014. The characterization of goat genetic diversity: Towards a genomic approach. *Small Ruminant Research,* 121.

ALKAN, C., COE, B. P. & EICHLER, E. E. 2011. Genome structural variation discovery and genotyping. *Nature reviews genetics,* 12**,** 363-376.

ÁLVAREZ, I., FERNÁNDEZ, I., TRAORÉ, A., PÉREZ-PARDAL, L., MENÉNDEZ-ARIAS, N. A. & GOYACHE, F. 2020. Genomic scan of selective sweeps in Djallonké (West African Dwarf) sheep shed light on adaptation to harsh environments. *Scientific Reports,* 10**,** 2824.

AMILLS, M. 2014. The Application of Genomic Technologies to Investigate the Inheritance of Economically Important Traits in Goats. *Advances in Biology,* 2014.

AMILLS, M., CAPOTE, J. & TOSSER-KLOPP, G. 2017. Goat domestication and breeding: a jigsaw of historical, biological and molecular data with missing pieces. *Anim Genet,* 48**,** 631-644.

AULCHENKO, Y. S., DE KONING, D. J. & HALEY, C. 2007. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics,* 177**,** 577-85.

AZIZ, M. 2010. Present status of the world goat populations and their productivity. *Lohmann Information,* 45.

BAE, J. S., CHEONG, H. S., KIM, L. H., NAMGUNG, S., PARK, T. J., CHUN, J.-Y., KIM, J. Y., PASAJE, C. F. A., LEE, J. S. & SHIN, H. D. 2010. Identification of copy number variations and common deletion polymorphisms in cattle. *BMC genomics,* 11**,** 1-10.

BERTOLINI, F., SERVIN, B., TALENTI, A., ROCHAT, E., KIM, E. S., OGET, C., PALHIÈRE, I., CRISÀ, A., CATILLO, G., STERI, R., AMILLS, M., COLLI, L., MARRAS, G., MILANESI, M., NICOLAZZI, E., ROSEN, B. D., VAN TASSELL, C. P., GULDBRANDTSEN, B., SONSTEGARD, T. S., TOSSER-KLOPP, G., STELLA, A., ROTHSCHILD, M. F., JOOST, S., CREPALDI, P. & THE ADAPTMAP, C. 2018. Signatures of selection and environmental adaptation across the goat genome post-domestication. *Genetics Selection Evolution,* 50**,** 57.

BHANUPRAKASH, V., CHHOTARAY, S., PRUTHVIRAJ, D. R., RAWAT, C., KARTHIKEYAN, A. & PANIGRAHI, M. 2018. Copy number variation in livestock: A mini review. *Veterinary world,* 11**,** 535-541.

BICKHART, D. M., HOU, Y., SCHROEDER, S. G., ALKAN, C., CARDONE, M. F., MATUKUMALLI, L. K., SONG, J., SCHNABEL, R. D., VENTURA, M., TAYLOR, J. F., GARCIA, J. F., VAN TASSELL, C. P., SONSTEGARD, T. S., EICHLER, E. E. & LIU, G. E. 2012. Copy number variation of individual cattle genomes using next-generation sequencing. *Genome Res,* 22**,** 778-90.

BOLORMAA, S., HAYES, B. J., SAVIN, K., HAWKEN, R., BARENDSE, W., ARTHUR, P. F., HERD, R. M. & GODDARD, M. E. 2011. Genome-wide association studies for feedlot and growth traits in cattle. *J Anim Sci,* 89**,** 1684-97.

BOMBA, L., NICOLAZZI, E. L., MILANESI, M., NEGRINI, R., MANCINI, G., BISCARINI, F., STELLA, A., VALENTINI, A. & AJMONE-MARSAN, P. 2015. Relative extended haplotype homozygosity signals across breeds reveal dairy and beef specific signatures of selection. *Genetics Selection Evolution,* 47**,** 25.

BOSMAN, L., VAN MARLE-KOSTER, E. & VISSER, C. 2015a. Population structure of South African commercial dairy goats. *European Regional Conference on Goats 2014*.

BOSMAN, L., VAN MARLE-KÖSTER, E. & VISSER, C. 2015b. Genetic diversity of South African dairy goats for genetic management and improvement. *Small Ruminant Research,* 123**,** 224-231.

BRAVERMAN, J. M., HUDSON, R. R., KAPLAN, N. L., LANGLEY, C. H. & STEPHAN, W. 1995. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. *Genetics,* 140**,** 783-96.

BRITO, L. F., KIJAS, J. W., VENTURA, R. V., SARGOLZAEI, M., PORTO-NETO, L. R., CÁNOVAS, A., FENG, Z., JAFARIKIA, M. & SCHENKEL, F. S. 2017. Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers. *BMC Genomics,* 18**,** 229.

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**,** 338-348.

CADZOW, M., BOOCOCK, J., NGUYEN, H. T., WILCOX, P., MERRIMAN, T. R. & BLACK, M. A. 2014. A bioinformatics workflow for detecting signatures of selection in genomic data. *Frontiers in Genetics,* 5.

CAMPBELL, Q. 2003. The origin and description of southern Africa's indigenous goats. *S Afr J Anim Sci,* 4**,** 18-22.

CARDON, L. R. & PALMER, L. J. 2003. Population stratification and spurious allelic association. *Lancet,* 361**,** 598-604.

CESARANI, A., SORBOLINI, S., CRISCIONE, A., BORDONARO, S., PULINA, G., BATTACONE, G., MARLETTA, D., GASPA, G. & MACCIOTTA, N. P. P. 2018. Genome-wide variability and selection signatures in Italian island cattle breeds. *Animal Genetics,* 49**,** 371-383.

CHEN, C., QIAO, R., WEI, R., GUO, Y., AI, H., MA, J., REN, J. & HUANG, L. 2012. A comprehensive survey of copy number variation in 18 diverse pig populations and identification of candidate copy number variable genes associated with complex traits. *BMC genomics,* 13**,** 1-10.

CHEN, M., PAN, D., REN, H., FU, J., LI, J., SU, G., WANG, A., JIANG, L., ZHANG, Q. & LIU, J.-F. 2016. Identification of selective sweeps reveals divergent selection between Chinese Holstein and Simmental cattle populations. *Genetics Selection Evolution,* 48**,** 76.

CHOKOE, T. C., MDLADLA-HADEBE, K., MUCHADEYI, F., DZOMBA, E., MATELELE, T., MPHAHLELE, T., MPOFU, T. J., NEPHAWE, K. & MTILENI, B. 2020. Genetic Diversity of South African Indigenous Goat Population from Four Provinces Using Genome-Wide SNP Data. *Sustainability,* 12**,** 10361.

COLELLA, S., YAU, C., TAYLOR, J. M., MIRZA, G., BUTLER, H., CLOUSTON, P., BASSETT, A. S., SELLER, A., HOLMES, C. C. & RAGOUSSIS, J. 2007. QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic acids research,* 35**,** 2013-2025.

COOPER, G. M., COE, B. P., GIRIRAJAN, S., ROSENFELD, J. A., VU, T. H., BAKER, C., WILLIAMS, C., STALKER, H., HAMID, R., HANNIG, V., ABDEL-HAMID, H., BADER, P., MCCRACKEN, E., NIYAZOV, D., LEPPIG, K., THIESE, H., HUMMEL, M., ALEXANDER, N., GORSKI, J., KUSSMANN, J., SHASHI, V., JOHNSON, K., REHDER, C., BALLIF, B. C., SHAFFER, L. G. & EICHLER, E. E. 2011. A copy number variation morbidity map of developmental delay. *Nat Genet,* 43**,** 838-46.

COSTER, A., BASTIAANSEN, J. W., CALUS, M. P., VAN ARENDONK, J. A. & BOVENHUIS, H. 2010. Sensitivity of methods for estimating breeding values using genetic markers to the number of QTL and distribution of QTL variance. *Genetics Selection Evolution,* 42**,** 1-11.

CROOIJMANS, R. P., FIFE, M. S., FITZGERALD, T. W., STRICKLAND, S., CHENG, H. H., KAISER, P., REDON, R. & GROENEN, M. A. 2013. Large scale variation in DNA copy number in chicken breeds. *BMC genomics,* 14**,** 1-10.

DEKKERS, J. C. 2004. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. *J Anim Sci,* 82 E-Suppl**,** E313-328.

DEKKERS, J. C. 2011. Use of high-density marker genotyping for genetic improvement of livestock by genomic selection. *Animal Science Reviews 2010***,** 197.

DEKKERS, J. C. & HOSPITAL, F. 2002. The use of molecular genetics in the improvement of agricultural populations. *Nat Rev Genet,* 3**,** 22-32.

DUBEUF, J.-P. & BOYAZOGLU, J. 2009. An international panorama of goat selection and breeds. *Livestock Science,* 120**,** 225-231.

DUIJVESTEIJN, N., KNOL, E. F., MERKS, J. W., CROOIJMANS, R. P., GROENEN, M. A., BOVENHUIS, H. & HARLIZIUS, B. 2010. A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6. *BMC Genet,* 11**,** 42.

EDEA, Z., DADI, H., DESSIE, T. & KIM, K.-S. 2019. Genomic signatures of high-altitude adaptation in Ethiopian sheep populations. *Genes & Genomics,* 41**,** 973-981.

EMERSON, J., CARDOSO-MOREIRA, M., BOREVITZ, J. O. & LONG, M. 2008. Natural selection shapes genome-wide patterns of copy-number polymorphism in Drosophila melanogaster. *science,* 320**,** 1629-1631.

FADISTA, J., NYGAARD, M., HOLM, L.-E., THOMSEN, B. & BENDIXEN, C. 2008. A snapshot of CNVs in the pig genome. *PloS one,* 3**,** e3916.

FARIELLO, M.-I., SERVIN, B., TOSSER-KLOPP, G., RUPP, R., MORENO, C., CONSORTIUM, I. S. G., CRISTOBAL, M. S. & BOITARD, S. 2014. Selection signatures in worldwide sheep populations. *PloS one,* 9**,** e103813.

FAY, J. C. & WU, C. I. 2000. Hitchhiking under positive Darwinian selection. *Genetics,* 155**,** 1405-13.

FIEGLER, H., REDON, R., ANDREWS, D., SCOTT, C., ANDREWS, R., CARDER, C., CLARK, R., DOVEY, O., ELLIS, P., FEUK, L., FRENCH, L., HUNT, P., KALAITZOPOULOS, D., LARKIN, J., MONTGOMERY, L., PERRY, G. H., PLUMB, B. W., PORTER, K., RIGBY, R. E., RIGLER, D., VALSESIA, A., LANGFORD, C., HUMPHRAY, S. J., SCHERER, S. W., LEE, C., HURLES, M. E. & CARTER, N. P. 2006. Accurate and reliable high-throughput detection of copy number variation in the human genome. *Genome research,* 16**,** 1566-1574.

FONTANESI, L., BERETTI, F., MARTELLI, P., COLOMBO, M., DALL'OLIO, S., OCCIDENTE, M., PORTOLANO, B., CASADIO, R., MATASSINO, D. & RUSSO, V. 2011. A first comparative map of copy number variations in the sheep genome. *Genomics,* 97**,** 158-165.

FONTANESI, L., MARTELLI, P. L., BERETTI, F., RIGGIO, V., DALL'OLIO, S., COLOMBO, M., CASADIO, R., RUSSO, V. & PORTOLANO, B. 2010. An initial comparative map of copy number variations in the goat (Capra hircus) genome. *BMC genomics,* 11**,** 1-15.

GODDARD, M. E., KEMPER, K. E., MACLEOD, I. M., CHAMBERLAIN, A. J. & HAYES, B. J. 2016. Genetics of complex traits: prediction of phenotype, identification of causal polymorphisms and genetic architecture. *Proc Biol Sci,* 283.

GORLOV, I. P., GORLOVA, O. Y., SUNYAEV, S. R., SPITZ, M. R. & AMOS, C. I. 2008. Shifting paradigm of association studies: value of rare single-nucleotide polymorphisms. *Am J Hum Genet,* 82**,** 100-12.

GROENEVELD, L. F., LENSTRA, J. A., EDING, H., TORO, M. A., SCHERF, B., PILLING, D., NEGRINI, R., FINLAY, E. K., JIANLIN, H., GROENEVELD, E. & WEIGEND, S. 2010. Genetic diversity in farm animals--a review. *Anim Genet,* 41 Suppl 1**,** 6-31.

GU, W., ZHANG, F. & LUPSKI, J. R. 2008. Mechanisms for human genomic rearrangements. *Pathogenetics,* 1**,** 1-17.

GU, X., FENG, C., MA, L., SONG, C., WANG, Y., DA, Y., LI, H., CHEN, K., YE, S., GE, C., HU, X. & LI, N. 2011. Genome-wide association study of body weight in chicken F2 resource population. *PLoS One,* 6**,** e21872.

GUO, J., TAO, H., LI, P., LI, L., ZHONG, T., WANG, L., MA, J., CHEN, X., SONG, T. & ZHANG, H. 2018. Whole-genome sequencing reveals selection signatures associated with important traits in six goat breeds. *Scientific Reports,* 8**,** 10405.

GURGUL, A., JASIELCZUK, I., ROPKA-MOLIK, K., SEMIK-GURGUL, E., PAWLINA-TYSZKO, K., SZMATOŁA, T., SZYNDLER-NĘDZA, M., BUGNO-PONIEWIERSKA, M., BLICHARSKI, T., SZULC, K., SKRZYPCZAK, E. & KRUPIŃSKI, J. 2018. A genome-wide detection of selection signatures in conserved and commercial pig breeds maintained in Poland. *BMC Genetics,* 19**,** 95.

HAASL, R. J. & PAYSEUR, B. A. 2016. Fifteen years of genomewide scans for selection: trends, lessons and unaddressed genetic sources of complication. *Molecular Ecology,* 25**,** 5-23.

HASTINGS, P. J., LUPSKI, J. R., ROSENBERG, S. M. & IRA, G. 2009. Mechanisms of change in gene copy number. *Nature Reviews Genetics,* 10**,** 551-564.

HAYES, B. J., BOWMAN, P. J., CHAMBERLAIN, A. J. & GODDARD, M. E. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science,* 92**,** 433-443.

HEFER, C., KOTZE, A., VAN MARLE-KOSTER, E. & VISSER, C. 2004. Genetic variation of three commercial and three indigenous goat populations in South Africa. *South African Journal of Animal Science,* 34**,** 24-27.

HERMISSON, J. & PENNINGS, P. 2017. Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods in Ecology and Evolution,* 8**,** 700-716.

HOU, Y., LIU, G. E., BICKHART, D. M., CARDONE, M. F., WANG, K., KIM, E.-S., MATUKUMALLI, L. K., VENTURA, M., SONG, J. & VANRADEN, P. M. 2011. Genomic characteristics of cattle copy number variations. *BMC genomics,* 12**,** 1-11.

HOWIE, B. N., DONNELLY, P. & MARCHINI, J. 2009. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLOS Genetics,* 5**,** e1000529.

HUBISZ, M., FALUSH, D., STEPHENS, M. & PRITCHARD, J. 2009. Hubisz MJ, Falush D, Stephens M, Pritchard JK.. Inferring weak population structure with the assistance of sample group information. Mol Ecol Res 9: 1322-1332. *Molecular ecology resources,* 9**,** 1322-32.

HUSON, H., SONSTEGARD, T., SILVERSTEIN, J., WOODWARD-GREENE, M., MASIGA, C., MUCHADEYI, F., REES, J., SAYRE, B., ELBETAGY, A. & ROTHSCHILD, M. African Goat Improvement Network (AGIN)(2014). Genetic and phenotypic characterization of African goat populations to prioritize conservation and production efforts for small-holder farmers in Sub Saharan Africa. Proceedings of the 10th World Congress of Genetics Applied to Livestock Production. 1-4.

ISLAM, R., LIU, X., GEBRESELASSIE, G., ABIED, A., MA, Q. & MA, Y. 2020. Genome-wide association analysis reveals the genetic locus for high reproduction trait in Chinese Arbas Cashmere goat. *Genes & Genomics,* 42.

JENSEN, J. D., FOLL, M. & BERNATCHEZ, L. 2016. The past, present and future of genomic scans for selection. *Mol Ecol,* 25**,** 1-4.

JIANG, L., JIANG, J., WANG, J., DING, X., LIU, J. & ZHANG, Q. 2012. Genome-wide identification of copy number variations in Chinese Holstein. *PloS one,* 7**,** e48732.

KLAMBAUER, G., SCHWARZBAUER, K., MAYR, A., CLEVERT, D.-A., MITTERECKER, A., BODENHOFER, U. & HOCHREITER, S. 2012. cn.MOPS: mixture of Poissons for discovering copy number variations in next-generation sequencing data with a low false discovery rate. *Nucleic acids research,* 40**,** e69-e69.

KÖNIG, S. & SWALVE, H. 2009. Application of selection index calculations to determine selection strategies in genomic breeding programs. *Journal of Dairy Science,* 92**,** 5292-5303.

KORN, J. M., KURUVILLA, F. G., MCCARROLL, S. A., WYSOKER, A., NEMESH, J., CAWLEY, S., HUBBELL, E., VEITCH, J., COLLINS, P. J. & DARVISHI, K. 2008. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nature genetics,* 40**,** 1253-1260.

KOTZE, A., SWART, H., GROBLER, J. & NEMAANGANI, A. 2004. A genetic profile of the Kalahari Red goat breed from southern Africa. *South African Journal of Animal Science,* 34.

LACHANCE, J. & TISHKOFF, S. A. 2013. SNP ascertainment bias in population genetic analyses: Why it is important, and how to correct it. *BioEssays,* 35**,** 780-786.

LAI, W. R., CHOUDHARY, V. & PARK, P. J. 2008. CGHweb: a tool for comparing DNA copy number segmentations from multiple algorithms. *Bioinformatics,* 24 7**,** 1014-5.

LANDE, R. & THOMPSON, R. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics,* 124**,** 743-756.

LASHMAR, S., VISSER, C. & VAN MARLE-KÖSTER, E. 2016. SNP-based genetic diversity of South African commercial dairy and fiber goat breeds. *Small Ruminant Research,* 136**,** 65-71.

LEE, J. A., CARVALHO, C. M. & LUPSKI, J. R. 2007. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *cell,* 131**,** 1235-1247.

LISCHER, H. E. L. & EXCOFFIER, L. 2011. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics,* 28**,** 298-299.

LIU, G. E., VENTURA, M., CELLAMARE, A., CHEN, L., CHENG, Z., ZHU, B., LI, C., SONG, J. & EICHLER, E. E. 2009. Analysis of recent segmental duplications in the bovine genome. *BMC genomics,* 10**,** 1-16.

MA, Y., ZHANG, Q., LU, Z., ZHAO, X. & ZHANG, Y. 2015. Analysis of copy number variations by SNP50 BeadChip array in Chinese sheep. *Genomics,* 106**,** 295-300.

MAHANJANA, A. & CRONJE, P. 2000. Factors affecting goat production in a communal farming system in the Eastern Cape region of South Africa. *South African Journal of Animal Science,* 30**,** 149-155.

MAIORANO, A. M., LOURENCO, D. L., TSURUTA, S., OSPINA, A. M. T., STAFUZZA, N. B., MASUDA, Y., FILHO, A. E. V., CYRILLO, J. N. D. S. G., CURI, R. A. & SILVA, J. A. I. I. D. V. 2018. Assessing genetic architecture and signatures of selection of dual purpose Gir cattle populations using genomic information. *PLOS ONE,* 13**,** e0200694.

MALAN, S. W. 2000. The improved Boer goat. *Small Rumin Res,* 36**,** 165-170.

MALOMANE, D. K., REIMER, C., WEIGEND, S., WEIGEND, A., SHARIFI, A. R. & SIMIANER, H. 2018. Efficiency of different strategies to mitigate ascertainment bias when using SNP panels in diversity studies. *BMC Genomics,* 19**,** 22.

MANUNZA, A., CARDOSO, T. F., NOCE, A., MARTÍNEZ, A., PONS, A., BERMEJO, L. A., LANDI, V., SÀNCHEZ, A., JORDANA, J., DELGADO, J. V., ADÁN, S., CAPOTE, J., VIDAL, O., UGARTE, E., ARRANZ, J. J., CALVO, J. H., CASELLAS, J. & AMILLS, M. 2016. Population structure of eleven Spanish ovine breeds and detection of selective sweeps with BayeScan and hapFLK. *Scientific Reports,* 6**,** 27296.

MAREE, C. & PLUG, I. 1993. Origin of southern African livestock and their potential role in the industry. *Livestock production systems: principles and practice***,** 5-13.

MASTRANGELO, S., TOLONE, M., SARDINA, M. T., SOTTILE, G., SUTERA, A. M., DI GERLANDO, R. & PORTOLANO, B. 2017. Genome-wide scan for runs of homozygosity identifies potential candidate genes associated with local adaptation in Valle del Belice sheep. *Genetics Selection Evolution,* 49**,** 84.

MATUKUMALLI, L. K., LAWLEY, C. T., SCHNABEL, R. D., TAYLOR, J. F., ALLAN, M. F., HEATON, M. P., O'CONNELL, J., MOORE, S. S., SMITH, T. P., SONSTEGARD, T. S. & VAN TASSELL, C. P. 2009. Development and characterization of a high density SNP genotyping assay for cattle. *PLoS One,* 4**,** e5350.

MDLADLA, K., DZOMBA, E. & MUCHADEYI, F. 2016a. P5039 A landscape genomic approach to unravel the genomic mechanism of adaptation in indigenous goats of South Africa. *Journal of Animal Science,* 94**,** 134-135.

MDLADLA, K., DZOMBA, E. F., HUSON, H. J. & MUCHADEYI, F. C. 2016b. Population genomic structure and linkage disequilibrium analysis of South African goat breeds using genome-wide SNP data. *Anim Genet,* 47**,** 471-82.

MDLADLA, K., DZOMBA, E. F. & MUCHADEYI, F. C. 2017. Characterization of the village goat production systems in the rural communities of the Eastern Cape, KwaZulu-Natal, Limpopo and North West Provinces of South Africa. *Trop Anim Health Prod,* 49**,** 515-527.

MEUWISSEN, T. H., HAYES, B. J. & GODDARD, M. E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics,* 157**,** 1819-29.

MOHLATLOLE, R. P., DZOMBA, E. F. & MUCHADEYI, F. C. 2015. Addressing production challenges in goat production systems of South Africa: the genomics approach. *Small Ruminant Research,* 131**,** 43-49.

MULLER, C. 2005. *Genetic parameter estimation and breeding plans for the South African dairy goat herd.* Stellenbosch: University of Stellenbosch.

NCUBE, K. T., HADEBE, K., DZOMBA, E. F., SOMA, P., FRYLINCK, L. & MUCHADEYI, F. C. 2020. Relationship between population genomic structure and growth profiles of South African goats under different production systems. *Trop Anim Health Prod,* 52**,** 1277-1286.

NICOLAZZI, E. L., BIFFANI, S., BISCARINI, F., OROZCO TER WENGEL, P., CAPRERA, A., NAZZICARI, N. & STELLA, A. 2015. Software solutions for the livestock genomics SNP array revolution. *Animal Genetics,* 46**,** 343-353.

NIELSEN, R. 2005. Molecular signatures of natural selection. *Annu Rev Genet,* 39**,** 197-218.

OLEKSYK, T. K., SMITH, M. W. & O'BRIEN, S. J. 2010. Genome-wide scans for footprints of natural selection. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences,* 365**,** 185-205.

OLIVIER, J., CLOETE, S., SCHOEMAN, S. & MULLER, C. 2005. Performance testing and recording in meat and dairy goats. *Small Ruminant Research,* 60**,** 83-93.

ONZIMA, R., UPADHYAY, M., MUKIIBI, R., KANIS, E., GROENEN, M. & CROOIJMANS, R. 2018. Genome-wide population structure and admixture analysis reveals weak differentiation among Ugandan goat breeds. *Animal genetics,* 49**,** 59-70.

OSTERTAG, E. M. & KAZAZIAN JR, H. H. 2001. Biology of mammalian L1 retrotransposons. *Annual review of genetics,* 35**,** 501-538.

OTTO, S. P. 2000. Detecting the form of selection from DNA sequence data. *Trends Genet,* 16**,** 526-9.

PATTERSON, N., PRICE, A. L. & REICH, D. 2006. Population structure and eigenanalysis. *PLoS Genet,* 2**,** e190.

PAUDEL, Y., MADSEN, O., MEGENS, H.-J., FRANTZ, L. A., BOSSE, M., CROOIJMANS, R. P. & GROENEN, M. A. 2015. Copy number variation in the speciation of pigs: a possible prominent role for olfactory receptors. *BMC genomics,* 16**,** 1-14.

PEIFFER, D. A., LE, J. M., STEEMERS, F. J., CHANG, W., JENNIGES, T., GARCIA, F., HADEN, K., LI, J., SHAW, C. A. & BELMONT, J. 2006. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome research,* 16**,** 1136-1148.

PIETERS, A., MARLE-KOSTER, V., VISSER, C. & KOTZE, A. 2009. South African developed meat type goats: A forgotten animal genetic resource?

PORTER, V., ALDERSON, L., HALL, S. J. & SPONENBERG, D. P. 2016. *Mason's World Encyclopedia of Livestock Breeds and Breeding, 2 Volume Pack*, Cabi.

PRICE, A. L., PATTERSON, N. J., PLENGE, R. M., WEINBLATT, M. E., SHADICK, N. A. & REICH, D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics,* 38**,** 904-909.

PRITCHARD, J. K., PICKRELL, J. K. & COOP, G. 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr Biol,* 20**,** R208-15.

PRITCHARD, J. K., STEPHENS, M. & DONNELLY, P. 2000. Inference of population structure using multilocus genotype data. *Genetics,* 155**,** 945-959.

PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M. A., BENDER, D., MALLER, J., SKLAR, P., DE BAKKER, P. I. & DALY, M. J. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics,* 81**,** 559-575.

QANBARI, S. & SIMIANER, H. 2014. Mapping signatures of positive selection in the genome of livestock. *Livestock Science,* 166**,** 133-143.

RAHMATALLA, S., ARENDS, D., REISSMANN, M., WIMMERS, K., REYER, H. & BROCKMANN, G. 2018. Genome-wide association study of body morphological traits in Sudanese goats. *Animal Genetics,* 49.

RAMAYO-CALDAS, Y., CASTELLÓ, A., PENA, R. N., ALVES, E., MERCADÉ, A., SOUZA, C. A., FERNÁNDEZ, A. I., PEREZ-ENCISO, M. & FOLCH, J. M. 2010. Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. *BMC genomics,* 11**,** 1-10.

RANDHAWA, I. A. S., KHATKAR, M. S., THOMSON, P. C. & RAADSMA, H. W. 2014. Composite selection signals can localize the trait specific genomic regions in multi-breed populations of cattle and sheep. *BMC Genetics,* 15**,** 34.

R CORE TEAM. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

ROSHYARA, N. R. & SCHOLZ, M. 2014. fcGENE: A Versatile Tool for Processing and Transforming SNP Datasets. *PLOS ONE,* 9**,** e97589.

SAMBRAUS, H. H. 1992. A colour atlas of livestock breeds: 220 breeds in words and pictures.

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**,** 104257.

SCHAEFFER, L. 2006. Strategy for applying genome-wide selection in dairy cattle. *Journal of animal Breeding and genetics,* 123**,** 218-223.

SCHMID, M. & BENNEWITZ, J. 2017. Invited review: Genome-wide association analysis for quantitative traits in livestock – a selective review of statistical models and experimental designs. *Arch. Anim. Breed.,* 60**,** 335-346.

SCHOLTENS, M., JIANG, A., SMITH, A., LITTLEJOHN, M., LEHNERT, K., SNELL, R., LOPEZ-VILLALOBOS, N., GARRICK, D. & BLAIR, H. 2020. Genome-wide association studies of lactation yields of milk, fat, protein and somatic cell score in New Zealand dairy goats. *Journal of Animal Science and Biotechnology,* 11**,** 55.

SHARMA, A., LEE, J. S., DANG, C. G., SUDRAJAD, P., KIM, H. C., YEON, S. H., KANG, H. S. & LEE, S.-H. 2015. Stories and Challenges of Genome Wide Association Studies in Livestock — A Review. *Asian-Australas J Anim Sci,* 28**,** 1371-1379.

SMITH, C. & SIMPSON, S. 2010. The use of genetic polymorphisms in livestock improvement. *Journal of Animal Breeding and Genetics,* 103**,** 205-217.

SMITH, J. M. & HAIGH, J. 2007. The hitch-hiking effect of a favourable gene. *Genet Res,* 89**,** 391-403.

SNYMAN, M. 2014. South African goat breeds: Angora goat. *Info-pack Ref,* 1.

SOLLER, M. 1978. The use of loci associated with quantitative effects in dairy cattle improvement. *Animal Science,* 27**,** 133-139.

SPONENBERG, D. P. 2020. Conserving the Genetic Diversity of Domesticated Livestock. *Diversity,* 12**,** 282.

TABERLET, P., COISSAC, E., PANSU, J. & POMPANON, F. 2011. Conservation genetics of cattle, sheep, and goats. *C R Biol,* 334**,** 247-54.

TABERLET, P., VALENTINI, A., REZAEI, H. R., NADERI, S., POMPANON, F., NEGRINI, R. & AJMONE-MARSAN, P. 2008. Are cattle, sheep, and goats endangered species? *Mol Ecol,* 17**,** 275-84.

TAM, V., PATEL, N., TURCOTTE, M., BOSSÉ, Y., PARÉ, G. & MEYRE, D. 2019. Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics,* 20**,** 467-484.

TOSSER-KLOPP, G., BARDOU, P., BOUCHEZ, O., CABAU, C., CROOIJMANS, R., DONG, Y., DONNADIEU-TONON, C., EGGEN, A., HEUVEN, H. C., JAMLI, S., JIKEN, A. J., KLOPP, C., LAWLEY, C. T., MCEWAN, J., MARTIN, P., MORENO, C. R., MULSANT, P., NABIHOUDINE, I., PAILHOUX, E., PALHIÈRE, I., RUPP, R., SARRY, J., SAYRE, B. L., TIRCAZES, A., WANG, J., WANG, W. & ZHANG, W. 2016. Correction: Design and Characterization of a 52K SNP Chip for Goats. *PLoS One,* 11**,** e0152632.

TURNER, S., ARMSTRONG, L. L., BRADFORD, Y., CARLSON, C. S., CRAWFORD, D. C., CRENSHAW, A. T., DE ANDRADE, M., DOHENY, K. F., HAINES, J. L., HAYES, G., JARVIK, G., JIANG, L., KULLO, I. J., LI, R., LING, H., MANOLIO, T. A., MATSUMOTO, M., MCCARTY, C. A., MCDAVID, A. N., MIREL, D. B., PASCHALL, J. E., PUGH, E. W., RASMUSSEN, L. V., WILKE, R. A., ZUVICH, R. L. & RITCHIE, M. D. 2011. Quality control procedures for genome-wide association studies. *Curr Protoc Hum Genet,* Chapter 1**,** Unit1.19.

VANRADEN, P. M. 2008. Efficient methods to compute genomic predictions. *Journal of dairy science,* 91**,** 4414-4423.

VANRADEN, P. M., VAN TASSELL, C. P., WIGGANS, G. R., SONSTEGARD, T. S., SCHNABEL, R. D., TAYLOR, J. F. & SCHENKEL, F. S. 2009. Invited Review: Reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science,* 92**,** 16-24.

VERBYLA, K. L., BOWMAN, P. J., HAYES, B. J. & GODDARD, M. E. Sensitivity of genomic selection to using different prior distributions. BMC proceedings, 2010. Springer, 1-4.

VISSER, C., HEFER, C., VAN MARLE-KOSTER, E. & KOTZE, A. 2004. Genetic variation of three commercial and three indigenous goat populations in South Africa. *South African Journal of Animal Science,* 34**,** 24-27.

VISSER, C., LASHMAR, S. F., VAN MARLE-KÖSTER, E., POLI, M. A. & ALLAIN, D. 2016. Genetic diversity and population structure in South African, French and Argentinian Angora goats from genome-wide SNP data. *PloS one,* 11**,** e0154353.

VISSER, C. & VAN MARLE-KOSTER, E. 2018. The Development and Genetic Improvement of South African Goats.

VISSER, C., VAN MARLE-KÖSTER, E., BOVENHUIS, H. & CROOIJMANS, R. 2011a. QTL for mohair traits in South African Angora goats. *Small ruminant research,* 100**,** 8-14.

VISSER, C., VAN MARLE-KÖSTER, E. & FRIEDRICH, H. 2011b. Parentage verification of South African Angora goats, using microsatellite markers. *South African Journal of Animal Science,* 41**,** 250-255.

VISSER, C., VAN MARLE-KÖSTER, E., SNYMAN, M., BOVENHUIS, H. & CROOIJMANS, R. 2013. Quantitative trait loci associated with pre-weaning growth in South African Angora goats. *Small Ruminant Research,* 112**,** 15-20.

VITTI, J. J., GROSSMAN, S. R. & SABETI, P. C. 2013. Detecting Natural Selection in Genomic Data. *Annual Review of Genetics,* 47**,** 97-120.

WANG, K., LI, M., HADLEY, D., LIU, R., GLESSNER, J., GRANT, S. F., HAKONARSON, H. & BUCAN, M. 2007. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome research,* 17**,** 1665-1674.

WANG, M. D., DZAMA, K., HEFER, C. A. & MUCHADEYI, F. C. 2015. Genomic population structure and prevalence of copy number variations in South African Nguni cattle. *BMC genomics,* 16**,** 1-16.

WEIGAND, H. & LEESE, F. 2018. Detecting signatures of positive selection in non-model species using genomic data. *Zoological Journal of the Linnean Society,* 184**,** 528-583.

WITTKE-THOMPSON, J. K., PLUZHNIKOV, A. & COX, N. J. 2005. Rational inferences about departures from Hardy-Weinberg equilibrium. *American journal of human genetics,* 76**,** 967-986.

YANG, L., XU, L., ZHOU, Y., LIU, M., WANG, L., KIJAS, J. W., ZHANG, H., LI, L. & LIU, G. E. 2018. Diversity of copy number variation in a worldwide population of sheep. *Genomics,* 110**,** 143-148.

YAVAŞ, G., KOYUTÜRK, M., ÖZSOYOĞLU, M., GOULD, M. P. & LAFRAMBOISE, T. 2009. An optimization framework for unsupervised identification of rare copy number variation from SNP array data. *Genome Biology,* 10**,** R119.

YURCHENKO, A. A., DAETWYLER, H. D., YUDIN, N., SCHNABEL, R. D., VANDER JAGT, C. J., SOLOSHENKO, V., LHASARANOV, B., POPOV, R., TAYLOR, J. F. & LARKIN, D. M. 2018. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. *Scientific Reports,* 8**,** 12984.

ZEDER, M. A. & HESSE, B. 2000. The Initial Domestication of Goats (Capra hircus) in the Zagros Mountains 10,000 Years Ago. *Science,* 287**,** 2254-2257.

ZHAO, F., MCPARLAND, S., KEARNEY, F., DU, L. & BERRY, D. P. 2015. Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genetics Selection Evolution,* 47**,** 49.

Chapter 3: The population structure of South African indigenous meat-type goats and inferences on its relationship with the variation of coat colour and coat colour patterns.

Abstract

Advances in molecular technologies and the recent availability of a complete, precise goat genome reference, now allow such studying of genetic diversity, population structure, and the detection of candidate genes with strategies that detect genetic markers, such as SNPs that are linked with economically important production traits. The genetic description of these populations and the identification of candidate genes that are responsible for a substantial amount of genetic variation of economical traits which influence physiological processes and phenotype expression can help open opportunities towards implementing improved selection schemes with improved accuracy, selection intensity, allowing the early selection of reproducers. Therefore, this chapter aims were to use SNP marker information to study the South African meat-type goats' population structure and infer its relationship with coat colour distribution. The dataset used in this study's computational analysis contained a total of 51 676 autosomal SNPs and 329 South African meat-type goats including the three commercial breeds such as the Boer (n = 97), Savanna (n = 38), and Kalahari Red (n = 52), as well as village goats with various coat colours and patterns including plain coats such as white (n = 25), red (n = 17), black (n = 19), grey (n = 10), and coat colour patterns such as speckled (n = 15), patchy (n = 15), belted (n = 16), white-sided (n = 14), and black-legs (n = 11). The clustering of the commercial breeds revealed by the PCA and Admixture plots suggested close genetic relatedness between the breeds. In contrast to their commercial counterparts, the ecotypes' clustering suggested higher levels of genetic diversity in the village goats. Furthermore, genetic variation within the populations (91.44%) was higher than among populations (4.77%), and small average F-statistics detected in all breeds indicated small genetic differentiation between population likely due to interpopulation gene flow. While genetic differences between these populations were likely the result of village goat adaptation to differences in climate. These finding also suggested that the coat colour distributions within the ecotype populations is likely due to the goats' adaptation to the different climatic conditions in these regions, as well as breeder preferences.

Keywords: genetic diversity, population structure, genetic SNP markers, coat colour

3.1 Introduction

Small ruminants such as goats are a vital part of the economy, especially that of farmers in rural communities, where people have successfully bred these animals for sustenance despite the limiting conditions found in extensive production systems which offer little protection from harsh environmental conditions and limited feed (Al-Dawood, 2017). South African meat-type ecotypes, which are primarily found in communal production systems have managed to strive under such adverse conditions due to their high genetic diversity and adaptability to a wide range of environments. Due to these characteristics, the ecotypes are economically important as a livestock species because they hold high production potential, especially in the face of drastically changing environmental conditions caused by global warming and an increase in meat demand due to the rising human population (Monau et al., 2020).

However, the lack of both information about proper breeding strategies, and investment capital within communal goat production systems inhibits the adoption of modern technologies, necessary to implement improved breeding schemes that utilise genetic, and phenotypic information to employ genomic selection. Thus, it is necessary to generate low-cost and easy-adoption selection technologies according to low-input communal production system requirements (Torres-Hernández et al., 2022; Torres-Hernández et al., 2020). A key step towards this goal is the description of the genetic composition of village goat populations. The genetic description of these populations is vital as it offers the opportunity to understand the goats' evolutionary history, and key points in their domestication that have resulted in their high genetic diversity and adaptability (Malan, 2000). To this extent, several studies have made efforts to describe South African goat populations' genetic diversity and identify numerous major genes linked to their reproductive, disease, and production traits (Ncube et al., 2020; Chokoe et al., 2020; Monau et al., 2018; Colli et al., 2018; Onzima et al., 2018; Mdladla et al., 2017; Mdladla et al., 2016).

The genetic description of these populations and the identification of candidate genes that are responsible for a substantial amount of genetic variation of economical traits which influence physiological processes and phenotype expression (Rajaganapathy et al., 2018) can help open opportunities towards implementing improved selection schemes with improved accuracy, selection intensity, allowing the early selection of reproducers (Monau et al., 2020; Torres-Hernández et al., 2020; Sevane et al., 2018). Advances in molecular technologies and the recent availability of a complete, precise goat genome reference, now allow such studying of genetic

diversity, population structure, and the detection of candidate genes with strategies that detect genetic markers, such as SNPs that are linked with economically important production traits (Gandini et al., 2017). Therefore, this chapter aims to use SNP marker information to study the South African meat-type goats' population structure and infer its relationship with coat colour distribution.

3.2 Materials and Methods

3.2.1 Genotypic data

The genotypic data used in this study was generated by combining six separate datasets. Two of these datasets were generated from samples retrieved from previous studies which genotyped the South African meat-type goat populations including the three commercial breeds of the Boer, Savanna and the Kalahari Red, as well as non-descript village goat populations using the Illumina Goat 50K SNP BeadChip. Samples in Dataset 1, which contained 128 genotyped samples, were retrieved from the study by (Mdladla et al., 2016), while those in Dataset two, which contained 5 genotyped samples, were retrieved from the study by (Ncube et al., 2020).

Datasets three to six were generated from archival goat blood samples found in the Agricultural Research Council-Biotechnology Platform in Onderstepoort (Pretoria, South Africa). From these blood samples, genomic DNA was extracted using the DNeasy" Blood and Tissue Kit (Qiagen), as per the manufacturer's instructions with a slight modification of increased lysis time to 90 min. DNA quality and quantity were determined using Qubit" 3.0 Fluorometer (Life Technologies). The samples were genotyped with the Illumina Goat 65K SNP genotyping Bead-chip using the Infinium assay at the ARC-BTP. SNP genotypes were called using genotyping module integrated in GenomeStudio ™ v2.05 (Illumina, Inc.). The SNP marker map files for each dataset was updated using the Illumina Goat 65K SNP genotyping bead-chip's genomic marker map in Golden Helix SNP and VARIATION SUITE (SVS) ™ version 8.9.1 (Golden Helix, Inc., 2020, Bozeman, MT, www.goldenhelix.com)

The goats used in this study represent the major meat type goat breeds and populations reared in South Africa as the samples of the three commercial goats were sampled from from various stud breeders and commercial farms around the country, while the non-descript local ecotype populations were sampled from smallholder farmers from villages in the Eastern Cape (Xhosa), KwaZulu-Natal (Zulu), Limpopo (Venda) and North-West (Tswana) provinces (Table 3.1).

3.2.2 Data quality control and merging

The genotype data was filtered using PLINK v1.9 (Purcell et al., 2007) in R (R Core Team, 2023). Quality control was carried for each dataset at a sample and SNP level to remove SNPs with call rates lower than 0.05 (MIND = 0.05). At the sample level, QC filtered out animals with genotyping rates less than 0.05 (GENO = 0.05), and those with minor allele frequencies lower than 0.05 (MAF < 0.05).

Dataset one:

Contained 51 854 SNPs and 128 goats before QC. Four goats were removed due to missing genotype data (mind > 0.05). Total genotyping rate in remaining samples was 0.991976. A total of 503 SNPs were removed due to missing genotype data (geno > 0.05) and minor allele frequencies (maf < 0.05). While 410 SNPs were removed due to the Hardy-Weinberg exact test (hwe: 0.000001). After QC, 50 941 SNPs and 124 goats remained.

Dataset two:

Contained 47 682 SNPs and 5 goats before QC. No goats were removed due to missing genotype data (mind > 0.05). Total genotyping rate was 0.987517. A total of 2686 SNPs were removed due to missing genotype data (geno > 0.05) and minor allele frequencies (maf < 0.05). While no SNPs were removed due to the Hardy-Weinberg exact test (hwe = 0.000001). After QC, 44 996 SNPs and 5 goats remained.

Dataset three:

Contained 59 100 SNPs and 128 goats before QC. Three goats were removed due to missing genotype data (mind > 0.05). Total genotyping rate in remaining samples was 0.984075. A total of 1 232 SNPs were removed due to missing genotype data (geno > 0.05) and minor allele frequencies (maf < 0.05). While 543 SNPs were removed due to the Hardy-Weinberg exact test (hwe = 0.000001). After QC, 57 325 SNPs and 125 goats remained.

Dataset four:

Contained 59 100 SNPs 20 goats before QC. No goats were removed due to missing genotype data (mind > 0.05). Total genotyping rate was 0.981325. A total of 1 238 SNPs were removed due to missing genotype data (geno > 0.05) and minor allele frequencies (maf < 0.05). While six SNPs were removed due to the Hardy-Weinberg exact test (hwe = 0.000001). After QC, 57 856 variants and 20 goats remained.

Dataset five:

Contained 59 100 variants and 16 goats before QC. One goat was removed due to missing genotype data (mind > 0.05). Total genotyping rate in remaining samples was 0.982732. A total of 984 SNPs were removed due to missing genotype data (geno > 0.05) and minor allele frequencies (maf < 0.05). While no SNPS were removed due to the Hardy-Weinberg exact test (hwe:0.000001). After QC, 58 116 SNPs and 15 goats remained.

Dataset six:

Contained 59 100 variants and 40 goats before QC. No goats were removed due to missing genotype data (mind > 0.05). Total genotyping rate was 0.987728. A total of 899 SNPs were removed due to missing genotype data (geno > 0.05) and minor allele frequencies (maf < 0.05). While 80 SNPs were removed due to the Hardy-Weinberg exact test (hwe:0.000001). After QC, 58 121 SNPs and 40 goats remained.

Merged dataset:

The six datasets were merged using PLINK v1.9 (Purcell et al. 2007) such that the combined dataset contained a total of 58 656 SNPs and 329 goats. However, an additional 1050 duplicated variants were removed. Another 3 577 SNPs in LD were removed using the indep-pairwise 50 5 0.8 filter. From the remaining 54 029 variants, 2 353 non-autosomal SNPs and those on contigs or not mapped to the latest reference assembly of the goat genome were removed. Therefore, the dataset used in this study's computational analysis contained a total of 51 676 autosomal SNPs and 329 South African meat-type goats with various coat colours and patterns including plain coats such as white, red, black, grey, and coat colour patterns such as speckled, patchy, belted, white-sided, and black-legs (Table 3.2). Other coat colour and coat colour patterns were considered for the ecotypes, however small sample sizes and high variation of coat colour patterns prevented their inclusion in this study.

Table 3.1: South African indigenous meat-type goat populations used in the study.

Breeds/Populations	Number of animals
Boer	97
Kalahari Red	52
Savanna	38
North-West (Village: Tswana)	30
Limpopo (Village: Venda)	33
Eastern Cape (Village: Xhosa)	34
KwaZulu-Natal (Village: Zulu)	45
Total no. of animals in combined dataset	329

Table 3.2: Coat colour and coat colour patterns of commercial and village meat-type goats.

Coat Colour/Pattern	Tswana	Xhosa	Venda	Zulu	Total for coat colour/coat colour pattern
White	15	10	-	-	25
Red	4	5	4	4	17
Black	-	-	11	8	19
Grey	-	-	3	7	10
Belted	5	2	2	7	16
Patchy	-	15	-	-	15
Speckled	6	2	1	6	15
White sided	-	-	8	6	14
Blacklegs	-	-	5	6	11
Boer	-	-	-	-	97
Savanna	-	-	-	-	38
Kalahari Red	-	-	-	-	52
Overall total no. of animals	**30**	**34**	**34**	**44**	**329**

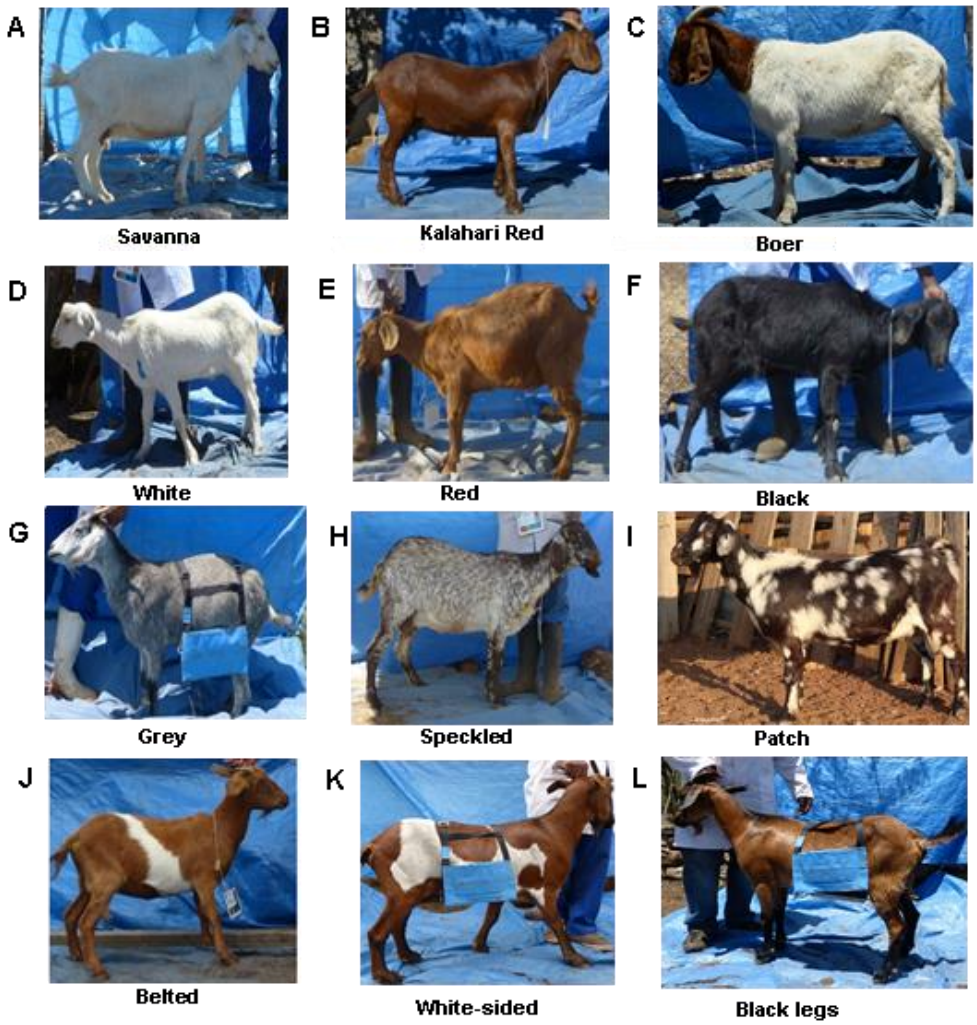

Figure 3.1: The different goat coat colours and patterns under investigation in the study. The commercial breeds, including the A) Savanna (n = 38), B) Kalahari Red (n =52), and C) Boer (n = 97) breeds were of their standard coat colours of white, red, and white body red head, respectively. Ecotypes were of various coat colours and patterns such as D) white (n = 25), E) red (n = 17), F) black (19), G) grey (n = 10), H) speckled (n = 15), I) patchy (n = 15), J) belted (n = 16), K) white-sided (n = 14), and L) black legs (11).

3.2.3 Principal component analysis

A principal component analysis of the quality-controlled data was carried out using PLINK (Purcell et al., 2007) in R (R Core Team, 2023). This was done using the R package adegenet (Jombart T, 2020) to assess the clustering of the populations.

3.2.4 Admixture

An admixture analysis was carried out for K = 1 – 20 in R (R Core Team, 2023) using the snmf function of the R package LEA (Frichot., 2015) and visualized in R studio to assess the goats' admixture proportions and possible ancestral populations.

3.2.5 AMOVA and Pairwise FST

PGDSpider v.2.1.1.5 software (Lischer and Excoffier, 2012) was used to convert PLINK files to the appropriate Arlequin formats. A locus-by-locus analysis of molecular variance (AMOVA) was then performed using ARLEQUIN v.3.5.2 (Lischer and Excoffier, 2010), with 1,000 permutations to assess the divergence among breeds and populations. Genetic differentiation among breed (FST) fixation indices were calculated using 20,000 permutations and a significance level of 0.05.

3.3 Results

3.3.1 Principal Component Analysis

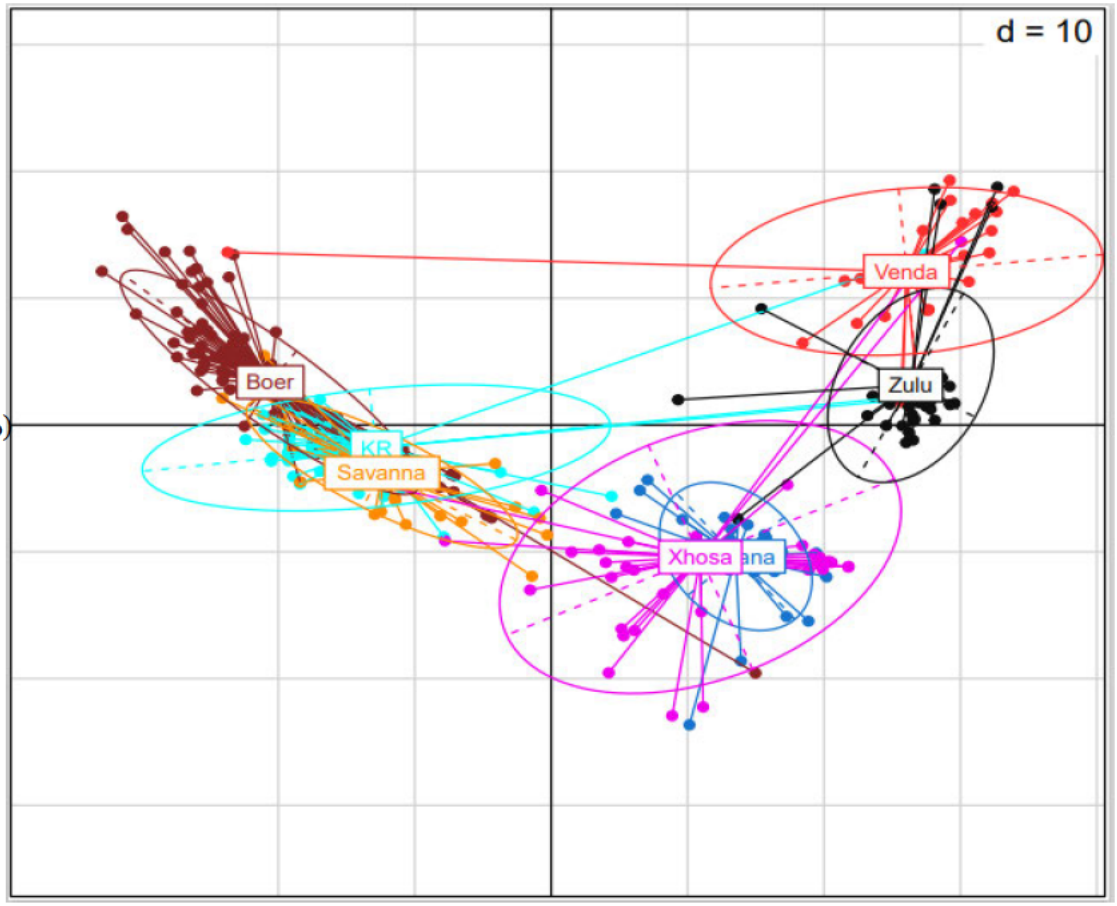

Figure 3.2: Principal component analysis (PCA) of the three commercial meat-type breeds and ecotype populations. The three meat-type breeds consisting of the Boer, Kalahari Red, and the Savanna cluster together separate from the four ecotypes, which included Xhosa, Zulu, Venda, and Tswana goat populations. The first and the second principal components explained 7.85% and 1.27% of the variation, respectively.

Population structure analysis carried out in this study provided evidence that the commercial meat-type breeds and the ecotypes are genetically distinct from each other likely due to their different production systems. A further separation within the ecotypes was revealed by PCs 1 and 2 which explained 7.58 and 1.27% of the variation, respectively. Although geographically separated, Tswana and Xhosa goats displayed some levels of relatedness clustering separately from the Zulu and Venda goats which formed a cluster of their own. The clustering is reflective of similarities in production systems of the different regions and the populations' historical dispersal across the country due to the migrations and settlements of the ethnic tribes that rear them. As such the higher levels of genetic diversity observed in the ecotype populations likely resulted from formal and informal trading networks between the different ethnic tribes, and adaptation to local production systems and environments.

3.3.2 Admixture analysis

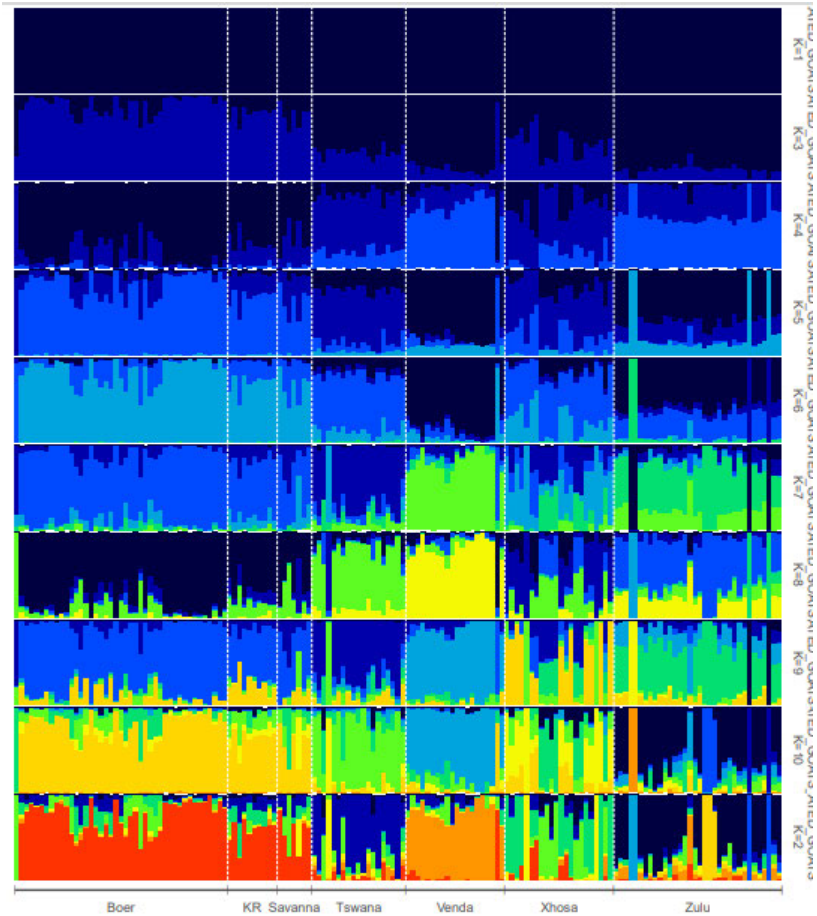

Figure 3.3: Admixture analysis for K = 1 - 10 of the three commercial meat-type breeds and ecotype populations. Optimum K value was 5. The admixture analysis of the South African goat populations included the three commercial meat-type populations of the Boer, Savanna and the Kalahari Red (KR) breeds, and four ecotype populations.

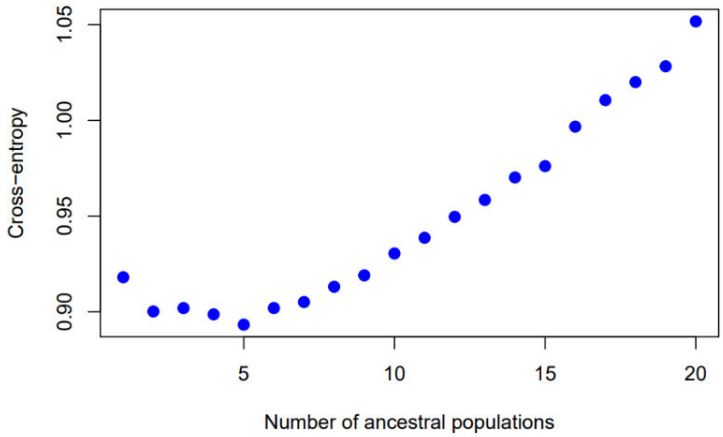

Figure 3.4: Cross validation error for optimal K value. The Cross-validation error was lowest at K = 5. This suggests that the optimal number of ancestral populations is 5.

Supporting the PCA which revealed a clear separation between the commercial meat-type breeds and the ecotypes, the genetic backgrounds revealed by the Admixture analysis (at K = 5) suggested a history of shared ancestry between the commercial breeds. The similar genetic backgrounds observed between the three breeds are reflective of the breeds' history of intense selection and separation into different populations from their indigenous unimproved ancestors. Admixture revealed subpopulation in the village goats, similar to the clustering revealed by the PCA, reflecting the different geographic locations and different ethnic groups that rear the goats.

3.3.3 AMOVA and pairwise-FST analysis

Table 3.3: Analysis of molecular variance (AMOVA) of the three commercial meat-type breeds and ecotype populations with different coat colours and patterns.

Source of variation	Sum of squares	Variance components	Percentage variation
Among populations	398296.010	517.88948	4.77341
Among individuals within populations	3299108.567	410.62110	3.78471
Within individuals	3166540.000	9920.95351	91.44188
Total	6863944.577	10998.23295	

Table 3.4: Pairwise FST comparison between the three commercial meat-type breeds and ecotype populations with different coat colours and patterns.

	Belted	Black	Blacklegs	Boer	Grey	Kalahari Red	Patchy	Red	Savanna	Speckled	White	White-sided
Belted	0											
Black	0.00762	0										
Blacklegs	0.00620	0.00244	0									
Boer	0.08449	0.09781	0.11091	0								
Grey	0.00368	0.00184	0.00032	0.10253	0							
Kalahari Red	0.05868	0.07195	0.08154	0.01063	0.07410	0						
Patchy	0.01773	0.02769	0.02906	0.05993	0.02608	0.03904	0					
Red	0.00110	0.00687	0.00921	0.07515	0.00513	0.05065	0.00979	0				
Savanna	0.06192	0.07729	0.08694	0.01249	0.07891	0.00899	0.04218	0.05385	0			
Speckled	0.00333	0.01081	0.00799	0.08779	0.00256	0.06045	0.01591	0.00334	0.06479	0		
White	0.00861	0.02297	0.02400	0.06088	0.01984	0.03922	0.01402	0.00837	0.04099	0.00684	0	
White-sided	0.00821	0.00046	0.00251	0.10624	0.00270	0.07866	0.03135	0.01221	0.08413	0.01295	0.026411	0

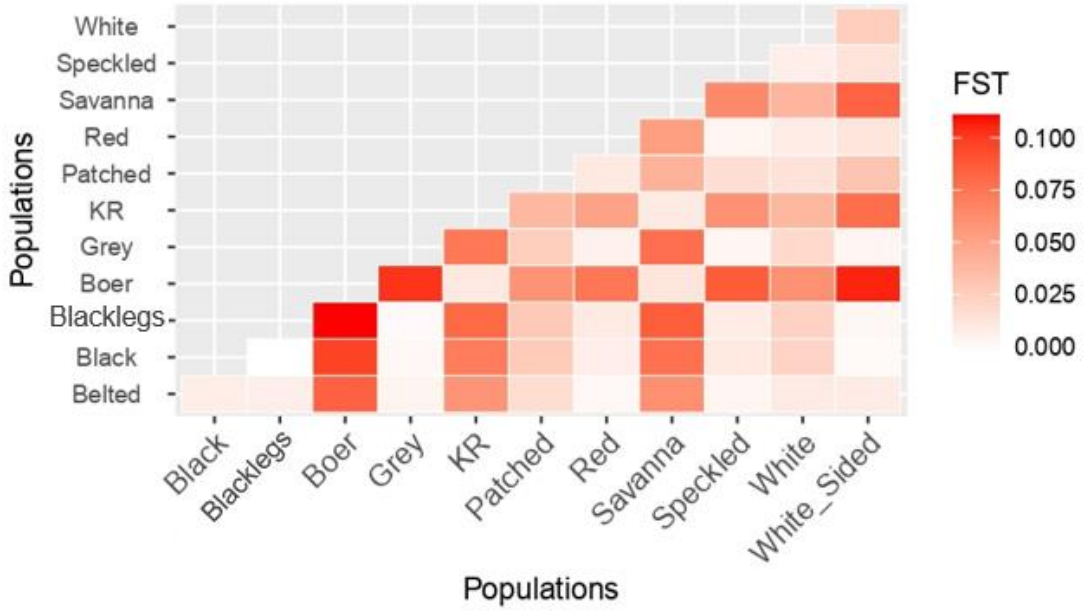

Figure 3.5: Pairwise FST values from the three commercial meat-type breeds and ecotype populations. FST values range from 0 to 1 where, low values are indicated by white and high values are indicated by red. FST values could indicate small (0–0.05), medium (0.05–0.15), high (0.15–0.25) and very high (FST > 0.25) differentiation between populations. Low to moderate FST values were detected between the commercial breeds and the ecotype populations, while low FST values were detected among the commercial breeds and the ecotype populations.

The AMOVA and pairwise FST analysis revealed trends similar to those revealed by the PCA and Admixture. Analysis of molecular variance (AMOVA) of the three commercial meat-type breeds and ecotype populations revealed variations of 4.77% and 91.44% among populations and within individuals, respectively (Table 3.3). Low FST values were detected among the commercial breeds and ecotype population, while low to moderate FST values were detected between the commercial breeds and the ecotype populations. The lowest pairwise FST value (0.00046) was between the white-sided and black village goats, while the highest pairwise FST value (0.11091) was between the Boer breed and village goats with the blacklegs/feet coat colour pattern (Table 3.4). The low FST values among commercial breeds (FST ranging from 0.00899 – 0.01249) and among the ecotype populations (FST ranging from 0.00046 – 0.03135) imply close relationships. While low to moderate genetic differentiation (range FST = 0.03904 - 0.11091) was observed between the commercial breeds and the ecotype populations (Figure 3.4). Furthermore, F-statistics including FIS (within-population inbreeding estimate), FIT (total inbreeding estimate) and FST (estimate of population differentiation) were used to

determine the inbreeding level in all the populations and for genetic differentiation among populations. Mean values of FIS, FIT and FST for all loci were 0.03974, 0.04773. and 0.08558, respectively. Therefore, genetic variation within the populations was higher than among populations, furthermore, the small average F-statistics detected in all the groups indicated small genetic differentiation between populations suggesting interpopulation gene flow.

3.4 Discussion

A clear genetic distinction between the three commercial meat-type breeds (i.e., Boer, Kalahari Red, and Savanna) and the four ecotypes (Xhosa, Zulu, Tswana, and Venda) was revealed by the PCA. This result was expected and is in line with the findings of previous studies (Chokoe et al., 2020; Mdladla et al., 2016), which demonstrated a similar genetic distinction between these populations. Goats from the commercial breeds clustered together suggesting low genetic diversity between the breeds. It's been proposed that this pattern is due to high levels of inbreeding between these breeds because of undergoing selective breeding programs for improvement (Chokoe et al., 2020; Mdladla et al., 2016). While goats from the ecotypes formed two separate clusters, suggesting a higher level of genetic diversity between the ecotypes compared to commercial breeds likely due to indiscriminate crossbreeding in communal production systems.

The commercial breeds' clustering can be explained by their breeding histories and the current breeding practises of their production systems. Literature on the history of the commercial breeds has reported that the Boer and the Savanna breeds were developed from indigenous unimproved goat populations in the Eastern Cape and Northern Cape respectively (Campbell, 2003). While the history of the Kalahari Red reported that the breed was developed from the redheaded Boer goat and the brown lop-eared unimproved local goats from South Africa and Namibia (Campbell, 2003).

This breed history suggests that the three breeds share similar genetic backgrounds, having been developed from the selection and goal-oriented crossbreeding of unimproved indigenous goats (Campbell, 2003). Thus, the clustering observed in the commercial breeds suggests close genetic relatedness between the breeds, a pattern that has been revealed by previous studies. For instance, Mdladla et al. (2016), revealed signs of a reduction in effective population size and higher levels of inbreeding within the commercial breeds due to intensive artificial selection implemented in the commercial production systems, which was absent in village goat production systems. The higher levels of inbreeding were a result of intensive artificial

selection which involves the use of only pure breeds which meet the standards defined by Breeders Associations with often no room for genetic variation (Mdladla et al., 2016).

In contrast to their commercial counterparts, the ecotypes' clustering suggested higher levels of genetic diversity in the village goats. This pattern has been previously linked to village goat farming systems. In their study, which characterized the village goat production systems of South Africa, Mdladla et al. (2016) revealed that village goat breeding is often uncontrolled and that communal production systems allows for indiscriminate breeding resulting in uncontrolled crossbreeding and mixed genotypes of goats (Rumosa Gwaze et al., 2008). In addition, the higher genetic diversity observed in the ecotypes was linked to smallholder goat farmers crossbreeding the commercial breeds with their local goats in an attempt to improve meat yield in their flocks (Mdladla et al., 2016). The two separate clusters revealed by the PCA within the ecotypes suggests close relatedness between the Tswana and Xhosa populations as well as between the Zulu and the Venda populations, likely due to gene flow between the populations.

These findings are supported in literature which has revealed that the population structure within village goat populations is linked to the cultural preferences for specific coat colours and patterns among the ethnical groups keeping these animals (Morrison 2007). The clustering of the Tswana and Xhosa goats, as well as the clustering between the Venda goats and the Zulu goats revealed by the PCA has also been observed in previous studies (Mdladla et al., 2016). The low levels of genetic diversity revealed between these goat populations suggests that they share similar genetic identities. The diversity of coat colour observed within and between the different goat populations is likely reflective of adaptions to environmental stresses under the extensive scavenging-type environments of communal production systems.

Furthermore, the close relatedness between these populations is also likely a result of shared bucks between different villages due to the low number of reproductive bucks in communal production systems. Mdladla et al. (2016) reported households that did not own bucks in Limpopo and KwaZulu-Natal suggesting that some farmers used communal bucks for their flocks. In addition, Mdladla et al. (2016) found that village goat farmers often castrate young bucks that are not needed for reproduction as a strategy to control breeding and improve meat quality (Rumosa Gwaze et al., 2008). The practice of castration results in their flocks often consisting of only one or two reproducing bucks for breeding, and these animals are often shared as communal bucks (Mahanjana & Cronje 2000).

The use of communal bucks for village production systems has also been reported in other parts of Africa (Rumosa Gwaze et al., 2008) suggesting that it's a common breeding strategy used by communal farmers. Although the practice of using communal bucks has the risk of increasing the chances of close relatives mating (Rumosa Gwaze et al., 2008), communal herding, which allows breeding does and bucks to mix among different flocks has been found to minimize levels of inbreeding in the ecotype populations (Jaitner et al., 2001). Communal herding thus explains the higher levels of genetic diversity and admixture revealed by the separate clusters observed between the ecotype populations compared to the commercial breeds.

The patterns of ancestral proportions observed in this study's Admixture analysis support the clustering revealed by the PCA. The commercial goat breeds including the Boer, the Kalahari Red and the Savanna clustered together, suggesting that these populations share similar genetic identities likely due to the effects of the intense selective breeding in their production systems which have resulted in lower levels of genetic diversity (Mdladla et al., 2016). Whereas the higher genetic diversity in the ecotype populations compared to the commercial breeds is likely reflective of the absence of strong artificial selection pressures, and gene flow between different populations due to the heterogeneous characteristics of communal production systems. As revealed by Mdladla et al. (2016), communal production systems are characterised by broad and multiple breeding objectives which result in larger effective population sizes than those for the commercial breeds which are raised as closed flocks with specific breeding goals.

The locus-by-locus AMOVA results from this study showed that a large part of genetic variation in the meat-type goats is accounted for within individuals (91.44%) than among populations (4.77%). The higher value of variance within-individuals in this study is likely due to continuous gene flow between individuals from the different ecotype populations. The ecotype populations are made up of highly heterogeneous animals, as such they likely account for majority of the within individual variance compared to the among population variance. In contrast, the commercial breeds have been shown to be highly related to each other which would explain the lower levels of between population variance. Furthermore, a history of trade and migration across geographical barriers between the ethnic groups that rear the ecotype goats suggests gene flow between the different populations. Therefore, the close genetic relationships observed in this study confirm the presence of considerable gene flow among these different meat-type populations.

In addition, the AMOVA results from this study were consistent with previous findings showing that there is higher within-population genetic variance among South African indigenous goat populations (Chokoe et al., 2020). Similar results were observed in Pakistani cattle breeds (Hussain et al., 2016) and yellow cattle in Taiwan, which indicated a 4.35% proportion of genetic variation attributed to population differentiation among the studied populations (Tu et al., 2014). Furthermore, Corredor et al. (2023) observed greater the genetic variation (75.71%) within Peruvian cattle breeds with low variability among the breeds (24.29%). Therefore, the higher genetic diversity in the ecotypes is the result of gene flow between different village goat populations and likely reflects the different goats' adaptations to very different production systems. This is further supported by Mdladla et al. (2016)'s findings that the genomic variation was explained more by climate variation compared to geographical distances.

Therefore, the coat colour diversity within the ecotype populations is likely due to goat adaptation to diverse climatic conditions. Coat colour and coat colour patterns varied within and amongst the goat populations. However, majority of the darker coats were found in goats within the Zulu-Venda cluster. While majority of the lighter coats were found in goat within the Tswana-Venda cluster. SA's diverse village goat populations are raised in extreme and severe climates and environmental conditions. However, the goats are well adapted to these conditions and have been reported to be superior in managing heat stress due to their thermoregulatory response mechanisms (Mdladla et al., 2016). In this sense, heat stress is an important selection pressure that has shaped the genome of the indigenous goats. Therefore, it is believed that the coat colour diversity observed in these local goat populations reflects their adaptation to the diverse climates and environments under which they are reared (Hagan et al., 2012; Yakubu et al., 2010a,b; Daramola & Adeloye, 2009).

Majority of the Zulu goats occur in the arid and subtropical wet areas of KZN, while Venda goats are found in the hot and drier desert-like areas of the Northern Cape (Mdladla et al., 2016; Morrison, 2007). This suggests that the village goats in the Zulu-Venda cluster represented by the PCA were mostly black because of their adaptations to extreme climates. Similar results were found in Ghanan goats, where dark coat colours were linked to environmental adaptation (Hagan et al. 2012) as they help the goats absorb more heat compared to other coat colours.

In addition, coat colour has important socio-cultural and economic value in veld goat communities, as a result, most farmers select their goats based on their coat colours and patterns

(Rumosa Gwaze et al. 2008; Mahanjana & Cronje 2000). For instance, Mahanjana and Cronje (2000) found that white coloured goats are in higher demand in the Eastern Cape due to their perceived value in traditional/cultural ceremonies, resulting in selection preferences towards white goats. Therefore the prevalence of specific coat colours at specific regions in the country (i.e., most white goats were found in the Tswana and Xhosa populations which were sampled from the North-West, and Eastern Cape, while majority of the black goats were found in Zulu and Venda populations which were sampled from KwaZulu-Natal, and Limpopo), is due to the goats' adaptation to the different climatic conditions in these regions, as well as the selective breeding practices of goat farmers which are informed by their socio-cultural and religious preferences. Similar reports have been indicated in other African agro-pastoral production systems, in which the selection of bucks is made by appearance causing the distribution of colour variations within local populations to be linked to the preferences of farmers and goat adaptability (Mataveia et al., 2021).

3.5 References

AL-DAWOOD, A. 2017. Towards heat stress management in small ruminants–a review. *Annals of Animal Science,* 17**,** 59-88.

CAMPBELL, Q. 2003. The origin and description of southern Africa's indigenous goats. *S Afr J Anim Sci,* 4**,** 18-22.

CHOKOE, T. C., MDLADLA-HADEBE, K., MUCHADEYI, F., DZOMBA, E., MATELELE, T., MPHAHLELE, T., MPOFU, T. J., NEPHAWE, K. & MTILENI, B. 2020. Genetic Diversity of South African Indigenous Goat Population from Four Provinces Using Genome-Wide SNP Data. *Sustainability,* 12**,** 10361.

COLLI, L., MILANESI, M., TALENTI, A., BERTOLINI, F., CHEN, M., CRISÀ, A., DALY, K. G., DEL CORVO, M., GULDBRANDTSEN, B. & LENSTRA, J. A. 2018. Genome-wide SNP profiling of worldwide goat populations reveals strong partitioning of diversity and highlights post-domestication migration routes. *Genetics Selection Evolution,* 50**,** 1-20.

CORREDOR, F.-A., FIGUEROA, D., ESTRADA, R., SALAZAR, W., QUILCATE, C., VÁSQUEZ, H. V., GONZALES, J., MAICELO, J. L., MEDINA, P. & ARBIZU, C. I. 2023. Genetic diversity and population structure of a Peruvian cattle herd using SNP data. *Frontiers in Genetics,* 14**,** 1073843.

DARAMOLA, J. O. & ADELOYE, A. A. 2009. Physiological adaptation to the humid tropics with special reference to the West African Dwarf (WAD) goat. *Trop Anim Health Prod,* 41**,** 1005-16.

FRICHOT, E. & FRANÇOIS, O. 2015. LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution,* 6**,** 925-929.

GANDINI, G., TURRI, F., RIZZI, R., CROTTA, M., MINOZZI, G. & PIZZI, F. 2017. Economic evaluation of genetic improvement in local breeds: The case of the Verzaschese goat. *Italian Journal of Animal Science,* 16**,** 1-12.

HAGAN, J., APORI, S., BOSOMPEM, M., ANKOBEA, G., MAWULI, A. & AL, H. 2012. Morphological characteristics of indigenous goats in the coastal savannah and forest eco-zones of Ghana. *J. Anim. Sci. Adv,* 2**,** 813-821.

JAITNER, J., SOWE, J., SECKA-NJIE, E. & DEMPFLE, L. 2001. Ownership pattern and management practices of small ruminants in The Gambia—implications for a breeding programme. *Small ruminant research,* 40**,** 101-108.

JOMBART, T., KAMVAR, Z. N., COLLINS, C., LUŠTRIK, R., BEUGIN, M.-P., KNAUS, B. J., SOLYMOS, P., MIKRYUKOV, V., SCHLIEP, K. & MAIÉ, T. 2020. Adegenet: exploratory analysis of genetic and genomic data.

LISCHER, H. E. L. & EXCOFFIER, L. 2011. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics,* 28**,** 298-299.

MAHANJANA, A. & CRONJE, P. 2000. Factors affecting goat production in a communal farming system in the Eastern Cape region of South Africa. *South African Journal of Animal Science,* 30**,** 149-155.
MALAN, S. W. 2000. The improved Boer goat. *Small Rumin Res,* 36**,** 165-170.
MATAVEIA, G. A., VISSER, C. & SITOE, A. 2021. *Smallholder goat production in Southern Africa: A review*, IntechOpen.
MDLADLA, K., DZOMBA, E. F., HUSON, H. J. & MUCHADEYI, F. C. 2016. Population genomic structure and linkage disequilibrium analysis of South African goat breeds using genome-wide SNP data. *Anim Genet,* 47**,** 471-82.
MDLADLA, K., DZOMBA, E. F. & MUCHADEYI, F. C. 2017. Characterization of the village goat production systems in the rural communities of the Eastern Cape, KwaZulu-Natal, Limpopo and North West Provinces of South Africa. *Tropical Animal Health and Production,* 49**,** 515-527.
MONAU, P. I., VISSER, C., MUCHADEYI, F. C., OKPEKU, M., NSOSO, S. J. & VAN MARLE-KÖSTER, E. 2020. Population structure of indigenous southern African goats based on the Illumina Goat50K SNP panel. *Trop Anim Health Prod,* 52**,** 1795-1802.
MONAU, P. I., VISSER, C., NSOSO, S. & VAN MARLE-KÖSTER, E. 2018. Phenotypic and genetic characterization of indigenous Tswana goats. *South African Journal of Animal Science,* 48**,** 925-934.
MORRISON, J. 2007. A guide to the identification of the natural african indigenous veld goats of southern africa. *Compiled for the Indigenous Veld Goat Breeders Society, Pretoria, South Africa*.
NCUBE, K. T., HADEBE, K., DZOMBA, E. F., SOMA, P., FRYLINCK, L. & MUCHADEYI, F. C. 2020. Relationship between population genomic structure and growth profiles of South African goats under different production systems. *Trop Anim Health Prod,* 52**,** 1277-1286.
ONZIMA, R., UPADHYAY, M., MUKIIBI, R., KANIS, E., GROENEN, M. & CROOIJMANS, R. 2018. Genome-wide population structure and admixture analysis reveals weak differentiation among Ugandan goat breeds. *Animal genetics,* 49**,** 59-70.
PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M. A., BENDER, D., MALLER, J., SKLAR, P., DE BAKKER, P. I. & DALY, M. J. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics,* 81**,** 559-575.
RAJAGANAPATHY, J., JAYASHANKAR, M., NAGARAJA, C., SHRIKRISHNA, I. & SATYANARAYANA, K. 2018. Genetic characterization of local goats of Karnataka by microsatellite marker analysis. *Indian Journal of Animal Research*.
R CORE TEAM. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
RUMOSA GWAZE, F., CHIMONYO, M. & DZAMA, K. 2008. Communal goat production in Southern Africa: a review. *Tropical Animal Health and Production,* 41**,** 1157.
SEVANE, N., CORTÉS, O., GAMA, L. T., MARTÍNEZ, A., ZARAGOZA, P., AMILLS, M., BEDOTTI, D. O., BRUNO DE SOUSA, C., CAÑON, J., DUNNER, S., GINJA, C., LANARI, M. R., LANDI, V., SPONENBERG, P. & DELGADO, J. V. 2018. Dissection of ancestral genetic contributions to Creole goat populations. *Animal,* 12**,** 2017-2026.
TORRES-HERNÁNDEZ, G., MALDONADO-JÁQUEZ, J. A., GRANADOS-RIVERA, L. D., WURZINGER, M. & CRUZ-TAMAYO, A. 2022. Creole goats in Latin America and the Caribbean: a priceless resource to ensure the well-being of rural communities. *International Journal of Agricultural Sustainability,* 20**,** 368-380.
TORRES-HERNÁNDEZ, G., VAQUERA HUERTA, H., MALDONADO-JÁQUEZ, J. A., SALINAS, H., PASTOR, F., GRANADOS-RIVERA, L. D., VAQUERA, H. & BAUTISTA, Y. 2020. Influence of milk production records on the estimation of typical and extended curve lactation parameters in local goats under grazing management. *The Indian journal of animal sciences,* 90**,** 643-646.
YAKUBU, A., RAJI, A. & OMEJE, J. 2010a. Genetic and phenotypic differentiation of qualitative traits in Nigerian indigenous goat and sheep populations. *Journal of Agricultural and Biological science,* 5**,** 58-66.
YAKUBU, A., SALAKO, A., IMUMORIN, I., IGE, A. & KINYEMI, M. 2010b. Discriminant analysis of morphometric differentiation in the West African Dwarf and Red Sokoto goats. *South African Journal of Animal Science,* 40**,** 381-387.

Chapter 4: Genome-wide association studies and copy-number variation analysis reveal genes for coat colour and coat colour patterns in South African indigenous meat-types goats.

Abstract

Genome-wide association studies have become routine in the mapping of loci that influence various traits considered to be economically important and have been successful in several livestock species. Production traits that have been targets of such studies include meat quantity and quality, milk yield, percentage of fat and protein, egg production, reproduction traits, and coat colour. The detection of genes responsible for economical traits such as coat colour can help open opportunities towards implementing improved selection schemes with improved selection accuracy and intensity allowing for the early selection of reproducers within communal goat production systems. This study used 51 767 SNPs from 329 indigenous goats with various coat colour and coat colour patterns to investigate genes responsible for goat coat colour. South African meat-type goats used in the study included the three commercial breeds: Boer (n = 97), Savanna (n = 38), and Kalahari Red (n = 52), as well as village goats with various coat colours and patterns including plain coats such as white (n = 25), red (n = 17), black (n = 19), grey (n = 10), and coat colour patterns such as speckled (n = 15), patchy (n = 15), belted (n = 16), white-sided (n = 14), and black-legs (n = 11). Case/control GWASs carried out for all the coat colours/patterns using the GAPIT revealed markers associated to several coat colours genes linked to coat colours such as white (*CDK5*), red (*CELF5*, *TLE6*), black (*CACNA2D1*), grey (*GSK3B*), and coat colour patterns such as white-body red head (*CADPS2*, *SLC13A1*), speckled (*KIT*, *TYRP1*), patchy (*GNAI3*), belted (*TYRP1*), white-sided (*AHCY*), and blacklegged (*AHCY*). Golden Helix SVS's univariate method from the copy number analysis module (CNAM) detected 2 047 CNVs and 279 CNVRs overlapping 5 222 genes involved in several biological processes, molecular functions, cellular components, and pathways. Several coat colour genes overlapping these CNVRs included *GSK3B, Notch1, Notch2, CDK5, ADAMTS20, TYRO3, MAP2K1, ITCH, ASIP, AHCY, SLC45A3, EDNRA, ADCY2,* and *TYR.* The study's findings need to be validated on larger sample sizes.

Keywords: Genome-wide association studies, Copy number variation, goats, coat colour, improved selection schemes

4.1 Introduction

The advent of high-density SNP Beadchips for goats have provided the opportunity for the identification of genomic regions related to the selection history of worldwide renowned goat breeds by genome-wide association studies (Brito et al., 2017; Benjelloun et al., 2015). GWASs have become routine in the mapping of loci that influence various traits considered to be economically important and have been successful in several livestock species (Martin, 2016; Matukumalli et al., 2009). Production traits that have been targets of such studies include meat quantity and quality, milk yield, percentage of fat and protein, egg production, reproduction traits, and coat colour.

Coat colour is an important production trait that is vital for breed identification, classification, and characterisation. In livestock production, variations in coat colour might be due to both the selection choices of breeders and environmental adaptations to different production systems and geographical regions (Mdladla et al., 2016). Consequently, selection carried out in the commercial goat sector and the indiscriminate breeding practices of the communal production systems have led to goat breeds with various coat colours and coat colour patterns, with higher coat variations seen in the communal systems due to the absence of strong artificial selection pressures (Mdladla et al., 2016). In addition, coat colour is linked to other livestock productivity traits including heat tolerance (Peters et al., 1980), disease resistance and susceptibility, and environmental adaptation (Mapholi et al., 2015). Therefore, identification of genomic regions underlying the variations in the trait is quite valuable for selection purposes.

While coat colour has been studied in different livestock species, genome-wide studies for goat coat colour have been limited to commercial goat breeds due to their high productivity as a result of intensive production systems (Stella et al., 2018; Nazari-Ghadikolaei et al., 2018). Several studies of the genetic diversity of South African indigenous goat breeds have been performed (Ncube et al., 2020; Monau et al., 2018; Colli et al., 2018; Onzima et al., 2018; Mdladla et al., 2018, Mdladla et al., 2017; Mdladla et al., 2016) due to their characteristic genetic resources and because the smallholder communal production systems contribute over 60% of goats in the country. However, few genome-wide association studies investigating the genetic mechanisms underlying economically important production traits such as coat colour have been carried out. This knowledge gap presents a major hinderance towards establishing improved breeding schemes in communal production systems where the indigenous breeds are at risk of erosion due to indiscriminate cross breeding and increasing levels of inbreeding

(Ncube et al., 2020; Monau et al., 2018; Colli et al., 2018; Onzima et al., 2018; Mdladla et al., 2018, Mdladla et al., 2017; Mdladla et al., 2016). Therefore, identifying the candidate genes responsible for coat colour will improve the efficiency of goat production in communal production systems through the development and implementation of improved selection and breeding programmes. In this regard, this chapter's first objective was to investigate genes associated with coat colours and coat colour patterns in South African indigenous meat-type breeds using genome-wide association studies.

Copy number variations (CNV) are genomic changes that encompass genomic inversions, translocations, duplications and deletions of various sizes, which display variable copy number in individuals when they are compared with a reference genome (Di Gerlando et al., 2019; Di Gerlando et al., 2018; Zarrei et al., 2015; Bickhart and Liu, 2014; Clop et al., 2012; Feuk et al., 2006). These genomic modifications have significantly shaped gene-trait associations by influencing gene dosage, regulation and transcription (da Silva et al., 2016; Beckmann et al., 2007; Feuk et al., 2006). As such CNVs offer an opportunity to enhance the characterisation and selection of livestock through the detection of variation in genes associated with phenotypic variation, similar to the use of genetic markers such as SNPs (Guan et al., 2020; Dekkers., 2004).

In the past, gene-trait association studies have mainly relied on SNP markers to explain genetic variation of economical traits (Wiggans et al., 2011, Hayes et al., 2009; Van Raden et al., 2009). Recently, there has been a rising interest in the study of CNVs as significant sources of genetic and phenotypic variation of economically important traits of several livestock including pigs (Paudel et al., 2015), sheep (Liu et al., 2013), cattle (Liu et al., 2009), and goats (Fontanesi et al., 2010a). These studies have revealed that the segregation of CNVs is responsible for the phenotypic variation of several traits related to morphology (Chen et al., 2018; Wright et al., 2009), pigmentation (Henkel et al., 2019; Menzi et al., 2016; Norris and Whan., 2008; Giuffra et al., 2002), sexual development (Pailhoux et al., 2001) and susceptibility to disease (Sundström et al., 2012). However, despite these efforts, the description of CNVs in some livestock is still lacking. For instance, in goats, technical limitations and the moderate quality of genome assemblies have hampered CNV mapping and resulted in fewer studies leaving many local goat breeds uncharacterized. (Liu et al., 2020; Liu et al., 2019; Henkel et al., 2019; Dong et al., 2015; Clop et al., 2012; Fontanesi et al., 2010a).

Local goat populations are under risk of erosion due to the indiscriminate crossbreeding and replacement with improved commercial breeds (Mdladla et al., 2016). This endangers the livelihood of goat farmers in communal production systems in which these goats are an important resource for their economy. Goats found in these communal production systems present great potential for genetic improvement as they display variation in production traits, and adaptation to harsh environmental conditions which offer limited feed. Their adaptability is expressed in their good grazing characteristics, longevity, resistance to diseases, and good fertility (Mdladla et al., 2016). However, due to the absence of improved selection breeding schemes within these communal production systems these local goat breeds remain uncharacterized, and their CNVs have not been characterised. Therefore, this chapter's second objective was to investigate the CNVs in South African indigenous meat-type goats with various coat colour and coat colour patterns. The detection of CNVs influencing coat colour is important because knowledge of their associations with economically important traits is extremely pertinent in the practical application of genomic selection for improved breeding and conservation programs.

4.2 Materials and Methods

4.2.1 Genotypic Data

The merged dataset from the previous chapter was used. The dataset contained a total of 51 676 autosomal SNPs and 329 South African meat-type goats with various coat colours and patterns including plain coats such as white, red, black, grey, and coat colour patterns such as speckled, patchy, belted, white-sided, and black-legs.

4.2.2 Genome-wide association studies

The R (R Core Team 2023) package GAPIT (Genome Association and Prediction Integrated Tool) Multiple Mixed Linear Model (Lipka. 2012) which accounts for multiple levels of relatedness by controlling for both population structure and unequal familial relationships using kinship matrix to reduce false positives was used to carry out several case-control GWASs for goat coat colour. GWASs were performed by comparing the identified coat colour (case) to all other coat colours combined (control). QQ plots were generated in R using ggplot2 (Wickham et al., 2016) to compare observed vs expected P-values to assess the influence of population structure on the GWAS. Furthermore, ggplot2 was used to generate Manhattan plots of the observed P-values (-log10) and the Bonferroni correction which accounts for multiple hypothesis testing was used to determine the threshold value for candidate markers.

4.2.3 Gene functional annotations

Candidate genes were retrieved from the Ensembl genome browser using the goat (ARS1: GCA_001704415.1.2) reference genome assembly. Identified genes were analysed to find biological pathways in which they are involved using the Kyoto Encyclopaedia of Genes and Genomes (KEGG) mapper (Kanehisa and Sato, 2020).

4.2.4 Copy number variation analysis

Signal intensity ratio (log R Ratio, LRR) retrieved from Illumina GenomeStudio was imported into Golden Helix SNP & Variation Suite (SVS) 8.9.1 (Golden Helix Inc., Bozeman, MT, USA) where they were normalized using the default goat GC correlation file to correct the waviness. In addition, principal component analysis (PCA) option was used to correct batch effects/stratification of the input data. The univariate method from the copy number analysis module (CNAM) was used to define the CNV segments using parameters: maximum 20 segments per 10 k markers, minimum 20 markers per segment, and the p-value equals to 0.005 for pairwise permutations. Finally, approximate copy number calls were estimated to genotype the CNVs as one of two types (gain or loss) across all samples.

4.2.5 CNVR gene enrichment and functional annotation

After filtering out CNVs over 5 Mb in length, CNV regions (CNVRs) were detected using the R package HandyCNV (Zhou et al., 2021). Furthermore, Handy CNV was used to reveal genes overlapping the CNVRs based on the Capra hircus (ARS1: GCA_001704415.1.2) reference genome assembly. The gene list was further analysed with the DAVID Gene Functional Classification Tool (http://david.abcc.ncifcrf.gov) (Huang, 2007) to identify the gene ontology (GO) terms for the molecular function, biological process, cellular component, and biological pathway.

4.3 Results

4.3.1 Genome-Wide Association Studies

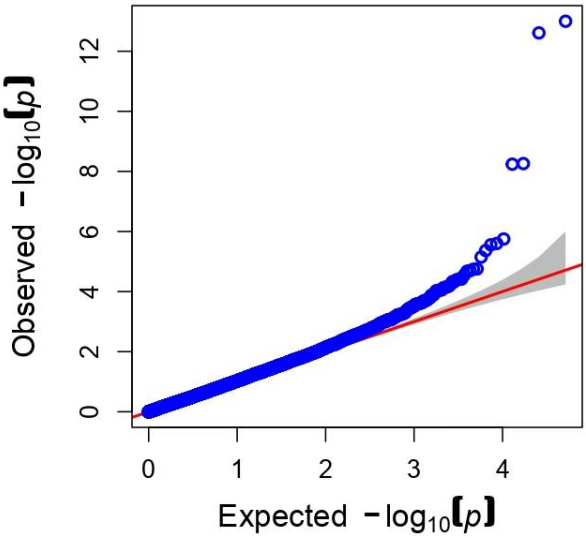

Figure 4.1: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for suggesting population stratification.

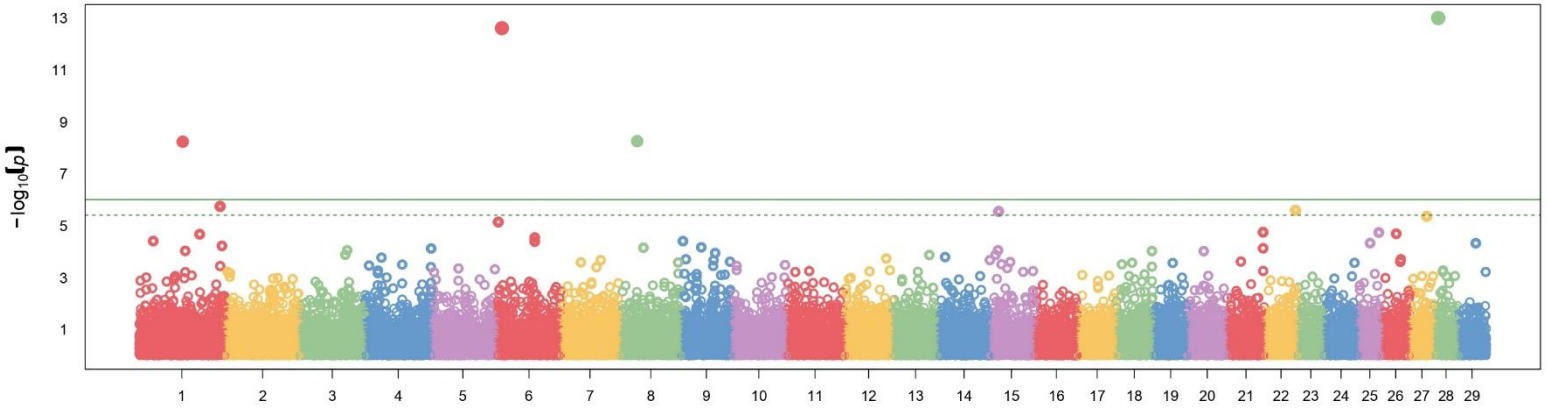

Figure 4.2: Manhattan plot (−log₁₀[P]) of a genome-wide association study on belted coat colour pattern in South African meat-type villages goats. The dotted line represents the genome-wide significance level.

For belted goats, markers for the phenotype were located on chromosomes 1, 6, 8, and 28. The marker on chromosome 1 was located at 79861226bp with a p-value of 5.753300e-09 and a MAF of 0.2659574. The marker on chromosome 6 was located at 10211421bp with a p-value of 2.451406e-13 and a MAF of 0.4863222. The marker on chromosome 8 at position 31516928bp with a p-value of 5.480419e-09 and MAF of 0.2720365. This marker was found to be associated with the *TYRP1* gene, which is involved in pathways including Tyrosine metabolism; Metabolic pathways; and Melanogenesis. The marker on 28 at position 7528711bp with a p-value of 1.001860e-13 and a MAF of 0.5000000.

Table 4.1: Genes detected by genome-wide association study on belted coat colour pattern in South African meat-type village goats.

Coat colour	Chr	Pos	P-value	MAF	Genes	Pathways
Belted	1	79861226	5.753300e-09	0.2659574		
	6	10211421	2.451406e-13	0.4863222		
	8	31516928	5.480419e-09	0.2720365	*TYRP1*	Tyrosine metabolism; Metabolic pathways; Melanogenesis
	28	7528711	1.001860e-13	0.5000000		

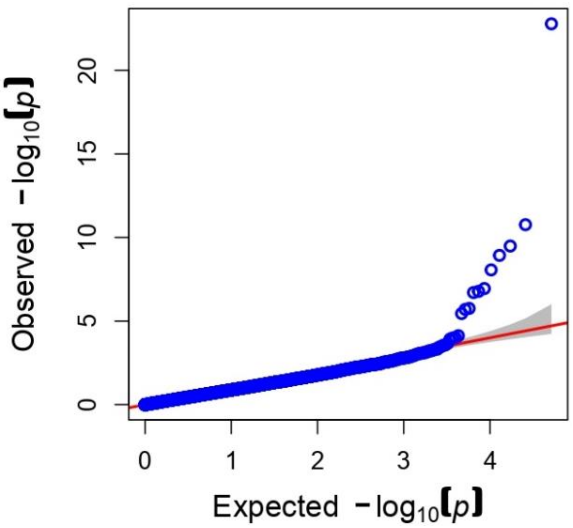

Figure 4.3: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for suggesting population stratification.

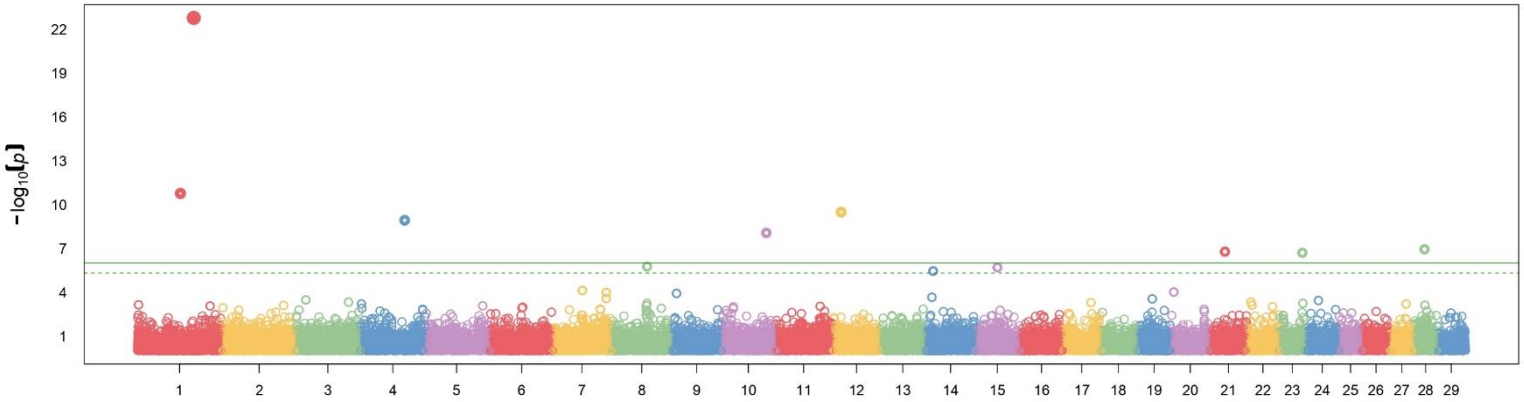

Figure 4.4: Manhattan plot (−log10[P]) of a genome-wide association study on black coat colour in South African meat-type villages goats. The dotted line represents the genome-wide significance level.

For black goats, markers for the phenotype were located on chromosomes 1, 4, 10, 12, 21, 23, and 28. Markers on chromosome 1 were located at positions 79861226bp and 104453302bp and had p-values equal to 1.705343e-11 and 1.642220e-23, as well as MAF of 0.26595745 and 0.44224924, respectively. The marker on 4 was located at position 82149528, with a p-value of 1.169672e-09 and a MAF of 0.10182371. The marker was associated with the *CACNA2D1* gene which is involved in several pathways including MAPK signalling pathway, Cardiac muscle contraction, Adrenergic signalling in cardiomyocytes, Oxytocin signalling pathway, Hypertrophic cardiomyopathy, Arrhythmogenic right ventricular cardiomyopathy, and Dilated cardiomyopathy. The marker on chromosome 10 was located at position 83465784bp, with a p-value of 8.678287e-09 and a MAF of 0.09422492. The marker was associated with the *LRRC49* gene. The marker on chromosome 12 was located on position 15136327bp and had a p-value of 3.240432e-10 and a MAF of 0.33586626. The marker on chromosome 21 was located at position 28774805bp with a p-value of 1.665563e-07 and a MAF of 0.06686930. The marker on chromosome 23 was located at position 42822352bp with a p-value of 1.959252e-07 and a MAF of 0.12462006. The marker on chromosome 28 was located on

position 19783507 bp with a p-value of 1.133841e-07 and a MAF of 0.22644377. This marker was linked to the *FAM241B* gene.

Table 4.2: Genes detected genome-wide association study on black coat colour in South African meat-type villages goats.

Phenotype	Chr	Position	P-value	MAF	Genes	Pathways
Black	1	79861226	1.705343e-11	0.26595745		
	1	104453302	1.642220e-23	0.44224924		
	4	82149528	1.169672e-09	0.10182371	*CACNA2D1*	MAPK signalling pathway, Cardiac muscle contraction, Adrenergic signalling in cardiomyocytes, Oxytocin signalling pathway, Hypertrophic cardiomyopathy, Arrhythmogenic right ventricular cardiomyopathy, Dilated cardiomyopathy
	10	83465784	8.678287e-09	0.09422492	*LRRC49*	
	12	15136327	3.240432e-10	0.33586626		
	21	28774805	1.665563e-07	0.06686930		
	23	42822352	1.959252e-07	0.12462006		
	28	19783507	1.133841e-07	0.22644377	*FAM241B*	

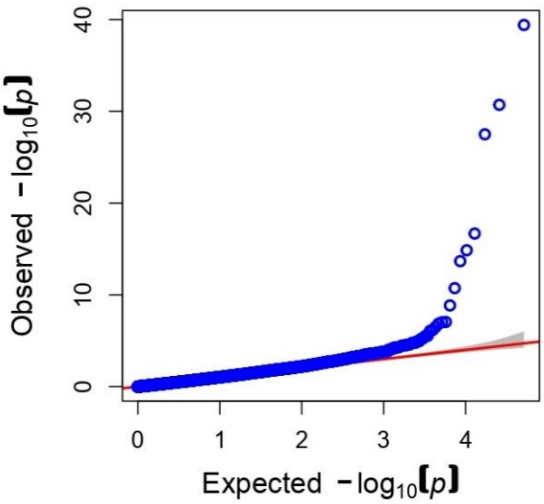

Figure 4.5: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for population stratification.

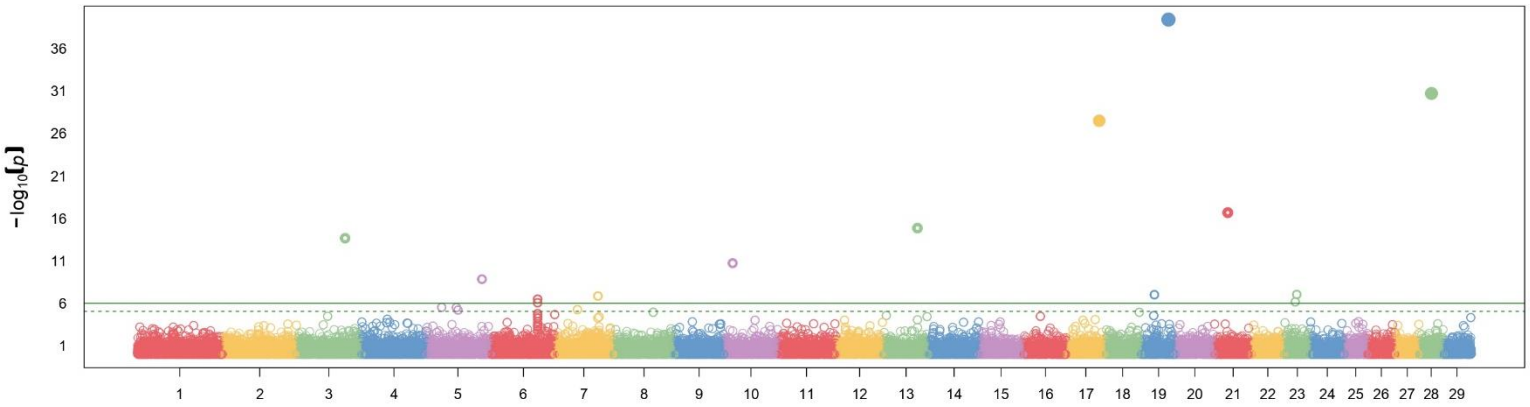

Figure 4.6: Manhattan plot (−log10[P]) of a genome-wide association study on blacklegs coat colour pattern in South African meat-type villages goats. The dotted line represents the genome-wide significance level.

For the goats with the blacklegs coat pattern phenotype, markers were detected on chromosomes 3, 5, 6, 7, 10, 1, 17, 19, 21, 23, and 28. The marker on chromosome 3 was located at position 89933466bp with a p-value of 2.115872e-14 and a MAF of 0.4848024. The marker on chromosome 5 was located at position 1,02E+08bp with a p-value of 1.415866e-09 and a MAF of 0.4300912. Markers on chromosome 6 were located at positions 85987197bp and 85988705bp with p-values equal to 3.174854e-07 and 8.428908e-07as well as MAF of 0.3829787 and 0.3844985, respectively. The marker on chromosome 7 was located at position 79754464bp with a p-value of 1.365552e-07 and a MAF of 0.2249240. The marker on chromosome 10 was located at position 16162904bp with a p-value of 1.850649e-11 and a MAF of 0.4984802. The marker on chromosome 13 was located at position 63273673bp with a p-value of 1.379939e-15 and a MAF of 0.4726444. This marker was linked to the *AHCY* gene which is involved in Cysteine and methionine metabolism, and Metabolic pathways. The marker on chromosome 17 was located on position 60223612bp with a p-value of 3.175654e-28 and a MAF of 0.4164134. The marker was linked to the *EDNRA* gene which is involved in several pathways including Calcium signalling pathway, cGMP-PKG signalling pathway, cAMP signalling pathway, Neuroactive ligand-receptor interaction, Vascular smooth muscle

contraction, Renin secretion, and Pathways in cancer. Markers on chromosome 19 were located at positions 23859595 bp and 49851241bp, with p-values equal to 9.627588e-08 and 3.855798e-40, and MAF of 0.2006079 and 0.4544073, respectively. The marker on chromosome 21 was located on position 25406760bp with a p-value of 2.088736e-17 and a MAF of 0.4772036. Markers on chromosome 23 was located at position 20456303bp and 23252192bp with p-values equal to 6.656304e-07 and 9.000976e-08, as well as MAF of 0.4908815 and 0.4620061. The marker detected on chromosome 28 was located at position 22354089bp with a p-value of 1.946135e-31 and a MAF of 0.4331307.

Table 4.3: Genes detected by genome-wide association study for blacklegs coat colour pattern in South African meat-type villages goats.

Coat colour	Chr	Pos	P-value	MAF	Genes	Pathways
Blacklegs	3	89933466	2.115872e-14	0.4848024		
	5	1,02E+08	1.415866e-09	0.4300912		
	6	85987197	3.174854e-07	0.3829787		
	6	85988705	8.428908e-07	0.3844985		
	7	79754464	1.365552e-07	0.2249240		
	10	16162904	1.850649e-11	0.4984802		
	13	63273673	1.379939e-15	0.4726444	*AHCY*	Cysteine and methionine metabolism, Metabolic pathways
	17	60223612	3.175654e-28	0.4164134	*EDNRA*	Calcium signalling pathway, cGMP-PKG signalling pathway, cAMP signalling pathway, Neuroactive ligand-receptor interaction, Vascular smooth muscle contraction, Renin secretion, Pathways in cancer
	19	23859595	9.627588e-08	0.2006079		
	19	49851241	3.855798e-40	0.4544073		
	21	25406760	2.088736e-17	0.4772036		
	23	20456303	6.656304e-07	0.4908815		
	23	23252192	9.000976e-08	0.4620061		
	28	22354089	1.946135e-31	0.4331307		

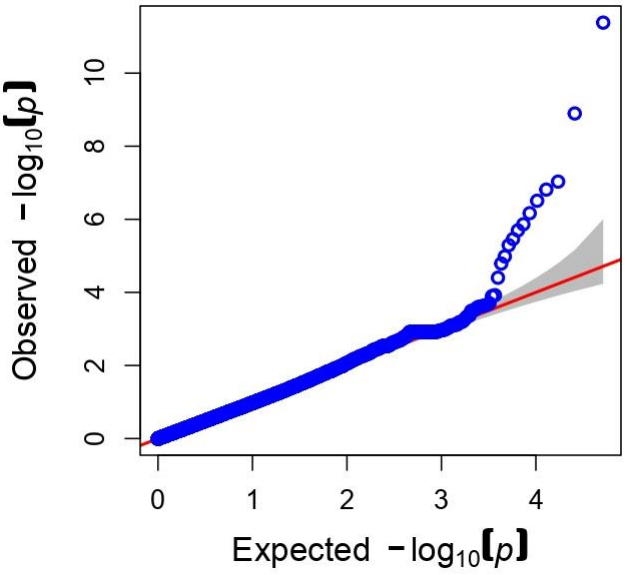

Figure 4.7: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for population stratification.

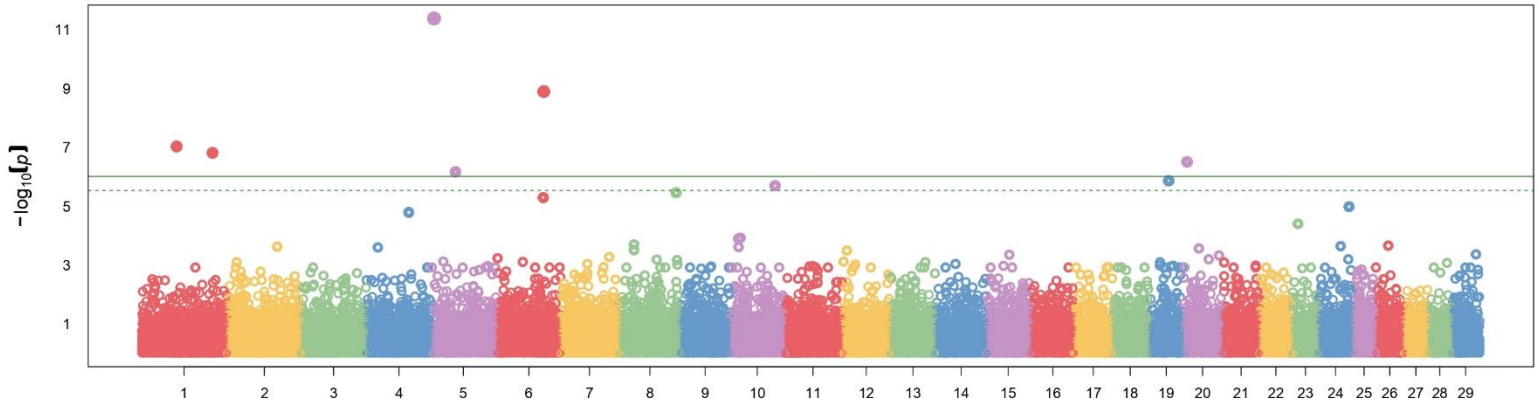

Figure 4.8: Manhattan plot (−log10[P]) of a genome-wide association study on grey coat colour in South African meat-type villages goats. The dotted line represents the genome-wide significance level.

For goats with the grey coat, markers were found on chromosome 1, 5, 6, and 20. The first of the markers on chromosome 1 was located at position 64366087bp with a p-value of 9.295806e-08 and a MAF of 0.11702128. This marker was linked to the *GSK3B* gene which is involved in several pathways including EGFR tyrosine kinase inhibitor resistance, ErbB signalling pathway, Chemokine signalling pathway, Cell cycle, mTOR signalling pathway, PI3K-Akt signalling pathway, Wnt signalling pathway, Hedgehog signalling pathway, Axon guidance, Hippo signalling pathway, Focal adhesion, Signalling pathways regulating pluripotency of stem cells, IL-17 signalling pathway, T cell receptor signalling pathway, B cell receptor signalling pathway, Neurotrophin signalling pathway, Dopaminergic synapse, Insulin signalling pathway, Melanogenesis, Prolactin signalling pathway, Thyroid hormone signalling

pathway, Insulin resistance, Non-alcoholic fatty liver disease, Cushing syndrome, Growth hormone synthesis, secretion and action, Alcoholic liver disease, Alzheimer disease, Prion disease, Pathways of neurodegeneration - multiple diseases, Yersinia infection, Hepatitis C, Measles, Human cytomegalovirus infection, Human papillomavirus infection, Kaposi sarcoma-associated herpesvirus infection, Pathways in cancer, Colorectal cancer, Endometrial cancer, Prostate cancer, Basal cell carcinoma, Breast cancer, Hepatocellular carcinoma, Gastric cancer, Diabetic cardiomyopathy, and Lipid and atherosclerosis. The second marker that was detected on chromosome 1 was located at position 130505864bp with a p-value of 1.537644e-07 and a MAF of 0.27507599. This marker was associated with the *ARMC8* gene, whose biological pathways have not yet been described in goats. Markers on chromosome 5 were located at positions 3978285bp and 43399844bp with p-values equal to 4.151461e-12 and 6.802867e-07, as well as MAF of 0.19756839 and 0.36018237, respectively. The marker on chromosome 6 was located at position 87026491bp with a p-value of 1.270636e-09 and a MAF of 0.07902736. The marker was linked to the *SLC4A4* gene which is involved in Proximal tubule bicarbonate reclamation, Pancreatic secretion, and Bile secretion. The marker on chromosome 20 was located at position 7339359bp with a p-value of 3.104850e-07 and a MAF of 0.05471125.

Table 4.4: Genes detected by genome-wide association for grey coat colour in South African meat-type villages goats.

Phenotype	Chr	Position	P-value	MAF	Genes	Pathways
Grey	1	64366087	9.295806e-08	0.11702128	*GSK3B*	EGFR tyrosine kinase inhibitor resistance, ErbB signalling pathway, Chemokine signalling pathway, Cell cycle, mTOR signalling pathway, PI3K-Akt signalling pathway, Wnt signalling pathway, Hedgehog signalling pathway, Axon guidance, Hippo signalling pathway, Focal adhesion, Signalling pathways regulating pluripotency of stem cells, IL-17 signalling pathway, T cell receptor signalling pathway, B cell receptor signalling pathway, Neurotrophin signalling pathway, Dopaminergic synapse, Insulin signalling pathway, Melanogenesis, Prolactin signalling pathway, Thyroid hormone signalling pathway, Insulin resistance, Non-alcoholic fatty liver disease, Cushing syndrome, Growth hormone synthesis, secretion and action, Alcoholic liver disease, Alzheimer disease, Prion disease, Pathways of neurodegeneration - multiple diseases, Yersinia infection, Hepatitis C, Measles, Human cytomegalovirus infection, Human papillomavirus infection, Kaposi sarcoma-associated herpesvirus infection, Pathways in cancer, Colorectal cancer, Endometrial cancer, Prostate cancer, Basal cell carcinoma, Breast cancer, Hepatocellular carcinoma, Gastric cancer, Diabetic cardiomyopathy, Lipid and atherosclerosis,
	1	130505864	1.537644e-07	0.27507599	*ARMC8*	
	5	3978285	4.151461e-12	0.19756839		
	5	43399844	6.802867e-07	0.36018237		
	6	87026491	1.270636e-09	0.07902736	*SLC4A4*	Proximal tubule bicarbonate reclamation, Pancreatic secretion, Bile secretion
	20	7339359	3.104850e-07	0.05471125		

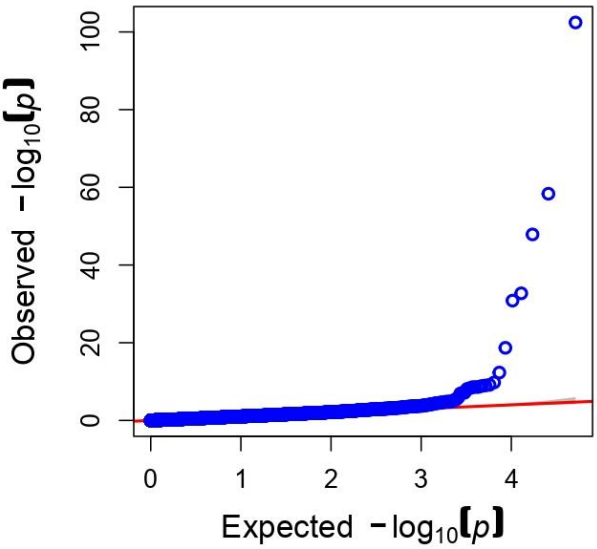

Figure 4.9: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for population stratification.

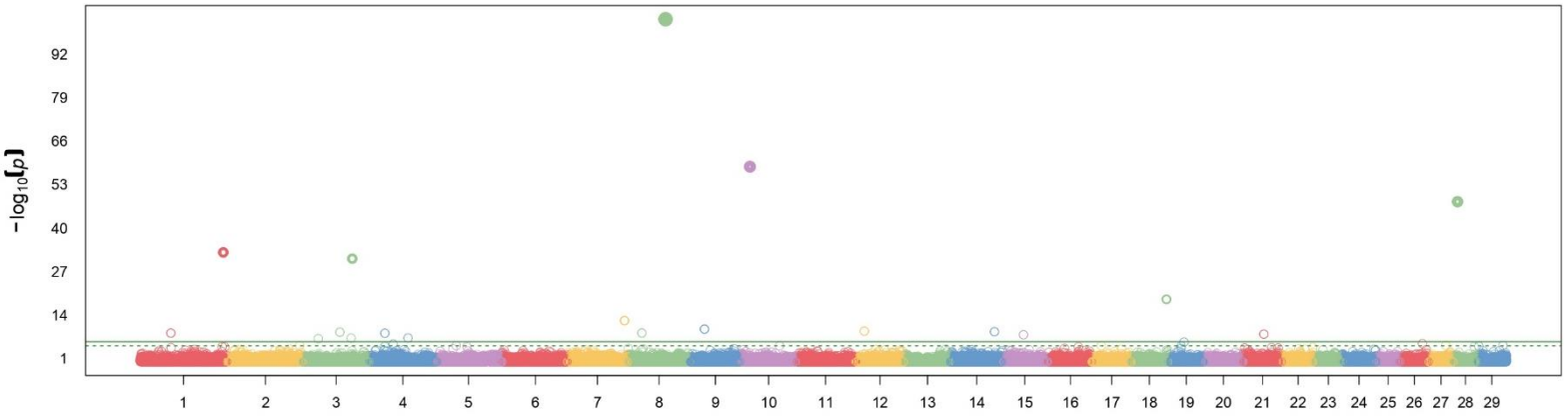

Figure 4.10: Manhattan plot (−log10[P]) of a genome-wide association study on patchy coat colour pattern in South African meat-type villages goats. The dotted line represents the genome-wide significance level.

For goats with the patchy coat colour pattern, markers were detected on chromosome 1, 3, 4, 7, 8, 9, 10, 12, 14, 15, 18, 21, and 28. Markers on chromosome 1 were located at positions 55130041bp and 1,5E+08bp with p-values equal to 2.619407e-09 and MAF of 0.2036474 and 0.4787234, respectively. Multiple markers were detected on chromosome 3, the first of these was located at position 27454000bp with a p-value of 1.219215e-07 and a MAF of 0.4969605. This marker was associated with the *SCP2* gene, which is involved in Primary bile acid biosynthesis, Biosynthesis of unsaturated fatty acids, Metabolic pathways, Fatty acid metabolism, PPAR signalling pathway, and Peroxisome. The second marker was located at position 66351175bp with a p-value of 1.370628e-09 and a MAF of 0.4939210. The third marker was located at position 86658898bp with a p-value of 7.987510e-08 and a MAF of 0.4878419. This marker was linked to the *GNAI3* gene which is involved in several pathways including Rap1 signalling pathway, cGMP-PKG signalling pathway, cAMP signalling pathway, Chemokine signalling pathway, Sphingolipid signalling pathway, Adrenergic signalling in cardiomyocytes, Axon guidance, Apelin signalling pathway, Gap junction,

Platelet activation, Leukocyte transendothelial migration, Circadian entrainment, Retrograde endocannabinoid signaling, Glutamatergic synapse, Cholinergic synapse, Serotonergic synapse, GABAergic synapse, Dopaminergic synapse, Long-term depression, Progesterone-mediated oocyte maturation, Estrogen signalling pathway, Melanogenesis, Oxytocin signalling pathway, Regulation of lipolysis in adipocytes, Renin secretion, Relaxin signalling pathway, Parathyroid hormone synthesis, secretion and action, Cushing syndrome, Growth hormone synthesis, secretion and action, Gastric acid secretion, Parkinson disease, Cocaine addiction, Morphine addiction, Alcoholism, Pertussis, Chagas disease, Toxoplasmosis, Human cytomegalovirus infection, Human immunodeficiency virus 1 infection, Pathways in cancer, and Chemical carcinogenesis - receptor activation. The last marker on chromosome 3 was located at position 88728896bp with a p-value of 1.443606e-31 and a MAF of 0.4954407. Markers on chromosome 4 were located at positions 27520531bp and 69328818bp, with p-values equal to 2.862607e-09 and 7.344395e-08, as well as MAF of 0.4832827 and 0.2036474, respectively. The marker on chromosome 7was located at position 1,02E+08 with a p-value of 4.832853e-13 and a MAF of 0.4878419. Markers on chromosome 8 were located at positions 25402021bp and 68153403bp with p-values equal to 2.431293e-09 and 3.658960e-103, as well as MAF of 0.2036474 and 0.4741641, respectively. The marker located at position 25402021bp was linked to the *ADAMTSL1* gene. The marker on chromosome 9 was located at position 25744891bp with a p-value of 1.659506e-10 and a MAF of 0.4969605. The marker at chromosome 10 was located at position 16255009bp with a p-value of 4.513714e-59 and a MAF of 0.4969605. The marker at chromosome 12 was located at position 15526787bp with a p-value of 6.927573e-10 and a MAF of 0.4984802. The marker at chromosome 14 was located at position 79879596bp with a p-value of 1.009905e-09 and a MAF of 0.4908815. The marker at chromosome 15 was located at position 37603025bp with a p-value of 8.341817e-09 and a MAF of 0.2340426. The marker at chromosome 18 was located at position 63143082bp with a p-value of 2.050970e-19 and a MAF of 0.4680851. The marker at chromosome 21 was located at position 37502153bp with a p-value of 5.400542e-09 and a MAF of 0.4711246. The marker at chromosome 28 was located at position 7488059bp with a p-value of 1.316510e-48 and a MAF of 0.4984802.

Table 4.5: Genes detected by genome-wide association study for patchy coat colour pattern in South African meat-type villages goats.

Coat colour	Chr	Pos	P-value	MAF	Genes	Pathways
Patchy	1	55130041	2.619407e-09	0.2036474		
	1	1,5E+08	1.917049e-33	0.4787234		
	3	27454000	1.219215e-07	0.4969605	*SCP2*	Primary bile acid biosynthesis, Biosynthesis of unsaturated fatty acids, Metabolic pathways, Fatty acid metabolism, PPAR signalling pathway, Peroxisome
	3	66351175	1.370628e-09	0.4939210		
	3	86658898	7.987510e-08	0.4878419	*GNAI3*	Rap1 signalling pathway, cGMP-PKG signalling pathway, cAMP signalling pathway, Chemokine signalling pathway, Sphingolipid signalling pathway, Adrenergic signalling in cardiomyocytes, Axon guidance, Apelin signalling pathway, Gap junction, Platelet activation, Leukocyte transendothelial migration, Circadian entrainment, Retrograde endocannabinoid signaling, Glutamatergic synapse, Cholinergic synapse, Serotonergic synapse, GABAergic synapse, Dopaminergic synapse, Long-term depression, Progesterone-mediated oocyte maturation, Estrogen signalling pathway, Melanogenesis, Oxytocin signalling pathway, Regulation of lipolysis in adipocytes, Renin secretion, Relaxin signalling pathway, Parathyroid hormone synthesis, secretion and action, Cushing syndrome, Growth hormone synthesis, secretion and action, Gastric acid secretion, Parkinson disease, Cocaine addiction, Morphine addiction, Alcoholism, Pertussis, Chagas disease, Toxoplasmosis, Human cytomegalovirus infection, Human immunodeficiency virus 1 infection, Pathways in cancer, Chemical carcinogenesis - receptor activation
	3	88728896	1.443606e-31	0.4954407		
	4	27520531	2.862607e-09	0.4832827		
	4	69328818	7.344395e-08	0.2036474		
	7	1,02E+08	4.832853e-13	0.4878419		
	8	25402021	2.431293e-09	0.2036474	*ADAMTSL1*	
	8	68153403	3.658960e-103	0.4741641		
	9	25744891	1.659506e-10	0.4969605		
	10	16255009	4.513714e-59	0.4969605		
	12	15526787	6.927573e-10	0.4984802		
	14	79879596	1.009905e-09	0.4908815		
	15	37603025	8.341817e-09	0.2340426		
	18	63143082	2.050970e-19	0.4680851		
	21	37502153	5.400542e-09	0.4711246		
	28	7488059	1.316510e-48	0.4984802		

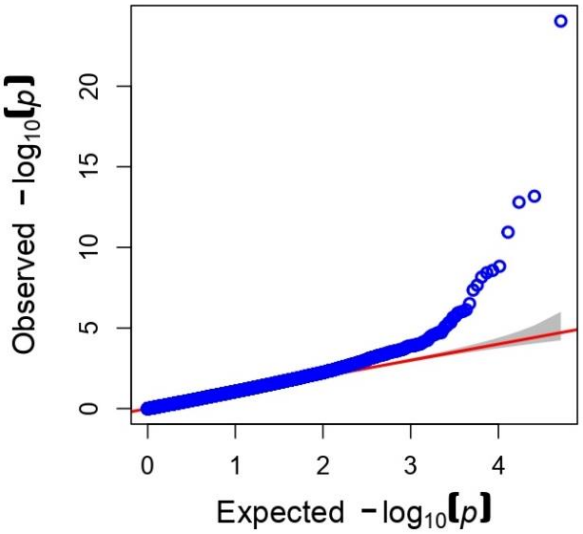

Figure 4.11: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for population stratification.

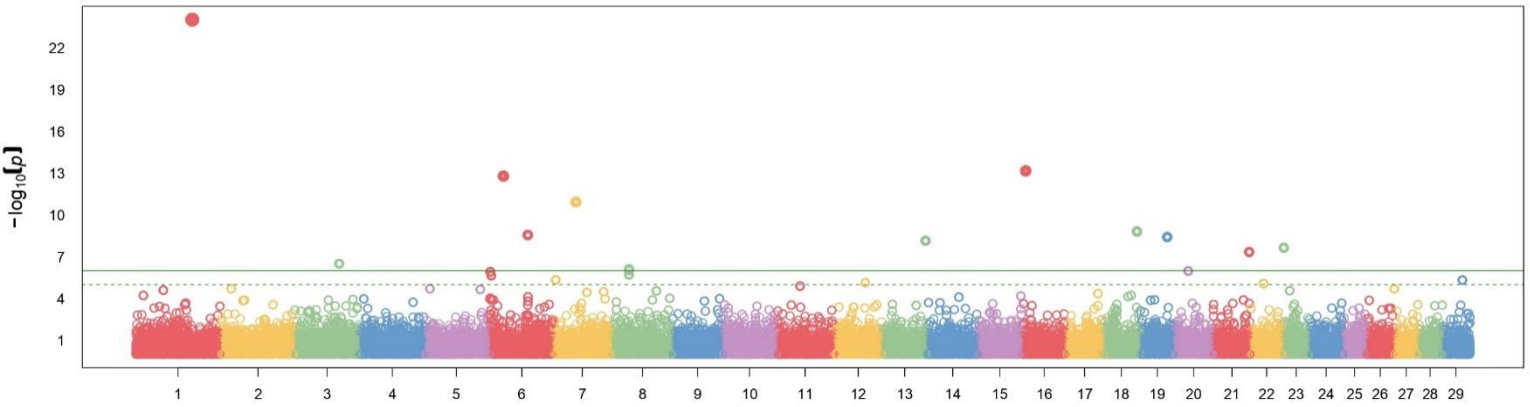

Figure 4.12: Manhattan plot (−log10[P]) of a genome-wide association study on speckled coat colour pattern in South African meat-type villages goats. The dotted line represents the genome-wide significance level.

For goats with the speckled coat pattern, markers were detected on chromosome 1, 3, 6, 7, 8, 13, 16, 18, 19, 21, and 23. The marker on chromosome 1 was located at position 104453302bp with a p-value of 9.211046e-25 and a MAF of 0.44224924. The marker on chromosome 3 was located at position 82018719bp with a p-value of 3.060380e-07 and a MAF of 0.09422492. The markers on chromosome 6 were located at positions 25688756bp and 70777648bp, with p-values equal to 1.586355e-13 and 2.689840e-09, as well as MAF equal to 0.47720365 and 0.29179331, respectively. The marker at position 70777648bp was associated to the *KIT* gene, which is involved several pathways including MAPK signalling pathway, Ras signalling pathway, Rap1 signalling pathway, Phospholipase D signalling pathway, PI3K-Akt signalling pathway, Hematopoietic cell lineage, Melanogenesis, Pathways in cancer, Acute myeloid leukemia, Breast cancer, and Central carbon metabolism in cancer. The marker on chromosome 7 was located at position 41939872bp with a p-value of 1.145001e-11 and a MAF of 0.49088146. The markers on chromosome 8 were located at positions 31504120bp and 31510435bp, respectively. The marker at position 31504120bp had a p-value of 9.206773e-07 and a MAF equal to 0.26899696. While the marker located at position 31510435bp had a p-

value equal to 7.402793e-07 and a MAF of 0.27355623. This marker was linked to the *TYRP1* gene, which is involved in Tyrosine metabolism, Metabolic pathways, and Melanogenesis. The marker on chromosome 13 was located at position 80644220bp with a p-value of 6.981786e-09 and a MAF of 0.39665653. The marker on chromosome 16 was located at position 6163018bp with a p-value of 6.682798e-14 and a MAF of 0.48024316. The marker on chromosome 18 was located at position 61040257bp with a p-value of 1.493746e-09 and a MAF of 0.44832827. The marker on chromosome 19 was located at position 49851241bp with a p-value of 3.760022e-09 and a MAF of 0.45440729. The marker on chromosome 21 was located at position 67119420bp with a p-value of 4.424409e-08 and a MAF of 0.44680851. The marker on chromosome 23 was located at position 1417518bp with a p-value of 2.219845e-08 and a MAF of 0.31155015.

Table 4.6: Genes detected by genome-wide association study for speckled coat colour pattern in South African meat-type villages goats.

Phenotype	Chr	Position	P-value	MAF	Genes	Pathways
Speckled	1	104453302	9.211046e-25	0.44224924		
	3	82018719	3.060380e-07	0.09422492		
	6	25688756; 70777648	1.586355e-13; 2.689840e-09	0.47720365; 0.29179331	*KIT*	MAPK signalling pathway, Ras signalling pathway, Rap1 signalling pathway, Phospholipase D signalling pathway, PI3K-Akt signalling pathway, Hematopoietic cell lineage, Melanogenesis, Pathways in cancer, Acute myeloid leukemia, Breast cancer, Central carbon metabolism in cancer
	7	41939872	1.145001e-11	0.49088146		
	8	31504120; 31510435	9.206773e-07; 7.402793e-07	0.26899696; 0.27355623	*TYRP1*	Tyrosine metabolism, Metabolic pathways, Melanogenesis
	13	80644220	6.981786e-09	0.39665653		
	16	6163018	6.682798e-14	0.48024316		
	18	61040257	1.493746e-09	0.44832827		
	19	49851241	3.760022e-09	0.45440729		
	21	67119420	4.424409e-08	0.44680851		
	23	1417518	2.219845e-08	0.31155015		

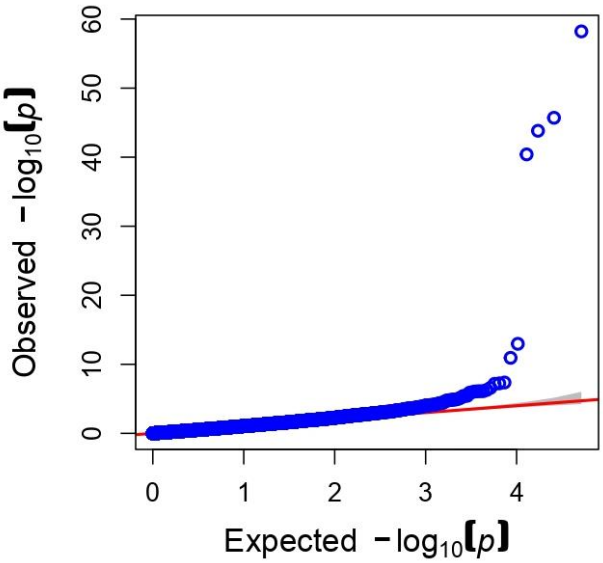

Figure 4.13: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for population stratification.

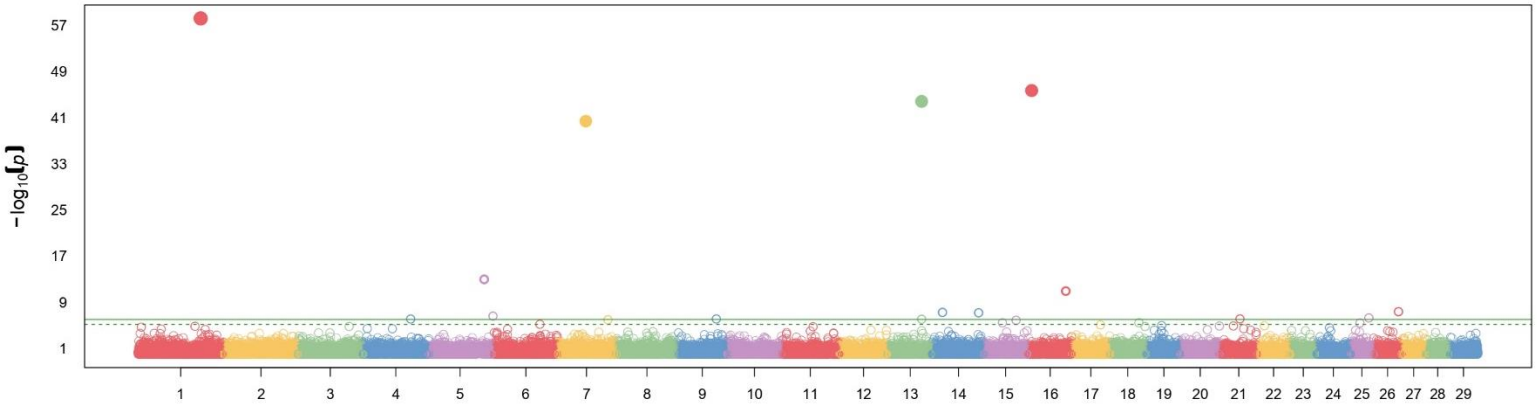

Figure 4.14: Manhattan plot (−log10[P]) of a genome-wide association study on white-sided coat colour pattern in South African meat-type villages goats. The dotted line represents the genome-wide significance level.

For goats with the white-sided coat colour pattern, markers were detected o chromosome 1, 4, 5, 7, 9, 13, 14, 16, 21, 25, and 26. The markers on chromosome 1 was located at position 1,15E+08bp with a p-value of 6.083501e-59 and a MAF of 0.48024316. This marker was linked to the *MBNL1* gene. The marker on chromosome 4 was located at position 87487318bp with a p-value of 8.177479e-07 and a MAF of 0.48480243. The markers on chromosome 6 were located at positions 1,02E+08bp and 1,18E+08bp, with p-values equal to 1.082369e-13, 2.547591e-07, and MAF equal to 0.43009119, 0.25075988, respectively. The marker on chromosome 7 was located at position 52400126bp with a p-value of 3.970844e-41 and a MAF equal to 0.46048632. The marker was linked to the *PPP2R2B* gene, which is involved in several pathways including mRNA surveillance pathway, Sphingolipid signalling pathway, PI3K-Akt signalling pathway, AMPK signalling pathway, Adrenergic signalling in cardiomyocytes, Hippo signalling pathway, Tight junction, Dopaminergic synapse, Chagas disease, Hepatitis C, and Human papillomavirus infection. The marker on chromosome 9 was located at position 71401588bp with a p-value of 7.894490e-07 and a MAF equal to 0.45440729. The markers on

chromosome 13 were located at positions 63229556bp and 63270242bp with p-values equal to 8.811772e-07 and 1.502263e-44 as well as MAF of 0.48328267 and 0.46960486, respectively. The marker at 63270242bp was linked to the *AHCY* gene, which is involved in Cysteine and methionine metabolism, and Metabolic pathways. Markers on chromosome 14 were located at positions 18532999bp and 85005904bp and had p-values equal to 5.784204e-08 and 6.736954e-08, with MAF equal to 0.24924012 and 0.43313070, respectively. Markers on chromosome 16 were located at positions 6163018bp and 68869136bp and had p-values equal to 1.968227e-46 and 1.162438e-11, with MAF equal to 0.48024316 and 0.05471125, respectively. The marker on chromosome 21 was located at position 37502153bp with a p-value of 7.779967e-07 and a MAF of 0.47112462. The marker on chromosome 25 was located at position 34107941bp with a p-value equal to 5.254017e-07 and a MAF of 0.46656535. The marker on chromosome 26 was located at position 45701546bp with a p-value of 4.325519e-08 and a MAF equal to 0.47264438.

Table 4.7: Genes detected by genome-wide association study for white-sided coat colour pattern in South African meat-type villages goats.

Coat colour	Chr	Pos	P-value	MAF	Genes	Pathways
White-sided	1	1,15E+08	6.083501e-59	0.48024316	*MBNL1*	
	4	87487318	8.177479e-07	0.48480243		
	5	1,02E+08	1.082369e-13	0.43009119		
	5	1,18E+08	2.547591e-07	0.25075988		
	7	52400126	3.970844e-41	0.46048632	*PPP2R2B*	mRNA surveillance pathway, Sphingolipid signalling pathway, PI3K-Akt signalling pathway, AMPK signalling pathway, Adrenergic signalling in cardiomyocytes, Hippo signalling pathway, Tight junction, Dopaminergic synapse, Chagas disease, Hepatitis C, Human papillomavirus infection
	9	71401588	7.894490e-07	0.45440729		
	13	63229556	8.811772e-07	0.48328267		
	13	63270242	1.502263e-44	0.46960486	*AHCY*	Cysteine and methionine metabolism, Metabolic pathways
	14	18532999	5.784204e-08	0.24924012		
	14	85005904	6.736954e-08	0.43313070		
	16	6163018	1.968227e-46	0.48024316		
	16	68869136	1.162438e-11	0.05471125		
	21	37502153	7.779967e-07	0.47112462		
	25	34107941	5.254017e-07	0.46656535		
	26	45701546	4.325519e-08	0.47264438		

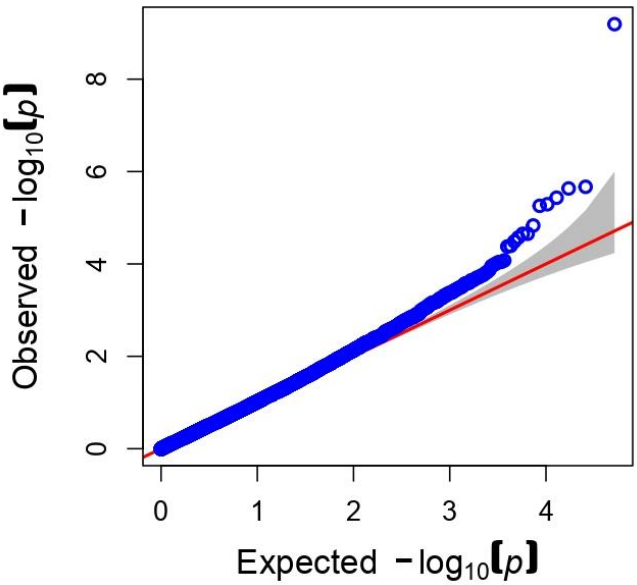

Figure 4.15: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for population stratification.

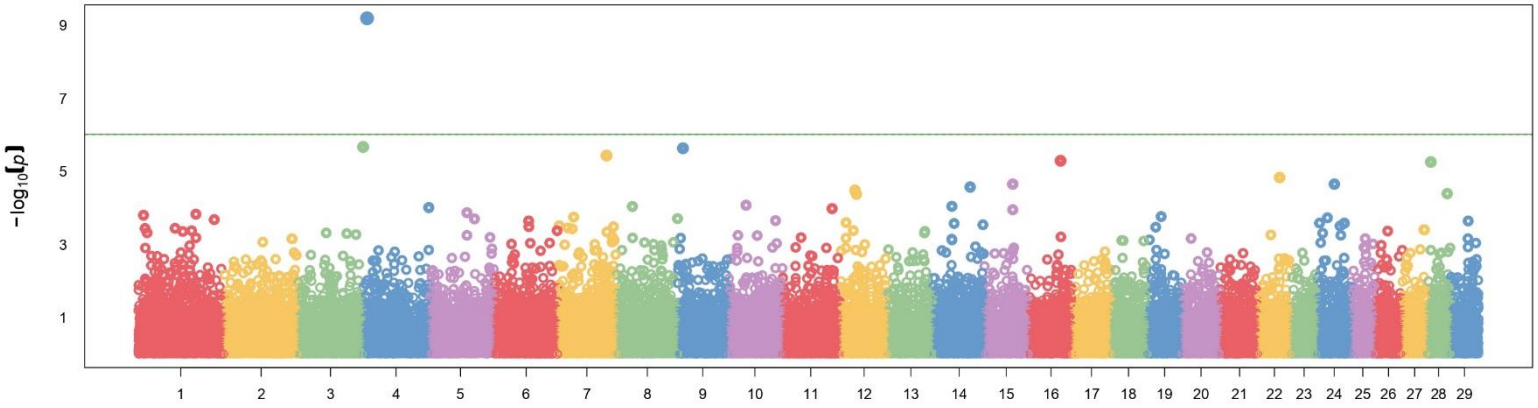

Figure 4.16: Manhattan plot (−log10[P]) of a genome-wide association study on white coat colour in South African meat-type villages goats. The dotted line represents the genome-wide significance level.

For goats with white coat, markers were detected on chromosome 4. The marker was located at position 6869291bp with a p-value of 6.440336e-10 and a MAF of 0.4863222. Several genes 1MB upstream and downstream the marker were found including the *CRYGN*; *WDR86*; *NUB1*; *CHPF2*; *SMARCD3* which is involved in Thermogenesis, and Hepatocellular carcinoma, *ABCF2*; *H2BK1*; *IQCA1L*; *ASB10*; *GBX1*; *AGAP3* which is involved in Endocytosis; *TMUB1*; *CDK5* which is involved in Axon guidance, Alzheimer disease, Pathways of neurodegeneration - multiple diseases, and Cocaine addiction; *ASIC3* which is involved in Inflammatory mediator regulation of TRP channels; *ABCB8* which is involved in ABC transporters; *ATG9B* which is involved in Autophagy - other, Mitophagy - animal, Autophagy – animal; *KCNH2*; *NOS3* which is involved in several pathways including Arginine biosynthesis, Arginine and proline metabolism, Metabolic pathways, Calcium signalling pathway, cGMP-PKG signalling pathway, HIF-1 signalling pathway, Sphingolipid signalling pathway, PI3K-Akt signalling pathway, VEGF signalling pathway, Apelin signalling pathway,

Platelet activation, Estrogen signalling pathway, Oxytocin signalling pathway, Relaxin signalling pathway, Insulin resistance, AGE-RAGE signalling pathway in diabetic complications, Diabetic cardiomyopathy, Lipid and atherosclerosis, and Fluid shear stress and atherosclerosis; AOC1 which is involved in Arginine and proline metabolism, Histidine metabolism, Tryptophan metabolism, and Metabolic pathways; *TMEM176A*; and *TMEM176B*.

Table 4.8: Genes detected by genome-wide association study for white coat colour in South African meat-type villages goats.

Coat colour	Chr	Pos	P-value	MAF	Genes	Pathways
White	4	6869291	6.440336e-10	0.4863222	*CRYGN*	
					WDR86	
					NUB1	
					CHPF2	
					SMARCD3	Thermogenesis, Hepatocellular carcinoma
					ABCF2	
					H2BK1	
					IQCA1L	
					ASB10	
					GBX1	
					AGAP3	Endocytosis
					TMUB1	
					CDK5	Axon guidance, Alzheimer disease, Pathways of neurodegeneration - multiple diseases, Cocaine addiction,
					ASIC3	Inflammatory mediator regulation of TRP channels
					ABCB8	ABC transporters
					ATG9B	Autophagy - other, Mitophagy - animal, Autophagy - animal
					KCNH2	
					NOS3	Arginine biosynthesis, Arginine and proline metabolism, Metabolic pathways, Calcium signalling pathway, cGMP-PKG signalling pathway, HIF-1 signalling pathway, Sphingolipid signalling pathway, PI3K-Akt signalling pathway, VEGF signalling pathway, Apelin signalling pathway, Platelet activation, Estrogen signalling pathway, Oxytocin signalling pathway, Relaxin signalling pathway, Insulin resistance, AGE-RAGE signalling

	pathway in diabetic complications, Diabetic cardiomyopathy, Lipid and atherosclerosis, Fluid shear stress and atherosclerosis,
AOC1	Arginine and proline metabolism, Histidine metabolism, Tryptophan metabolism, Metabolic pathways
TMEM176A	
TMEM176B	

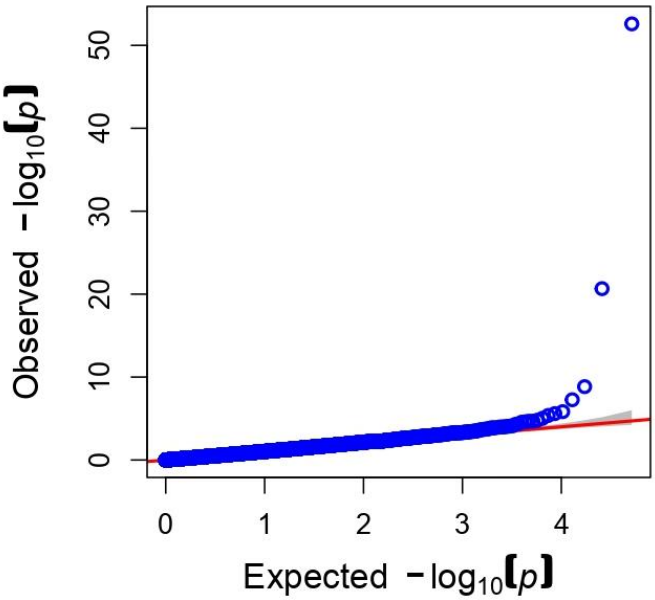

Figure 4.17: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for population stratification.

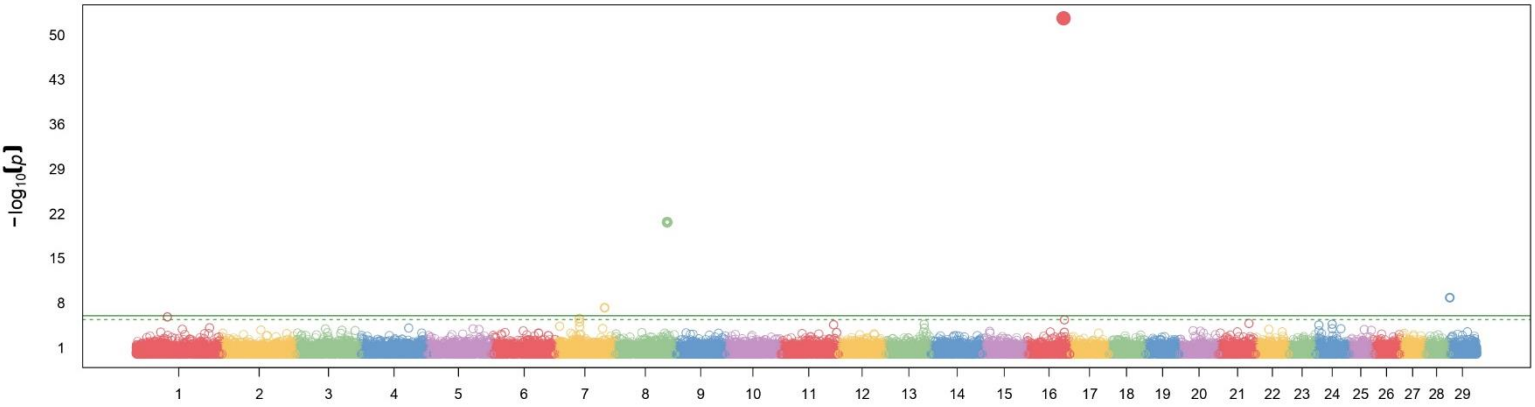

Figure 4.18: Manhattan plot (−log10[P]) of a genome-wide association study on red coat colour in South African meat-type villages goats. The dotted line represents the genome-wide significance level.

For goats with the red coat, markers were detected on chromosome 7, 8, 16, and 29. The marker on chromosome 7 was located at 89971016bp with a p-value equal to 5.431019e-08 and a MAF of 0.1261398. This marker was linked to the *CELF5* gene, and the *TLE6* gene which is involved in the Wnt signalling pathway, and the Notch signalling pathway. The marker on chromosome 8 was located at position 96479822bp with a p-value of 2.184835e-21 and a MAF of 0.1990881. This marker was linked to the *ZNF462* gene. The marker on chromosome 16 was located at position 67305118bp with a p-value of 2.574106e-53 and a MAF of 0.1975684. The marker on chromosome 29 was located at position 1412375bp with a p-value equal to 1.422977e-09 and a MAF of 0.2142857.

Table 4.9: Genes detected by genome-wide association study for red coat colour in South African meat-type villages goats.

Coat colour	Chr	Pos	P-value	MAF	Genes	Pathways
Red	7	89971016	5.431019e-08	0.1261398	*CELF5*, *TLE6;*	Wnt signalling pathway, Notch signalling pathway,
	8	96479822	2.184835e-21	0.1990881	*ZNF462*	
	16	67305118	2.574106e-53	0.1975684		
	29	1412375	1.422977e-09	0.2142857		

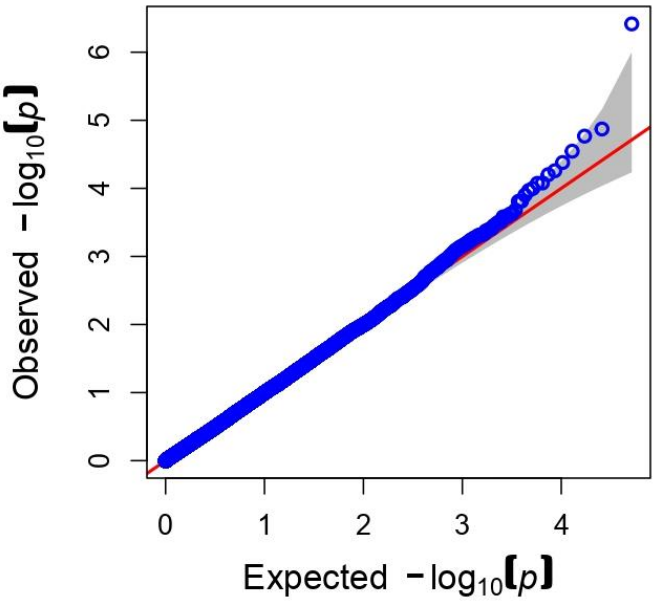

Figure 4.19: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for population stratification.

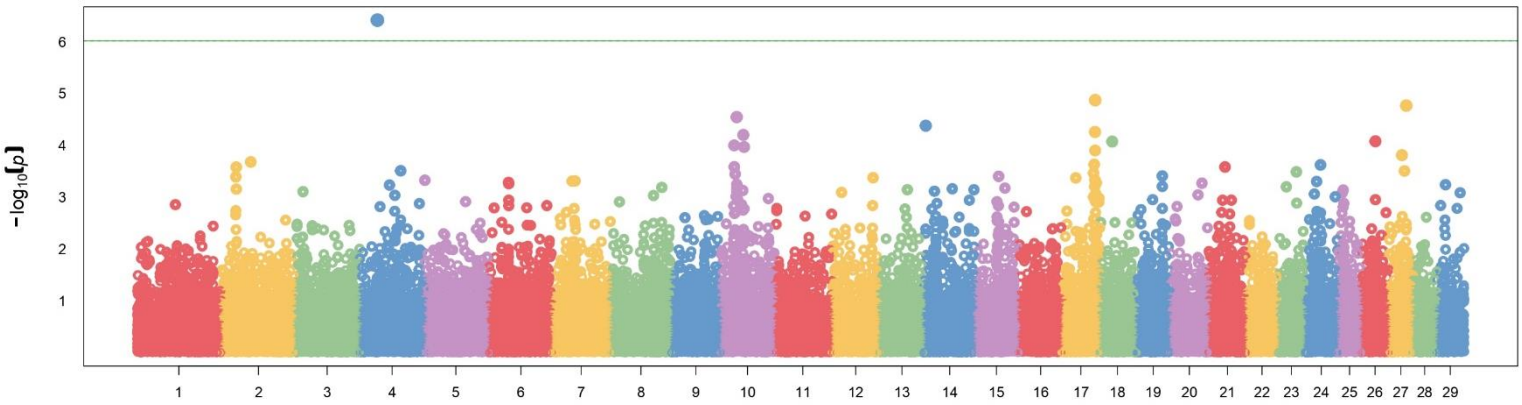

Figure 4.20: Manhattan plot (−log10[P]) of a genome-wide association study on red-head and white body patterns of South African Boer goats. The dotted line represents the genome-wide significance level.

For Boer goats with the white-body and red-head coat colour pattern, a marker was detected on chromosome 4 at position 32890854bp with a p-value of 3.838605e-07and a MAF of 0.4163636. This marker was found to be associated with the *CADPS2*; and the *SLC13A1* genes.

Table 4.10: Genes detected by genome-wide association study for red-head white-body coat colour pattern of the South African Boer breed.

Coat colour	Chr	Pos	P-value	MAF	Genes	Pathways
White-body red-head	4	32890854	3.838605e-07	0.4163636	*CADPS2; SLC13A1*	

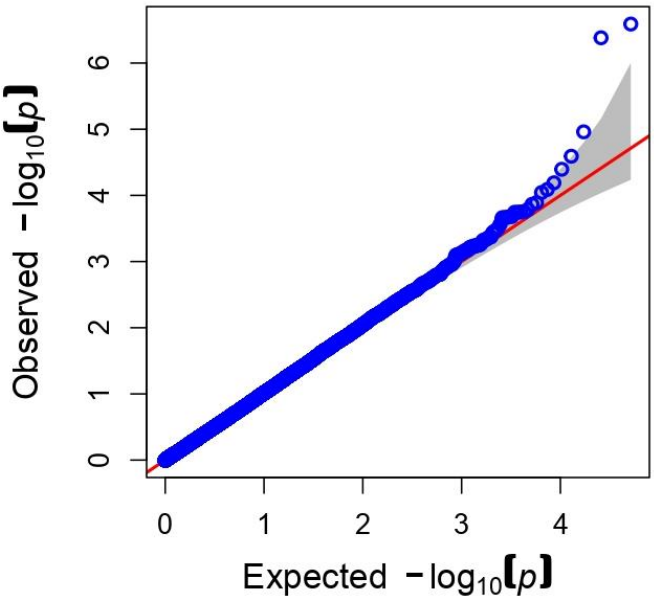

Figure 4.21: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for population stratification.

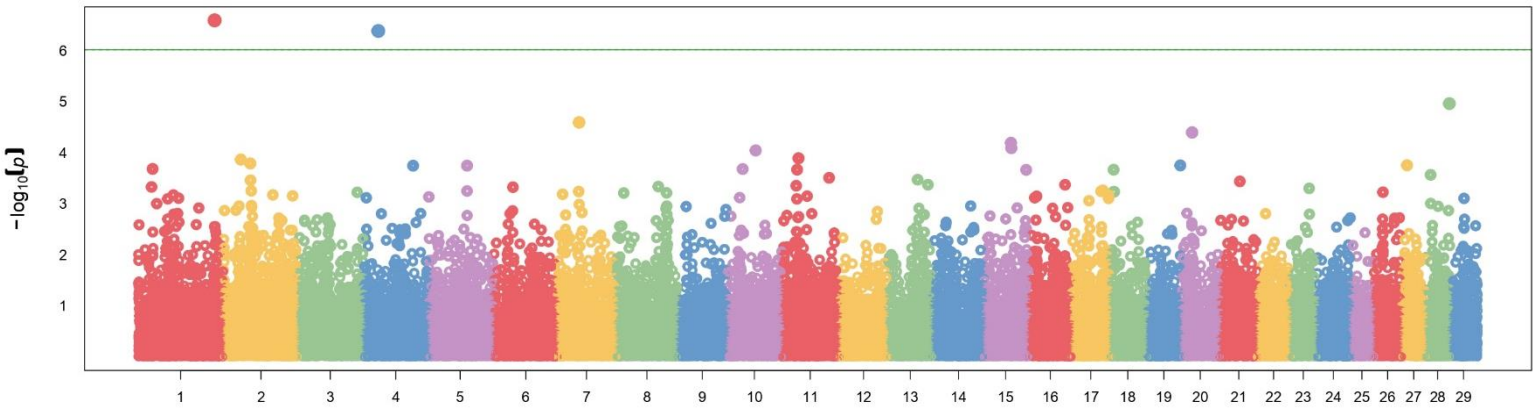

Figure 4.22: Manhattan plot (−log10[P]) of a genome-wide association study on white coat colour of South African Savanna goats. The dotted line represents the genome-wide significance level.

For Savanna goats, markers were detected on chromosome 1 and 4. The marker on chromosome 1 was located at position 1,41E+08bp with a p-value of 2.574560e-07 and a MAF equal to 0.1504559. The marker on chromosome 4 was located at position 27952097bp with a p-value of 4.152916e-07 and a MAF equal to 0.0881459. This marker was linked to the *SND1* gene which is involved in Viral carcinogenesis.

Table 4.11: Genes detected by genome-wide association study for the white coat colour of the South African Savanna breed.

Coat colour	Chr	Pos	P-value	MAF	Genes	Pathways
White	1	1,41E+08	2.574560e-07	0.1504559		
	4	27952097	4.152916e-07	0.0881459	*SND1*	Viral carcinogenesis

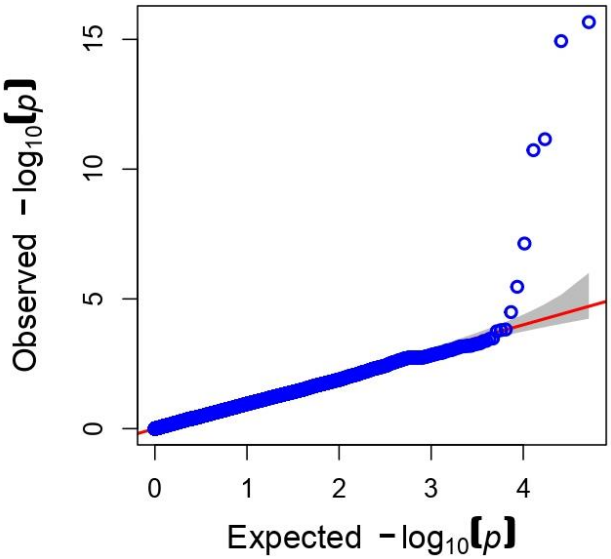

Figure 4.23: Quantile-quantile (QQ) plot of the data shown in the Manhattan plot showing the deviation of the observed P values from the expected values (i.e., the null hypothesis) to check for population stratification.

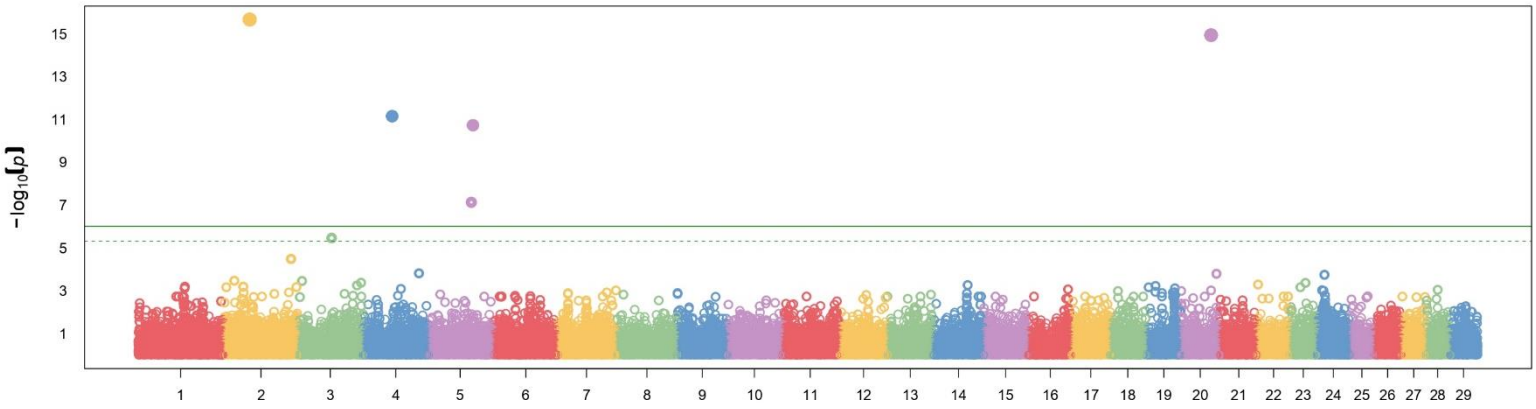

Figure 4.24: Manhattan plot (−log10[P]) of a genome-wide association study on red coat colour of South African Kalahari red goats. The dotted line represents the genome-wide significance level.

For Kalahari Red goats, markers were found on chromosome 2, 4, 5, and 20. The marker on chromosome 2 was located at 47866299bp with a p-value of 2.171808e-16 and a MAF equal to 0.2803030. The marker on chromosome 4 was located at position 53445995bp with a p-value of 7.023353e-12 and a MAF equal to 0.4772727. This marker was linked to the *CHN2* gene. The markers on chromosome 5 were located at positions 78352550bp and 81352549bp with a p-values of 7.355898e-08 and 1.862449e-11 as well MAF equal to 0.3333333 and 0.3484848, respectively. The marker at position 81352549bp was linked to the *TMTC1* gene. The marker on chromosome 20 was located at position 56025700bp with a p-value of 1.166134e-15 and a MAF of 0.3030303.

Table 4.12: Genes detected by genome-wide association study for the red coat colour of the South African Kalahari Red breed.

Coat colour	Chr	Pos	P-value	MAF	Genes	Pathways
Red	2	47866299	2.171808e-16	0.2803030		
	4	53445995	7.023353e-12	0.4772727	*CHN2*	
	5	78352550	7.355898e-08	0.3333333		
	5	81352549	1.862449e-11	0.3484848	*TMTC1*	
	20	56025700	1.166134e-15	0.3030303		

4.3.2 Copy number variation analysis

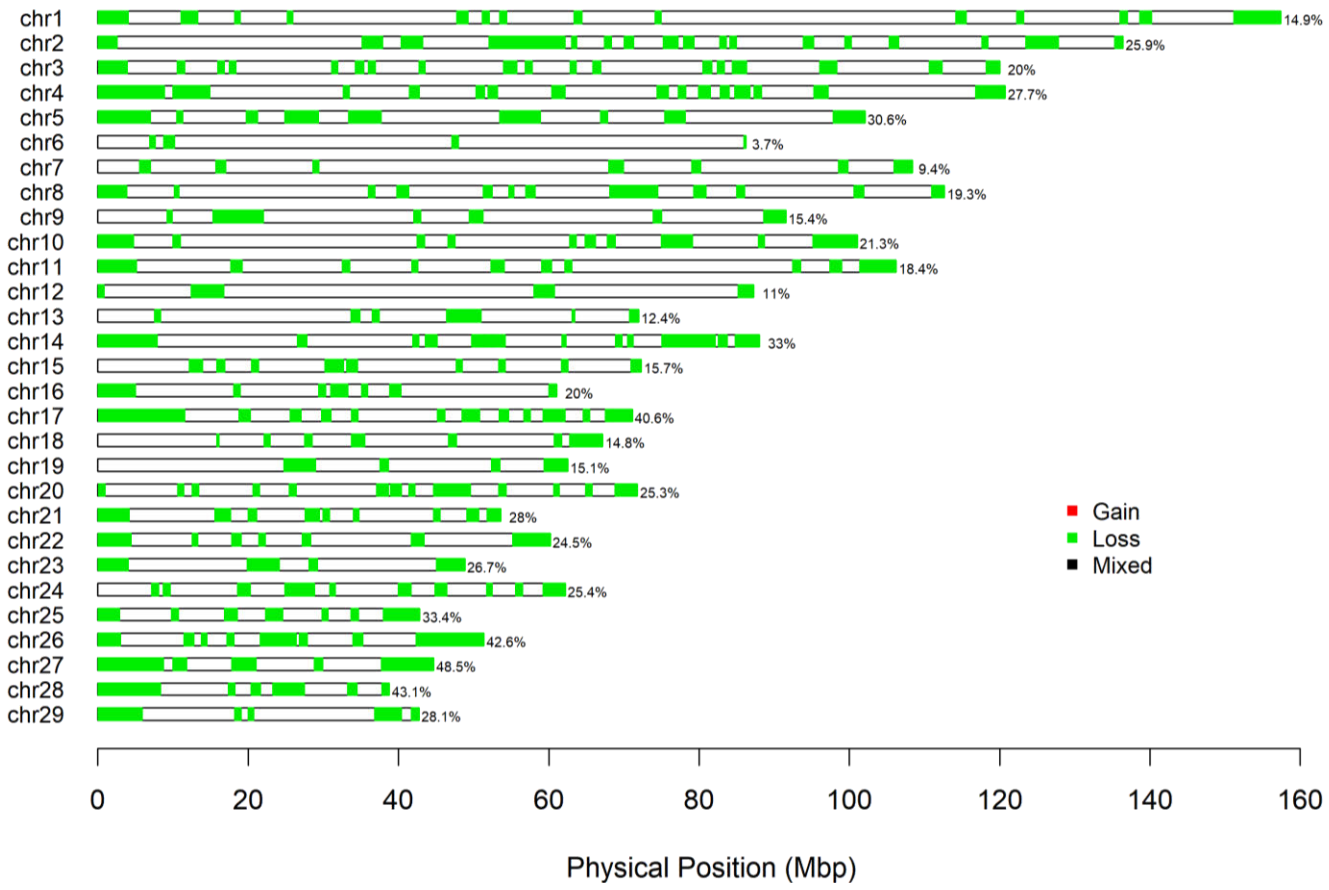

Figure 4.25: Genomic distribution of copy number variation regions on 29 autosomes in South African meat-type goats with various coat colours and coat colour patterns. Green is for loss; red is for gain, and black is for both loss and gain.

In this study a total of 2 047 CNVs were identified with an average number of 12.11 CNVs per individual. Among these CNVs, 115 were gains with an average length of 831464 bp and 1 932 were losses with an average length of 927516 bp. The trend of higher numbers of deletions than duplications observed in this study is in line with previous studies and suggests that there is less selection against duplication CNVs than deletions. From these CNVs, a total of 279 CNVRs with an average length of 1865633 bp were found. Among these, 262 CNVRs overlapped 5 222 genes.

Table 4.13: The gene ontology (GO) categories (biological process) containing genes overlapping the CNVRs identified in SA meat-type goats.

Term	Biological Process	Gene Count
GO:0006355	regulation of transcription, DNA-templated	106
GO:0019372	lipoxygenase pathway	7
GO:0043066	negative regulation of apoptotic process	56
GO:0009952	anterior/posterior pattern specification	22
GO:0010596	negative regulation of endothelial cell migration	8
GO:0031639	plasminogen activation	7
GO:0008630	intrinsic apoptotic signalling pathway in response to DNA damage	11
GO:0006749	glutathione metabolic process	15
GO:0043651	linoleic acid metabolic process	5
GO:0086073	bundle of His cell-Purkinje myocyte adhesion involved in cell communication	5
GO:0021612	facial nerve structural organization	5
GO:0045907	positive regulation of vasoconstriction	6
GO:0030953	astral microtubule organization	6
GO:0031424	keratinization	10
GO:0007032	endosome organization	11
GO:0034122	negative regulation of toll-like receptor signalling pathway	7
GO:0045444	fat cell differentiation	16
GO:0007098	centrosome cycle	10
GO:0002752	cell surface pattern recognition receptor signalling pathway	4
GO:0018101	protein citrullination	4
GO:0033209	tumor necrosis factor-mediated signalling pathway	11
GO:0019827	stem cell population maintenance	13
GO:0061760	antifungal innate immune response	6
GO:0006471	protein ADP-ribosylation	6
GO:0046626	regulation of insulin receptor signalling pathway	6
GO:2000649	regulation of sodium ion transmembrane transporter activity	6

GO:0010332	response to gamma radiation	8
GO:0035265	organ growth	5
GO:0035280	miRNA loading onto RISC involved in gene silencing by miRNA	5
GO:0002223	stimulatory C-type lectin receptor signalling pathway	5
GO:0007160	cell-matrix adhesion	15
GO:0045931	positive regulation of mitotic cell cycle	9
GO:0001938	positive regulation of endothelial cell proliferation	14
GO:1903078	positive regulation of protein localization to plasma membrane	14
GO:0006541	glutamine metabolic process	7
GO:0019369	arachidonic acid metabolic process	7
GO:0042733	embryonic digit morphogenesis	16
GO:0006260	DNA replication	18
GO:0016525	negative regulation of angiogenesis	18
GO:0009954	proximal/distal pattern formation	9
GO:0071345	cellular response to cytokine stimulus	9
GO:0031054	pre-miRNA processing	6
GO:0051895	negative regulation of focal adhesion assembly	6
GO:0006974	cellular response to DNA damage stimulus	27
GO:2000811	negative regulation of anoikis	7
GO:1900027	regulation of ruffle assembly	5
GO:0031666	positive regulation of lipopolysaccharide-mediated signalling pathway	5
GO:0043542	endothelial cell migration	8
GO:0050793	regulation of developmental process	4
GO:0050919	negative chemotaxis	4
GO:0050994	regulation of lipid catabolic process	4
GO:0044818	mitotic G2/M transition checkpoint	4
GO:0007283	spermatogenesis	42
GO:0034116	positive regulation of heterotypic cell-cell adhesion	6
GO:0044790	negative regulation by host of viral release from host cell	6
GO:0060749	mammary gland alveolus development	6
GO:0070534	protein K63-linked ubiquitination	10

GO:0042552	myelination	10
GO:0048839	inner ear development	10
GO:0006865	amino acid transport	9
GO:0090051	negative regulation of cell migration involved in sprouting angiogenesis	5
GO:0006983	ER overload response	5
GO:0009395	phospholipid catabolic process	5
GO:0048511	rhythmic process	11
GO:0009058	biosynthetic process	9
GO:0097191	extrinsic apoptotic signalling pathway	9
GO:0016601	Rac protein signal transduction	6
GO:0032693	negative regulation of interleukin-10 production	6
GO:0032695	negative regulation of interleukin-12 production	6
GO:0006520	cellular amino acid metabolic process	7
GO:0002250	adaptive immune response	14
GO:0042752	regulation of circadian rhythm	11
GO:0010501	RNA secondary structure unwinding	4
GO:0090129	positive regulation of synapse maturation	4
GO:0035095	behavioral response to nicotine	4
GO:0046784	viral mRNA export from host cell nucleus	4
GO:0006287	base-excision repair, gap-filling	4
GO:0002315	marginal zone B cell differentiation	4
GO:0043653	mitochondrial fragmentation involved in apoptotic process	4
GO:0043117	positive regulation of vascular permeability	4
GO:0030422	production of siRNA involved in RNA interference	4
GO:0007219	Notch signalling pathway	21
GO:0050714	positive regulation of protein secretion	10
GO:0060928	atrioventricular node cell development	3
GO:0021697	cerebellar cortex formation	3
GO:0006450	regulation of translational fidelity	3
GO:1901201	regulation of extracellular matrix assembly	3
GO:0090310	negative regulation of methylation-dependent chromatin silencing	3

GO:0060666	dichotomous subdivision of terminal units involved in salivary gland branching	3
GO:0002326	B cell lineage commitment	3
GO:2001303	lipoxin A4 biosynthetic process	3
GO:0045637	regulation of myeloid cell differentiation	3
GO:0046888	negative regulation of hormone secretion	3
GO:0050728	negative regulation of inflammatory response	17
GO:0045662	negative regulation of myoblast differentiation	8
GO:0032007	negative regulation of TOR signaling	8
GO:0022008	neurogenesis	8
GO:0001501	skeletal system development	11
GO:0070266	necroptotic process	5
GO:0010762	regulation of fibroblast migration	5
GO:0035278	miRNA mediated inhibition of translation	5
GO:0060020	Bergmann glial cell differentiation	5
GO:0046847	filopodium assembly	5
GO:1902042	negative regulation of extrinsic apoptotic signalling pathway via death domain receptors	7
GO:1990573	potassium ion import across plasma membrane	6
GO:0048566	embryonic digestive tract development	6
GO:0043129	surfactant homeostasis	6
GO:0050901	leukocyte tethering or rolling	6
GO:0030838	positive regulation of actin filament polymerization	10
GO:0086091	regulation of heart rate by cardiac conduction	10
GO:0009617	response to bacterium	17

Table 4.14: The gene ontology (GO) categories (cellular component) containing genes overlapping the CNVRs identified in SA meat-type goats.

Term	Cellular Component	Gene Count
GO:0005886	plasma membrane	567
GO:0005833	hemoglobin complex	15
GO:0016021	integral component of membrane	1017
GO:0036464	cytoplasmic ribonucleoprotein granule	16
GO:0031091	platelet alpha granule	8
GO:0005654	nucleoplasm	417
GO:0001518	voltage-gated sodium channel complex	8
GO:0014704	intercalated disc	11
GO:0005829	cytosol	490
GO:0005813	centrosome	88
GO:0030057	desmosome	8
GO:0070578	RISC-loading complex	5
GO:0031253	cell projection membrane	5
GO:0016442	RISC complex	6
GO:0031965	nuclear membrane	32
GO:0005925	focal adhesion	32
GO:0071556	integral component of lumenal side of endoplasmic reticulum membrane	7
GO:0009897	external side of plasma membrane	36
GO:0005887	integral component of plasma membrane	83
GO:0030008	TRAPP complex	5
GO:0045095	keratin filament	25
GO:0005765	lysosomal membrane	30
GO:0014069	postsynaptic density	22
GO:0101031	chaperone complex	7
GO:0005577	fibrinogen complex	4
GO:0005788	endoplasmic reticulum lumen	14

GO:0034451	centriolar satellite	23
GO:1990913	sperm head plasma membrane	3
GO:0031225	anchored component of membrane	10
GO:0016323	basolateral plasma membrane	32

Table 4.15: The gene ontology (GO) categories (molecular function) containing genes overlapping the CNVRs identified in SA meat-type goats.

Term	Molecular Function	Gene Count
GO:0004930	G-protein coupled receptor activity	277
GO:0004984	olfactory receptor activity	240
GO:0005344	oxygen transporter activity	15
GO:0019825	oxygen binding	15
GO:0004190	aspartic-type endopeptidase activity	27
GO:0004519	endonuclease activity	15
GO:0004222	metalloendopeptidase activity	35
GO:0022848	acetylcholine-gated cation-selective channel activity	10
GO:0050839	cell adhesion molecule binding	13
GO:1990837	sequence-specific double-stranded DNA binding	51
GO:0016829	lyase activity	8
GO:0043621	protein self-association	18
GO:0080025	phosphatidylinositol-3,5-bisphosphate binding	9
GO:0000981	RNA polymerase II transcription factor activity, sequence-specific DNA binding	42
GO:0017110	nucleoside-diphosphatase activity	4
GO:0016165	linoleate 13S-lipoxygenase activity	4
GO:0004668	protein-arginine deiminase activity	4
GO:0004052	arachidonate 12-lipoxygenase activity	4
GO:0010181	FMN binding	7
GO:0022857	transmembrane transporter activity	32
GO:0046872	metal ion binding	308
GO:0004364	glutathione transferase activity	10
GO:0042802	identical protein binding	209
GO:0005244	voltage-gated ion channel activity	11
GO:0001671	ATPase activator activity	6
GO:0005539	glycosaminoglycan binding	5
GO:0005212	structural constituent of eye lens	7

GO:0015464	acetylcholine receptor activity	5
GO:0005247	voltage-gated chloride channel activity	5
GO:0031625	ubiquitin protein ligase binding	50
GO:0017080	sodium channel regulator activity	6
GO:0005246	calcium channel regulator activity	6
GO:0005507	copper ion binding	14
GO:0005355	glucose transmembrane transporter activity	4
GO:0008607	phosphorylase kinase regulator activity	3
GO:0008506	sucrose:proton symporter activity	3
GO:0045174	glutathione dehydrogenase (ascorbate) activity	3
GO:0050610	methylarsonate reductase activity	3
GO:0050473	arachidonate 15-lipoxygenase activity	3
GO:0047756	chondroitin 4-sulfotransferase activity	3
GO:0046976	histone methyltransferase activity (H3-K27 specific)	3
GO:0035033	histone deacetylase regulator activity	3
GO:0086006	voltage-gated sodium channel activity involved in cardiac muscle cell action potential	3

Table 4.16: The gene ontology (GO) categories (pathways) containing genes overlapping the CNVRs identified in SA meat-type goats.

Term	Kegg Pathway	Gene count
chx04740	Olfactory transduction	487
chx05168	Herpes simplex virus 1 infection	152
chx05332	Graft-versus-host disease	34
chx04612	Antigen processing and presentation	38
chx01523	Antifolate resistance	24
chx04974	Protein digestion and absorption	52
chx04650	Natural killer cell mediated cytotoxicity	57
chx04940	Type I diabetes mellitus	27
chx05330	Allograft rejection	24
chx02010	ABC transporters	26
chx05320	Autoimmune thyroid disease	26
chx04145	Phagosome	53
chx05144	Malaria	19
chx04514	Cell adhesion molecules	49
chx04136	Autophagy - other	13
chx00592	alpha-Linolenic acid metabolism	12

The gene ontology analysis carried on these genes revealed that the functions of the genes included several biological processes, molecular functions, cellular components, and pathways (Tables 4.12 – 4.15). Some of the candidate genes found among the CNVRs were associated with important breeding traits such as coat colour.

4.4 Discussion

4.4.1 Genome-wide association studies

Several key proteins such as *TYR, TYRP1, TYRP2,* and *MITF* have been revealed to catalyse reactions critical to melanin production (Yamaguchi et al., 2007). In mammals, melanin production begins with the production of l-*DOPA* from 1-tyrosine which increases tyrosinase activity, upregulating melanin synthesis and the production of brownish black eumelanin (D'Mello et al., 2016; Moellmann et al., 1988). *TYRP1,* and *TYRP2* have been revealed to increase the production eumelanin (Gibbs et al., 2000). While the synthesis of reddish yellow pheomelanin is dependent on the synthesis of pheomelanin which is produced only in the presence of sulfhydryl compounds such as cysteine/glutathione (Lamoreux et al., 2001; Thody et al., 1991). In this sense, the eumelanin to pheomelanin ratio determines coat pigmentation, with a higher overall eumelanin resulting in darker skin (Lin and Fisher et al., 2007), while individuals with higher pheomelanin tend to have lighter skin (Gupta, 2014).

In goats, the Agouti Signalling Protein (*ASP*), is responsible for white coat because it limits the production of eumelanin by outcompeting the α-melanocyte-stimulating hormone (*α-MSH*) to the binding site of *MC1R,* increasing the production of pheomelanin (Peng et al., 2019; Nasti et al., 2015). When *ASIP* binds to *MC1R* it lowers the production of cAMP, inactivating pigment cells which results in increased pheomelanin (Peng et al., 2019; Nasti et al., 2015). In the absence of *ASIP*, *α-MSH* binds to *MC1R* and leads to increases of intracellular cAMP levels and the production of eumelanin (Peng et al., 2019; Nasti et al., 2015). In this sense, coat pattern in goats is determined by the ratio of eumelanin and pheomelanin. The production and distribution of these two types of melanin are regulated by the several genes involved in melanogenesis such as *ASIP, MC1R, KIT, MITF,* and *TYRP1/2* (Arenas-Báez et al., 2023) as well as melanocyte transport genes such as *MYO5A, RAB27A, MLPH, MATP,* and *SLC24A5*. However, since coat colour is a polygenic trait, that is often influenced by environmental factors such as UV radiation and predetermined genetic factors such as age and ethnicity (Videira et al., 2013), the precise genetic mechanisms of the trait are still not fully understood. Hence, this chapter's first objective was to investigate genes influencing coat colour and coat colour patterns in indigenous South African meat-type goats using genome-wide association studies that utilise SNP data generated by the Illumina Goat 65K SNP genotyping Bead-chip.

In this study, GWAS detected several markers associated with coat colour genes in indigenous South African meat-type goats. For goats with white coat, several genes were detected 1MB

up-and-downstream the marker found on chromosome 4. These included *CDK5* (cyclin-dependent kinase 5) which is involved in Axon guidance, Alzheimer disease, Pathways of neurodegeneration - multiple diseases, and Cocaine addiction. *CDK5* has been revealed to be an essential regulator of several genes which play critical roles in melanogenesis such as *TYR*, *MITF*, and *MC1R* (Dong et al., 2017). In addition, the gene influences melanogenesis signalling pathways such as cAMP (Guan et al. 2011), MAPK (Hisanaga and Endo et al., 2010) and Wnt signalling pathways (Li et al., 2010). In mice, *CDK5* knockdown causes a lightened coat colour while overexpression the gene produces brown patches in sheep (Dong et al., 2017). The detection of the gene in white goats suggests that it may be involved in goat coat colour.

For goats with the white-sided coat colour pattern, a marker associates to the *AHCY* gene, which is involved in Cysteine and methionine metabolism, and Metabolic pathways was detected on chromosome 13. The gene has been identified in a genomic region including *ASIP* and *ITCH* which causes white coat colour in sheep (Norris and Whan et al., 2008). The *PPP2R2B* gene which was detected on chromosome 7 is involved in several pathways including the PI3K-Akt signalling pathway which reduces melanogenesis by downregulating the transcription of tyrosinase and by inactivating the MAPK pathway (Zhou et al., 2017). The inhibition of tyrosinase by the PI3K/Akt pathway results in skin hypopigmentation (Tadokoro et al., 2005) which may explain the white-sided coat colour pattern in goats.

The *TYRP1* gene found in belted and in speckled goats is involved in pathways including Tyrosine metabolism; Metabolic pathways; and Melanogenesis. The upregulation of *TYRP1* by *MITF* has been shown to contribute to black/brown coat colour. In black goats, *TYRP1* shows higher levels of expression in black and brown goats compared to white goats (Bhat et al., 2019). A similar finding was reported for coat colour in sheep (Fan et al., 2013). However, the downregulation of the *TYRP1* gene has been shown to contribute to white phenotypes. In addition, mutations of the gene have been associated with white colour in geese (Wang et al., 2014), and goats (Bhat et al., 2019) as well as the white spotting phenotype in cattle (Liu et al., 2009) and dogs (Rothschild et al., 2006). King et al. (2003) revealed that mutations or dysfunction of *TYRP1* cause extensive loss of pigmentation. It's detection in this study may suggest that it's involved in the white spotting seen in speckled and belted goats.

In addition, a marker linked to *KIT* (Proto-oncogene receptor tyrosine kinase) was detected in goats with the speckled coat colour pattern. The *KIT* gene is involved in several pathways including melanogenesis and has been shown to influence different white colour patterns in

pigs, cats, cows, mice, horses, rabbits, dogs, and camels (Durig et al., 2017; Holl et al., 2017; Haase et al., 2015; David et al., 2014; Fontanesi et al., 2014; Wong et al., 2013; Fontanesi et al., 2010b; Haase et al., 2009; Pielberg et al., 2002; Marklund et al., 1998; Geissler et al., 1988). The genetic mechanism by which the *KIT* gene affects coat colour have been revealed in a few studies. Durkin et al. (2012) revealed that the *KIT* gene is responsible for colour sidedness and white spotting. According to Grosz and MacNeil (1999) the dominant state of *KIT* results in white coats while the expression of the gene in the heterozygous state results in variations in white coats. This has been supported by several studies which have revealed that the *KIT* gene is associated to white spotting in domestic cats (Cooper et al., 2006), colour-sidedness in Swiss cattle breeds (Durkin et al., 2012) and African Nguni cattle (Szczerbal et al., 2017). Bhat et al. (2019) revealed that *KIT* is responsible for white coats in Pashmina goats. While Nazari-Ghadikolaei et al. (2018) found it in white Iranian Markhoz goats. Furthermore, Haase et al. (2009), revealed that a mutation of *KIT* is involved in coat colour variation in horses. Hence, it's detection in this study suggests it is linked to white sidedness in goats.

In goats with the patchy coat colour pattern, a maker associated to the *GNAI3* gene was detected. The gene is involved in several pathways including cAMP signalling pathway, and melanogenesis. In sheep and goats, *GNAI3* has been associated with thermotolerance (Mishra and Mohapatra, 2020). In addition, the gene has been associated with ocular albinism in humans (Young et al., 2016). Mishra and Mohapatra, (2020) revealed that *GNAI3* interacts with the MAPK and PI3K-Akt signalling pathways which are involved in the regulation of melanogenesis. Its detection in this study may implicate it in the loss of pigment seen in the patchy coat colour pattern in goats. In grey goats, the *GSK3B* gene which is involved in several pathways including the PI3K-Akt signalling pathway, Wnt signalling pathway, and Melanogenesis was detected. Khaled et al. (2002) revealed that *GSK3B* is activated by cAMP and interacts with *MITF* to facilitate its binding to the tyrosinase promoter (Khaled et al., 2002) thereby leading to the stimulation of melanogenesis. In this regard, *GSK3B* plays an active role in cAMP-induced melanogenesis and in melanocyte differentiation. Its detection in this study suggests it may be involved in grey coat colour in goats.

In red goats, a marker associated with the *TLE6* gene was detected. The gene is involved in the Wnt signalling pathway, and the Notch signalling pathway. The Wnt signalling pathway is involved in melanogenesis through inducing melanocyte stem cell (MelSC) differentiation into melanocytes (D'mello et al., 2016). On the other hand, the Notch signalling pathway is vital for proper coat pigmentation (Schouwey and Beermann et al., 2008). *TLE6* involvement with

melanogenesis pathways and its detection in this study suggests it may be involved in red coat colour in goats. In black goats, the *CACNA2D1* gene which is involved in several pathways including MAPK signalling pathway, was detected. In mammals, the MAPK pathway regulates melanogenesis (D'Mello et al., 2016; Gupta, 2014). The MAPK pathway was found to lead to the activation of MITF which upregulates the two melanogenesis genes: TYR and TYRP1 (Tadokoro et al., 2005). The detection of CACNA2D1 in black goats suggests the gene likely plays a role in goat coat colour.

In goats with the blackleg coat colour pattern, markers associated to the *AHCY* and *EDNRA* genes were detected. the *AHCY* gene is involved in Cysteine and methionine metabolism, and Metabolic pathways. The gene has also been linked to coat pigmentation (Norris and Whan et al., 2008). The *EDNRA* gene is involved in several pathways including the cAMP signalling pathway. The cAMP pathway upregulates tyrosine expression by inducing the binding of α-MSH to the *MC1R* which leads to the production of produce eumelanin (Natale et al., 2016). Hence, the gene may play a role in the blacklegs coat colour pattern. The markers found in Boer goats were associated with *CADPS2*; and *SLC13A1*, whose biological functions have not yet been revealed in goats. In Savanna goats, candidate markers were linked to *SND1* which is involved in Viral carcinogenesis. While in Kalahari Red goats, the marker was linked to *CHN2* and *TMTC1*.

As coat colour is a polygenic trait whose variation is influenced by the epistatic effects of multiple genes as well as the environment, it can be difficult for GWAS to detect causal variants (Sturm et al., 2001). This study's findings detected genes that have also been found by previous studies which have revealed that over 127 loci influence coat colour in mammals (Bennett and Lamoreux et al., 2003). The genes identified regulate key functions that are vital not only for melanocyte differentiation and migration, but also melanocyte signalling and regulation which all affect melanin production and transport (D'Mello et al., 2016). As such major mechanisms responsible for coat colour variation also include the numerous regulatory pathways that affect melanogenesis such as the MAPK, PI3K-Akt, cAMP, and Wnt signalling pathways (D'Mello et al., 2016). However, the identification and characterisation of the gene-gene interactions responsible for colour variations still need deeper insights for the effective implementation of genomic selection in livestock animals. Furthermore, due to the small sample sizes for each of the coat colour/patterns investigated in this study the analysis needs to be repeated with larger sample sizes to verify findings and draw overall conclusions on the genetic mechanisms responsible for coat colour variation in the broader SA goat population.

4.4.2 Copy number variation

In recent years, CNVs have been revealed to influence the phenotypic diversity of various economically important traits (Zhou et al., 2016). To date, CNVs have been revealed to regulate important physiological functions by influencing gene dosage, interruption, and expression in several livestock animals (Paudel et al., 2015; Liu et al., 2013; Fontanesi et al., 2010a; Liu et al., 2009). For instance, copy number variation is involved in coat pigmentation in sheep, goats, pigs, and cattle (Brenig et al., 2013; Durkin et al., 2012; Norris and Whan, 2008; Pielberg et al., 2002). Moreover, copy number variation regions covering coat colour genes have been shown to influence coat colour variation by affecting the production, and distribution of melanogenesis (Brenig et al., 2013; Fontanesi and Russo, 2013; Fontanesi et al., 2011; Fontanesi et al., 2010a). For instance, CNVRs overlapping the *MC1R* gene were revealed to be responsible for the coat of cattle (Hou et al., 2012; Hou et al., 2011). While CNVs on *ASIP* determine coat colours in sheep and goats (Han et al., 2015; Fontanesi et al., 2011; Norris and Whan, 2008). However, only a few CNV studies have been carried out using SNP array data (Di Gerlando et al., 2020; Liu et al., 2020; Liu et al., 2019). Hence, the second objective of this chapter was to detect copy number variations related to the coat colour and patterns of goats genotyped with the Illumina GoatSNP65K BeadChip.

A total of 2 047 CNVs were detected with an average number of 12.11 CNVs per individual. Among these CNVs, 115 were gains with an average length of 831464 bp and 1 932 were losses with an average length of 927516 bp. From these CNVs, a total of 279 CNVRs with an average length of 1865633 bp were found. Among these, 262 CNVRs overlapped 5 222 genes. The gene ontology analysis carried on these genes revealed that the genes were involved in several biological processes, molecular functions, cellular components, and pathways (Tables 4.12 – 4.15). In addition, several coat colour candidate genes were found overlapping some of the CNVRs (Appendix A). For instance, CNVR_9 (chr1: 63378629 - 64366087) overlapped with *GSK3B* which is involved in ER overload response (GO:0006983), rhythmic process (GO:0048511), negative regulation of extrinsic apoptotic signalling pathway via death domain receptors (GO:1902042), plasma membrane (GO:0005886), nucleoplasm (GO:0005654), centrosome (GO:0005813), and ubiquitin protein ligase binding (GO:0031625). In addition, *GSK3B* plays an active role in cAMP-induced melanogenesis and in melanocyte differentiation (Khaled et al., 2002).

CNVR_49 (chr3: 96069301- 98281805) overlapped with *NOTCH2* which is involved in regulation of transcription DNA-templated (GO:0006355), regulation of developmental process (GO:0050793), marginal zone B cell differentiation (GO:0002315), Notch signalling pathway (GO:0007219), plasma membrane (GO:0005886), integral component of membrane (GO:0016021), nucleoplasm (GO:0005654), and integral component of membrane (GO:0016021). While *NOTCH1* was found on CNVR_131 (chr11: 101402129 -106216353). The gene is involved in regulation of transcription DNA-templated (GO:0006355), negative regulation of anoikis (GO:2000811), regulation of developmental process (GO:0050793), negative regulation of cell migration involved in sprouting angiogenesis (GO:0090051), Notch signalling pathway (GO:0007219), regulation of extracellular matrix assembly (GO:1901201), negative regulation of myoblast differentiation (GO:0045662), plasma membrane (GO:0005886), integral component of membrane (GO:0016021), and identical protein binding (GO:0042802). *NOTCH1*, and *NOTCH2* influence coat colour through the Notch signalling pathway which regulates melanocyte development (Schouwey and Beermann et al., 2008). Furthermore, the deletion of *NOTCH1*, and *NOTCH2* results in a severe coat colour dilution in mice (Schouwey and Beermann et al., 2008).

CNVR_53 (chr4: 50909 - 88316002) overlapped *CDK5* which is involved in cell-matrix adhesion (GO:0007160), cerebellar cortex formation (GO:0021697), plasma membrane (GO:0005886), nucleoplasm (GO:0005654), cytosol (GO:0005829), and centrosome (GO:0005813). In addition, CDK5 has been revealed to be an essential regulator of *MITF* and *PAX3* which play critical roles in melanogenesis (Dong et al., 2017). In mice, *CDK5* knockdown causes a lightened coat colour, while in sheep, overexpression of the gene produces brown patches on a white background (Dong et al., 2017).

CNVR_73 (chr5: 33365253 - 37650260) overlapped with *ADAMTS20* which is involved in negative regulation of apoptotic process (GO:0043066), and metalloendopeptidase activity (GO:0004222). Drögemüller et al. (2009) revealed that *ADAMTS20* was related to cattle coat colour. In goats, the gene was identified as the strong candidate gene under selection for breed diversity (Oget et al., 2018). In addition, *ADAMTS20* regulates melanogenesis in goats (Crepaldi and Nicoloso 2010). Dong et al. (2015) revealed that *ADAMTS20* was involved in goat coat colour variation during domestication. These results were further corroborated by additional studies that identified the gene overlapping with a CNVR in worldwide goat populations (Di Gerlando et al., 2020; Liu et al., 2019). This result is in line with previous

findings and suggests that the *ADAMTS20* CNV is responsible for coat colour variation in goats.

CNVR_115 (chr10: 64866327 - 66185170) overlapped with *TYRO3* which is involved in negative regulation of toll-like receptor signalling pathway (GO:0034122), spermatogenesis (GO:0007283), negative regulation of inflammatory response (GO:0050728), plasma membrane (GO:0005886), and integral component of membrane (GO:0016021). *TYRO3* has been identified as a candidate gene linked to skin pigmentation through regulating *MITF* expression in a *SOX10*-dependent manner in melanoma cells. CNVR_118 (chr10: 87926611 - 88674300) overlapped with *MAP2K1* which is involved in cerebellar cortex formation (GO:0021697), Bergmann glial cell differentiation (GO:0060020), plasma membrane (GO:0005886), and cytosol (GO:0005829). In addition, Gutiérrez-Gil et al. (2015) revealed that this gene is linked to coat colour in cattle.

CNVR_141 (chr13: 63077479 - 63416242) overlapped with *ITCH*, *ASIP*, and *AHCY*. *ITCH* is involved in negative regulation of apoptotic process (GO:0043066), protein K63-linked ubiquitination (GO:0070534), plasma membrane (GO:0005886), and nucleoplasm (GO:0005654). While *AHCY* is involved in cytosol (GO:0005829). *ITCH*, *ASIP*, and *AHCY* genomic regions have been identified in previous studies (Liu et al., 2019; Fontanesi et al., 2009; Norris and Whan. 2008). In sheep, Norris and Whan (2008), revealed that a CNV on genomic regions of *ITCH*, *ASIP*, and *AHCY* caused white coat colour. In goats, Fontanesi et al. (2009) found that variation in coat colour was due to a CNV overlapping with *ASIP* and *AHCY*. In addition, the study by Liu et al. (2019) validated copy number variations of *ASIP*, *AHCY*, and *ITCH* in African indigenous goats.

CNVR_183 (chr17: 59249949 - 62145068) overlapped with *EDNRA* which is involved in plasma membrane (GO:0005886), and integral component of membrane (GO:0016021). The genomic amplification of *EDNRA* is responsible for white spotting (Menzi et al., 2016). In addition, Liu et al. (2019) detected CNVs on coat colour genes such as EDNRA, ASIP, *AHCY*, and *ITCH* in African indigenous goats with different coat colours. CNVR_204 (chr20: 38947614 – 40340043) overlapped with *SLC45A2* which is involved in integral component of membrane (GO:0016021), and sucrose: proton symporter activity (GO:0008506). *SLC45A2* influences coat colour by regulating the synthesis and transportation of melanin (Wang et al., 2015). Furthermore, researchers have determined that mutations in *SLC45A2* can lead to oculocutaneous albinism type IV (Inagaki et al., 2006). These mutations also affect the

pigmentation of mice (Du and Fisher et al., 2002), chickens (Gunnarsson et al., 2007) and horses (Mariat et al., 2003).

CNVR_209 (chr20: 64902194 - 65745399) overlapped with *ADCY2* which is involved in plasma membrane (GO:0005886), integral component of membrane (GO:0016021), and metal ion binding (GO:0046872). *ADCY2* is closely related to melanogenesis and has been shown to influence black beak colour in ducks (Pan et al., 2023). CNVR_229 (chr22: 55234708 - 60193727) overlapped with *WNT7A* which is involved in negative regulation of apoptotic process (GO:0043066), embryonic digit morphogenesis (GO:0042733). Similar to *ADCY2*, *WNT7A* is closely related to melanogenesis (Pan et al., 2023). CNVR_275 (chr29: 12994 - 5900083) overlapped with *TYR* which is involved in integral component of membrane (GO:0016021), and copper ion binding (GO:0005507). *TYR* encodes tyrosinase which catalyses melanogenesis (Sulem et al., 2008). Two mutations in *TYR* have been associated with the Siamese and Burmese coat colour phenotypes (Sulem et al., 2008).

In conclusion, this study identified CNVRs which overlapped with genes related to animal coat colour. These findings suggest a biological link between CNVRs and coat colour which has been associated with the goat adaptability to harsh climatic and environmental conditions. Coat colour influences the economical productivity of indigenous goats. Therefore, this knowledge is vital as it will allow farmers to utilize breeds with specific environmental adaptations in selection strategies for genetic improvement. However, it is noted that the SNP probes used to detect CNVs are neither dense enough nor uniformly distributed to achieve an unbiased high resolution CNV map (Jiang et al., 2013; Di Gerlando et al., 2020). In addition, a portion of goat genome is not annotated, and this could have influenced the results of the gene GO analysis. Finally, differences in the CNVs detected between this study and previous studies may have been caused by the use of different quality control parameter, different CNV detection algorithms, and differences in sample sizes. Therefore, results from this study warrant further validations and investigations in the future.

4.5 References

ARENAS-BÁEZ, P., TORRES-HERNÁNDEZ, G., CASTILLO-HERNÁNDEZ, G., HERNÁNDEZ-RODRÍGUEZ, M., SÁNCHEZ-GUTIÉRREZ, R. A., VARGAS-LÓPEZ, S., GONZÁLEZ-MALDONADO, J., DOMÍNGUEZ-MARTÍNEZ, P. A., GRANADOS-RIVERA, L. D. & MALDONADO-JÁQUEZ, J. A. 2023. Coat Color in Local Goats: Influence on Environmental Adaptation and Productivity, and Use as a Selection Criterion. *Biology,* 12**,** 929.

BECKMANN, J. S., ESTIVILL, X. & ANTONARAKIS, S. E. 2007. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nature Reviews Genetics,* 8**,** 639-646.

BENJELLOUN, B., ALBERTO, F. J., STREETER, I., BOYER, F., COISSAC, E., STUCKI, S., BENBATI, M., IBNELBACHYR, M., CHENTOUF, M., BECHCHARI, A., LEEMPOEL, K., ALBERTI, A., ENGELEN, S., CHIKHI, A., CLARKE, L., FLICEK, P., JOOST, S., TABERLET, P., POMPANON, F. & , N. C. 2015. Characterizing neutral genomic diversity and selection signatures in indigenous populations of Moroccan goats (Capra hircus) using WGS data. *Frontiers in Genetics,* 6.

BENNETT, D. C. & LAMOREUX, M. L. 2003. The Color Loci of Mice – A Genetic Century. *Pigment Cell Research,* 16**,** 333-344.

BHAT, B., SINGH, A., IQBAL, Z., KAUSHIK, J. K., RAO, A., AHMAD, S. M., BHAT, H., AYAZ, A., SHEIKH, F. & KALRA, S. 2019. Comparative transcriptome analysis reveals the genetic basis of coat color variation in Pashmina goat. *Scientific Reports,* 9**,** 6361.

BICKHART, D. M. & LIU, G. E. 2014. The challenges and importance of structural variation detection in livestock. *Front Genet,* 5**,** 37.

BRENIG, B., BECK, J., FLOREN, C., BORNEMANN-KOLATZKI, K., WIEDEMANN, I., HENNECKE, S., SWALVE, H. & SCHÜTZ, E. 2013. Molecular genetics of coat colour variations in White Galloway and White Park cattle. *Animal Genetics,* 44**,** 450-453.

BRITO, L. F., KIJAS, J. W., VENTURA, R. V., SARGOLZAEI, M., PORTO-NETO, L. R., CÁNOVAS, A., FENG, Z., JAFARIKIA, M. & SCHENKEL, F. S. 2017. Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers. *BMC Genomics,* 18**,** 229.

CHEN, C., LIU, C., XIONG, X., FANG, S., YANG, H., ZHANG, Z., REN, J., GUO, Y. & HUANG, L. 2018. Copy number variation in the MSRB3 gene enlarges porcine ear size through a mechanism involving miR-584-5p. *Genet Sel Evol,* 50**,** 72.

CLOP, A., VIDAL, O. & AMILLS, M. 2012. Copy number variation in the genomes of domestic animals. *Anim Genet,* 43**,** 503-17.

COLLI, L., MILANESI, M., TALENTI, A., BERTOLINI, F., CHEN, M., CRISÀ, A., DALY, K. G., DEL CORVO, M., GULDBRANDTSEN, B. & LENSTRA, J. A. 2018. Genome-wide SNP profiling of worldwide goat populations reveals strong partitioning of diversity and highlights post-domestication migration routes. *Genetics Selection Evolution,* 50**,** 1-20.

COOPER, M., FRETWELL, N., BAILEY, S. & LYONS, L. 2006. White spotting in the domestic cat (Felis catus) maps near KIT on feline chromosome B1. *Animal genetics,* 37**,** 163-165.

D'MELLO, S. A., FINLAY, G. J., BAGULEY, B. C. & ASKARIAN-AMIRI, M. E. 2016. Signaling Pathways in Melanogenesis. *Int J Mol Sci,* 17.

DA SILVA, J. M., GIACHETTO, P. F., DA SILVA, L. O., CINTRA, L. C., PAIVA, S. R., YAMAGISHI, M. E. B. & CAETANO, A. R. 2016. Genome-wide copy number variation (CNV) detection in Nelore cattle reveals highly frequent variants in genome regions harboring QTLs affecting production traits. *BMC genomics,* 17**,** 1-14.

DAVID, V. A., MENOTTI-RAYMOND, M., WALLACE, A. C., ROELKE, M., KEHLER, J., LEIGHTY, R., EIZIRIK, E., HANNAH, S. S., NELSON, G. & SCHÄFFER, A. A. 2014. Endogenous retrovirus insertion in the KIT oncogene determines white and white spotting in domestic cats. *G3: Genes, Genomes, Genetics,* 4**,** 1881-1891.

DEKKERS, J. C. 2004. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. *J Anim Sci,* 82 E-Suppl**,** E313-328.

DI GERLANDO, R., MASTRANGELO, S., MOSCARELLI, A., TOLONE, M., SUTERA, A. M., PORTOLANO, B. & SARDINA, M. T. 2020. Genomic Structural Diversity in Local Goats: Analysis of Copy-Number Variations. *Animals,* 10**,** 1040.

DI GERLANDO, R., SARDINA, M. T., TOLONE, M., SUTERA, A. M., MASTRANGELO, S. & PORTOLANO, B. 2018. Genome-wide detection of copy-number variations in local cattle breeds. *Animal Production Science,* 59**,** 815-822.

DI GERLANDO, R., SUTERA, A. M., MASTRANGELO, S., TOLONE, M., PORTOLANO, B., SOTTILE, G., BAGNATO, A., STRILLACCI, M. G. & SARDINA, M. T. 2019. Genome-wide association study between CNVs and milk production traits in Valle del Belice sheep. *PLoS One,* 14**,** e0215204.

DONG, C., YANG, S., FAN, R., JI, K., ZHANG, J., LIU, X., HU, S., XIE, J., LIU, Y., GAO, W., WANG, H., YAO, J., SMITH, G. W. & HERRID, M. 2017. Functional Role of Cyclin-Dependent Kinase 5 in the Regulation of Melanogenesis and Epidermal Structure. *Scientific Reports,* 7**,** 13783.

DONG, Y., ZHANG, X., XIE, M., AREFNEZHAD, B., WANG, Z., WANG, W., FENG, S., HUANG, G., GUAN, R., SHEN, W., BUNCH, R., MCCULLOCH, R., LI, Q., LI, B., ZHANG, G., XU, X., KIJAS, J. W., SALEKDEH, G. H., WANG, W. & JIANG, Y. 2015. Reference genome of wild goat (capra aegagrus) and sequencing of goat breeds provide insight into genic basis of goat domestication. *BMC Genomics,* 16**,** 431.

DRÖGEMÜLLER, C., ENGENSTEINER, M., MOSER, S., RIEDER, S. & LEEB, T. 2009. Genetic mapping of the belt pattern in Brown Swiss cattle to BTA3. *Animal Genetics,* 40**,** 225-229.

DU, J. & FISHER, D. E. 2002. Identification of Aim-1 as the underwhite mouse mutant and its transcriptional regulation by MITF. *J Biol Chem,* 277**,** 402-6.

DÜRIG, N., JUDE, R., HOLL, H., BROOKS, S. A., LAFAYETTE, C., JAGANNATHAN, V. & LEEB, T. 2017. Whole genome sequencing reveals a novel deletion variant in the KIT gene in horses with white spotted coat colour phenotypes. *Animal genetics,* 48**,** 483-485.

DURKIN, K., COPPIETERS, W., DRÖGEMÜLLER, C., AHARIZ, N., CAMBISANO, N., DRUET, T., FASQUELLE, C., HAILE, A., HORIN, P. & HUANG, L. 2012. Serial translocation by means of circular intermediates underlies colour sidedness in cattle. *Nature,* 482**,** 81-84.

FAN, R., XIE, J., BAI, J., WANG, H., TIAN, X., BAI, R., JIA, X., YANG, L., SONG, Y., HERRID, M., GAO, W., HE, X., YAO, J., SMITH, G. W. & DONG, C. 2013. Skin transcriptome profiles associated with coat color in sheep. *BMC Genomics,* 14**,** 389.

FEUK, L., CARSON, A. R. & SCHERER, S. W. 2006. Structural variation in the human genome. *Nature Reviews Genetics,* 7**,** 85-97.

FONTANESI, L., BERETTI, F., MARTELLI, P., COLOMBO, M., DALL'OLIO, S., OCCIDENTE, M., PORTOLANO, B., CASADIO, R., MATASSINO, D. & RUSSO, V. 2011. A first comparative map of copy number variations in the sheep genome. *Genomics,* 97**,** 158-165.

FONTANESI, L., BERETTI, F., RIGGIO, V., DALL'OLIO, S., GONZÁLEZ, E. G., FINOCCHIARO, R., DAVOLI, R., RUSSO, V. & PORTOLANO, B. 2009. Missense and nonsense mutations in melanocortin 1 receptor (MC1R) gene of different goat breeds: association with red and black coat colour phenotypes but with unexpected evidences. *BMC Genetics,* 10**,** 47.

FONTANESI, L., MARTELLI, P. L., BERETTI, F., RIGGIO, V., DALL'OLIO, S., COLOMBO, M., CASADIO, R., RUSSO, V. & PORTOLANO, B. 2010a. An initial comparative map of copy number variations in the goat (Capra hircus) genome. *BMC genomics,* 11**,** 1-15.

FONTANESI, L. & RUSSO, V. 2013. Molecular genetics of coat colour in pigs. *Acta Agric. Slov.,* 4.

FONTANESI, L., TAZZOLI, M., RUSSO, V. & BEEVER, J. 2010b. Genetic heterogeneity at the bovine KIT gene in cattle breeds carrying different putative alleles at the spotting locus. *Animal genetics,* 41**,** 295-303.

FONTANESI, L., VARGIOLU, M., SCOTTI, E., LATORRE, R., FAUSSONE PELLEGRINI, M. S., MAZZONI, M., ASTI, M., CHIOCCHETTI, R., ROMEO, G. & CLAVENZANI, P. 2014. The KIT gene is associated with the English spotting coat color locus and congenital megacolon in Checkered Giant rabbits (Oryctolagus cuniculus). *PLoS One,* 9**,** e93750.

GEISSLER, E. N., RYAN, M. A. & HOUSMAN, D. E. 1988. The dominant-white spotting (W) locus of the mouse encodes the c-kit proto-oncogene. *Cell,* 55**,** 185-192.

GIBBS, S., MURLI, S., DE BOER, G., MULDER, A., MOMMAAS, A. M. & PONEC, M. 2000. Melanosome Capping of Keratinocytes in Pigmented Reconstructed Epidermis – Effect of Ultraviolet Radiation and 3-Isobutyl-1-Methyl-Xanthine on Melanogenesis. *Pigment Cell Research,* 13**,** 458-466.

GIUFFRA, E., TÖRNSTEN, A., MARKLUND, S., BONGCAM-RUDLOFF, E., CHARDON, P., KIJAS, J. M., ANDERSON, S. I., ARCHIBALD, A. L. & ANDERSSON, L. 2002. A large duplication associated with dominant white color in pigs originated by homologous recombination between LINE elements flanking KIT. *Mamm Genome,* 13**,** 569-77.

GROSZ, M. & MACNEIL, M. 1999. The" spotted" locus maps to bovine chromosome 6 in a Hereford-Cross population. *The Journal of heredity,* 90**,** 233-236.

GUAN, D., MARTÍNEZ, A., CASTELLÓ, A., LANDI, V., LUIGI-SIERRA, M. G., FERNÁNDEZ-ÁLVAREZ, J., CABRERA, B., DELGADO, J. V., SUCH, X. & JORDANA, J. 2020. A genome-wide analysis of copy number variation in Murciano-Granadina goats. *Genetics Selection Evolution,* 52**,** 1-10.

GUAN, J. S., SU, S. C., GAO, J., JOSEPH, N., XIE, Z., ZHOU, Y., DURAK, O., ZHANG, L., ZHU, J. J., CLAUSER, K. R., CARR, S. A. & TSAI, L. H. 2011. Cdk5 is required for memory function and hippocampal plasticity via the cAMP signaling pathway. *PLoS One,* 6**,** e25735.

GUNNARSSON, U., HELLSTRÖM, A. R., TIXIER-BOICHARD, M., MINVIELLE, F., BED'HOM, B., ITO, S., JENSEN, P., RATTINK, A., VEREIJKEN, A. & ANDERSSON, L. 2007. Mutations in SLC45A2 cause plumage color variation in chicken and Japanese quail. *Genetics,* 175**,** 867-77.

GUPTA, S. 2014. Skin colour: No hiding in the dark. *Nature,* 515**,** S121-S123.

GUTIÉRREZ-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 signals across breeds. *Front Genet,* 6**,** 167.

HAASE, B., BROOKS, S. A., TOZAKI, T., BURGER, D., PONCET, P.-A., RIEDER, S., HASEGAWA, T., PENEDO, C. & LEEB, T. 2009. Seven novel KIT mutations in horses with white coat colour phenotypes. *Animal Genetics,* 40**,** 623-629.

HAASE, B., RIEDER, S. & LEEB, T. 2015. Two variants in the KIT gene as candidate causative mutations for a dominant white and a white spotting phenotype in the donkey. *Animal genetics,* 46**,** 321-324.

HAN, J.-L., MIN, Y., GUO, T.-T., YUE, Y.-J., LIU, J.-B., NIU, C.-E., WANG, C.-F. & YANG, B.-H. 2015. Molecular characterization of two candidate genes associated with coat color in Tibetan sheep (Ovis arise). *Journal of integrative agriculture,* 14**,** 1390-1397.

HAYES, B. J., BOWMAN, P. J., CHAMBERLAIN, A. J. & GODDARD, M. E. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of Dairy Science,* 92**,** 433-443.

HENKEL, J., SAIF, R., JAGANNATHAN, V., SCHMOCKER, C., ZEINDLER, F., BANGERTER, E., HERREN, U., POSANTZIS, D., BULUT, Z. & AMMANN, P. 2019. Selection signatures in goats reveal copy number variants underlying breed-defining coat color phenotypes. *PLoS genetics,* 15**,** e1008536.

HISANAGA, S.-I. & ENDO, R. 2010. Regulation and role of cyclin-dependent kinase activity in neuronal survival and death. *Journal of Neurochemistry,* 115**,** 1309-1321.

HOLL, H., ISAZA, R., MOHAMOUD, Y., AHMED, A., ALMATHEN, F., YOUCEF, C., GAOUAR, S., ANTCZAK, D. F. & BROOKS, S. 2017. A frameshift mutation in KIT is associated with white spotting in the Arabian camel. *Genes,* 8**,** 102.

HOU, Y., LIU, G. E., BICKHART, D. M., CARDONE, M. F., WANG, K., KIM, E.-S., MATUKUMALLI, L. K., VENTURA, M., SONG, J. & VANRADEN, P. M. 2011. Genomic characteristics of cattle copy number variations. *BMC genomics,* 12**,** 1-11.

HOU, Y., LIU, G. E., BICKHART, D. M., MATUKUMALLI, L. K., LI, C., SONG, J., GASBARRE, L. C., VAN TASSELL, C. P. & SONSTEGARD, T. S. 2012. Genomic regions showing copy number variations associate with resistance or susceptibility to gastrointestinal nematodes in Angus cattle. *Functional & Integrative Genomics,* 12**,** 81-92.

HUANG, D. W., SHERMAN, B. T., TAN, Q., COLLINS, J. R., ALVORD, W. G., ROAYAEI, J., STEPHENS, R., BASELER, M. W., LANE, H. C. & LEMPICKI, R. A. 2007. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome biology,* 8**,** 1-16.

INAGAKI, K., SUZUKI, T., ITO, S., SUZUKI, N., ADACHI, K., OKUYAMA, T., NAKATA, Y., SHIMIZU, H., MATSUURA, H., OONO, T., IWAMATSU, H., KONO, M. & TOMITA, Y. 2006. Oculocutaneous albinism type 4: six novel mutations in the membrane-associated transporter protein gene and their phenotypes. *Pigment Cell Res,* 19**,** 451-3.

JIANG, L., JIANG, J., YANG, J., LIU, X., WANG, J., WANG, H., DING, X., LIU, J. & ZHANG, Q. 2013. Genome-wide detection of copy number variations using high-density SNP genotyping platforms in Holsteins. *BMC Genomics,* 14**,** 131.

KANEHISA, M. & SATO, Y. 2020. KEGG Mapper for inferring cellular functions from protein sequences. *Protein Science,* 29**,** 28-35.

KHALED, M., LARRIBERE, L., BILLE, K., ABERDAM, E., ORTONNE, J. P., BALLOTTI, R. & BERTOLOTTO, C. 2002. Glycogen synthase kinase 3beta is activated by cAMP and plays an active role in the regulation of melanogenesis. *J Biol Chem,* 277**,** 33690-7.

KING, R. A., WILLAERT, R. K., SCHMIDT, R. M., PIETSCH, J., SAVAGE, S., BROTT, M. J., FRYER, J. P., SUMMERS, C. G. & OETTING, W. S. 2003. MC1R mutations modify the classic phenotype of oculocutaneous albinism type 2 (OCA2). *The American Journal of Human Genetics,* 73**,** 638-645.

LAMOREUX, M. L., WAKAMATSU, K. & ITO, S. 2001. Interaction of major coat color gene functions in mice as studied by chemical analysis of eumelanin and pheomelanin. *Pigment Cell Research,* 14**,** 23-31.

LI, Q., LIU, X., ZHANG, M., YE, G., QIAO, Q., LING, Y., WU, Y., ZHANG, Y. & YU, L. 2010. Characterization of a novel human CDK5 splicing variant that inhibits Wnt/β-catenin signaling. *Molecular biology reports,* 37**,** 2415-2421.

LIN, J. Y. & FISHER, D. E. 2007. Melanocyte biology and skin pigmentation. *Nature,* 445**,** 843-850.

LIPKA, A. E., TIAN, F., WANG, Q., PEIFFER, J., LI, M., BRADBURY, P. J., GORE, M. A., BUCKLER, E. S. & ZHANG, Z. 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics,* 28**,** 2397-2399.

LIU, J., ZHANG, L., XU, L., REN, H., LU, J., ZHANG, X., ZHANG, S., ZHOU, X., WEI, C. & ZHAO, F. 2013. Analysis of copy number variations in the sheep genome using 50K SNP BeadChip array. *BMC genomics,* 14**,** 1-11.

LIU, L., HARRIS, B., KEEHAN, M. & ZHANG, Y. 2009. Genome scan of pigmentation traits in Friesian-Jersey crossbred cattle. *Journal of Genetics and Genomics,* 36**,** 661-666.

LIU, M., WOODWARD-GREENE, J., KANG, X., PAN, M. G., ROSEN, B., VAN TASSELL, C. P., CHEN, H. & LIU, G. E. 2020. Genome-wide CNV analysis revealed variants associated with growth traits in African indigenous goats. *Genomics,* 112**,** 1477-1480.

LIU, M., ZHOU, Y., ROSEN, B. D., VAN TASSELL, C. P., STELLA, A., TOSSER-KLOPP, G., RUPP, R., PALHIÈRE, I., COLLI, L. & SAYRE, B. 2019. Diversity of copy number variation in the worldwide goat population. *Heredity,* 122**,** 636-646.

MAPHOLI, N. O. 2015. *Exploring genetic architecture of tick resistance in South African Nguni cattle.* Stellenbosch: Stellenbosch University.

MARIAT, D., TAOURIT, S. & GUÉRIN, G. 2003. A mutation in the MATP gene causes the cream coat colour in the horse. *Genetics Selection Evolution,* 35**,** 119.

MARKLUND, S., KIJAS, J., RODRIGUEZ-MARTINEZ, H., RÖNNSTRAND, L., FUNA, K., MOLLER, M., LANGE, D., EDFORS-LILJA, I. & ANDERSSON, L. 1998. Molecular basis for the dominant white phenotype in the domestic pig. *Genome research,* 8**,** 826-833.

MARTIN, P. M., PALHIÈRE, I., RICARD, A., TOSSER-KLOPP, G. & RUPP, R. 2016. Genome Wide Association Study Identifies New Loci Associated with Undesired Coat Color Phenotypes in Saanen Goats. *PloS one,* 11**,** e0152426-e0152426.

MATUKUMALLI, L. K., LAWLEY, C. T., SCHNABEL, R. D., TAYLOR, J. F., ALLAN, M. F., HEATON, M. P., O'CONNELL, J., MOORE, S. S., SMITH, T. P., SONSTEGARD, T. S. & VAN TASSELL, C. P. 2009. Development and characterization of a high density SNP genotyping assay for cattle. *PLoS One,* 4**,** e5350.

MDLADLA, K., DZOMBA, E. F., HUSON, H. J. & MUCHADEYI, F. C. 2016. Population genomic structure and linkage disequilibrium analysis of South African goat breeds using genome-wide SNP data. *Anim Genet,* 47**,** 471-82.

MDLADLA, K., DZOMBA, E. F. & MUCHADEYI, F. C. 2017. Characterization of the village goat production systems in the rural communities of the Eastern Cape, KwaZulu-Natal, Limpopo and North West Provinces of South Africa. *Tropical Animal Health and Production,* 49**,** 515-527.

MDLADLA, K., DZOMBA, E. F. & MUCHADEYI, F. C. 2018. Landscape genomics and pathway analysis to understand genetic adaptation of South African indigenous goat populations. *Heredity (Edinb),* 120**,** 369-378.

MENZI, F., KELLER, I., REBER, I., BECK, J., BRENIG, B., SCHÜTZ, E., LEEB, T. & DRÖGEMÜLLER, C. 2016. Genomic amplification of the caprine EDNRA locus might lead to a dose dependent loss of pigmentation. *Sci Rep,* 6**,** 28438.

MISHRA, C. & MOHAPATRA, S. 2020. An integrative bioinformatics analysis of caprine GNAI3 gene. *Journal of Genetics,* 99.

MOELLMANN, G., SLOMINSKI, A., KUKLINSKA, E. & LERNER, A. B. 1988. Regulation of melanogenesis in melanocytes. *Pigment Cell Research,* 1**,** 79-87.

MONAU, P. I., VISSER, C., NSOSO, S. & VAN MARLE-KÖSTER, E. 2018. Phenotypic and genetic characterization of indigenous Tswana goats. *South African Journal of Animal Science,* 48**,** 925-934.

NASTI, T. H. & TIMARES, L. 2015. MC1R, Eumelanin and Pheomelanin: Their Role in Determining the Susceptibility to Skin Cancer. *Photochemistry and Photobiology,* 91**,** 188-200.

NATALE, C. A., DUPERRET, E. K., ZHANG, J., SADEGHI, R., DAHAL, A., O'BRIEN, K. T., COOKSON, R., WINKLER, J. D. & RIDKY, T. W. 2016. Sex steroids regulate skin pigmentation through nonclassical membrane-bound receptors. *Elife,* 5.

NAZARI-GHADIKOLAEI, A., MEHRABANI-YEGANEH, H., MIAREI-AASHTIANI, S. R., STAIGER, E. A., RASHIDI, A. & HUSON, H. J. 2018. Genome-Wide Association Studies Identify Candidate Genes for Coat Color and Mohair Traits in the Iranian Markhoz Goat. *Front Genet,* 9**,** 105.

NCUBE, K. T., HADEBE, K., DZOMBA, E. F., SOMA, P., FRYLINCK, L. & MUCHADEYI, F. C. 2020. Relationship between population genomic structure and growth profiles of South African goats under different production systems. *Trop Anim Health Prod,* 52**,** 1277-1286.

NORRIS, B. J. & WHAN, V. A. 2008. A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. *Genome Res,* 18**,** 1282-93.

OGET, C., SERVIN, B. & PALHIÈRE, I. 2019. Genetic diversity analysis of French goat populations reveals selective sweeps involved in their differentiation. *Animal Genetics,* 50**,** 54-63.

ONZIMA, R. B., UPADHYAY, M. R., MUKIIBI, R., KANIS, E., GROENEN, M. A. M. & CROOIJMANS, R. P. M. A. 2018. Genome-wide population structure and admixture analysis reveals weak differentiation among Ugandan goat breeds. *Animal Genetics,* 49**,** 59-70.

PAILHOUX, E., VIGIER, B., CHAFFAUX, S., SERVEL, N., TAOURIT, S., FURET, J. P., FELLOUS, M., GROSCLAUDE, F., CRIBIU, E. P., COTINOT, C. & VAIMAN, D. 2001. A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nat Genet,* 29**,** 453-8.

PAN, R., HUA, T., GUO, Q., BAI, H., JIANG, Y., WANG, Z., BI, Y., CHEN, G., WU, X. & CHANG, G. 2023. Identification of SNPs in MITF associated with beak color of duck. *Front Genet,* 14**,** 1161396.

PAOLA, C. & L, N. 2010. SNPs in coat colour genes in goats. *Italian Journal of Animal Science,* 6.

PAUDEL, Y., MADSEN, O., MEGENS, H.-J., FRANTZ, L. A., BOSSE, M., CROOIJMANS, R. P. & GROENEN, M. A. 2015. Copy number variation in the speciation of pigs: a possible prominent role for olfactory receptors. *BMC genomics,* 16**,** 1-14.

PENG, Y., WANG, Y., WANG, R., GENG, L., MA, R., ZHANG, C., LIU, Z., GONG, Y., LI, J. & LI, X. 2019. Exploring differentially expressed genes associated with coat color in goat skin using RNA-seq. *Canadian Journal of Animal Science,* 99**,** 357-366.

PETERS, K. J. 1980. The importance of coat colour and coat type for productive performance of beef cattle under subtropical conditions. *Annales de Génétique et de Sélection Animale,* 12**,** 131-132.

PIELBERG, G., OLSSON, C., SYVÄNEN, A.-C. & ANDERSSON, L. 2002. Unexpectedly high allelic diversity at the KIT locus causing dominant white color in the domestic pig. *Genetics,* 160**,** 305-311.

R CORE TEAM. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

ROTHSCHILD, M. F., VAN CLEAVE, P. S., GLENN, K. L., CARLSTROM, L. P. & ELLINWOOD, N. M. 2006. Association of MITF with white spotting in Beagle crosses and Newfoundland dogs. *Animal Genetics,* 37**,** 606-607.

SCHOUWEY, K. & BEERMANN, F. 2008. The Notch pathway: hair graying and pigment cell homeostasis. *Histol Histopathol,* 23**,** 609-19.

STELLA, A., NICOLAZZI, E. L., VAN TASSELL, C. P., ROTHSCHILD, M. F., COLLI, L., ROSEN, B. D., SONSTEGARD, T. S., CREPALDI, P., TOSSER-KLOPP, G., JOOST, S., AMILLS, M., AJMONE-MARSAN, P., BERTOLINI, F., BOETTCHER, P., BOYLE ONZIMA, R., BRADLEY, D., BUJA, D., CANO PEREIRA, M. E., CARTA, A., CATILLO, G., COLLI, L., CREPALDI, P., CRISÀ, A., DEL CORVO, M., DALY, K., DROEGEMUELLER, C., DURUZ, S., ELBELTAGI, A., ESMAILIZADEH, A., FACO, O., FIGUEIREDO CARDOSO, T., FLURY, C., GARCIA, J. F., GULDBRANDTSEN, B., HAILE, A., HALLSTEINN HALLSSON, J., HEATON, M., HUNNICKE NIELSEN, V., HUSON, H., JOOST, S., KIJAS, J., LENSTRA, J. A., MARRAS, G., MILANESI, M., MINHUI, C., MOAEEN-UD-DIN, M., MORRY O'DONNELL, R., MOSES DANLAMI, O., MWACHARO, J., NICOLAZZI, E. L., PALHIÈRE, I., PILLA, F., POLI, M., REECY, J., RISCHKOWSKY, B. A., ROCHAT, E., ROSEN, B., ROTHSCHILD, M., RUPP, R., SAYRE, B., SERVIN, B., SILVA, K., SONSTEGARD, T., SPANGLER, G., STELLA, A., STERI, R., TALENTI, A., TORTEREAU, F., TOSSER-KLOPP, G., VAJANA, E., VAN TASSELL, C. P., ZHANG, W. & THE ADAPTMAP, C. 2018. AdaptMap: exploring goat diversity and adaptation. *Genetics Selection Evolution,* 50**,** 61.

STURM, R. A., TEASDALE, R. D. & BOX, N. F. 2001. Human pigmentation genes: identification, structure and consequences of polymorphic variation. *Gene,* 277**,** 49-62.

SULEM, P., GUÐBJARTSSON, D., STACEY, S., HELGASON, A., RAFNAR, T., MAGNUSSON, K., MANOLESCU, A., KARASON, A., PALSSON, A., THORLEIFSSON, G., JAKOBSDOTTIR, M., STEINBERG, S., PÁLSSON, S., JÓNASSON, F., SIGURGEIRSSON, B., THORISDOTTIR, K., RAGNARSSON, R., BENEDIKTSDOTTIR, K., ABEN, K. & STEFANSSON, K. 2008. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nature genetics,* 39**,** 1443-52.

SUNDSTRÖM, E., IMSLAND, F., MIKKO, S., WADE, C., SIGURDSSON, S., PIELBERG, G. R., GOLOVKO, A., CURIK, I., SELTENHAMMER, M. H., SÖLKNER, J., LINDBLAD-TOH, K. & ANDERSSON, L. 2012. Copy number expansion of the STX17 duplication in melanoma tissue from Grey horses. *BMC Genomics,* 13**,** 365.

SZCZERBAL, I., ROBINSON, T. J., DURAN, A., DAVEY, S. C., ANDERSSON, M. & SWITONSKI, M. 2017. Ectopic position of duplicated KIT gene in African Nguni cattle, associated with color sidedness, confirms its shared ancestry with the Bos taurus lineage. *Animal Genetics,* 48**,** 122-123.

TADOKORO, T., OZAWA, K., BONNET-DUQUENNOY, M. & BONTÉ, F. 2005. Hormone influence on melanogenesis and spots formation. *International Journal of Cosmetic Science,* 27**,** 56-59.

THODY, A. J., HIGGINS, E. M., WAKAMATSU, K., ITO, S., BURCHILL, S. A. & MARKS, J. M. 1991. Pheomelanin as well as Eumelanin Is Present in Human Epidermis. *Journal of Investigative Dermatology,* 97**,** 340-344.

VANRADEN, P., VAN TASSELL, C., WIGGANS, G., SONSTEGARD, T., SCHNABEL, R., TAYLOR, J. & SCHENKEL, F. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. *Journal of dairy science,* 92**,** 16-24.

VIDEIRA, I. F., MOURA, D. F. & MAGINA, S. 2013. Mechanisms regulating melanogenesis. *An Bras Dermatol,* 88**,** 76-83.

WANG, M. D., DZAMA, K., HEFER, C. A. & MUCHADEYI, F. C. 2015. Genomic population structure and prevalence of copy number variations in South African Nguni cattle. *BMC genomics,* 16**,** 1-16.

WANG, Y., LI, S. M., HUANG, J., CHEN, S. Y. & LIU, Y. P. 2014. Mutations of TYR and MITF Genes are Associated with Plumage Colour Phenotypes in Geese. *Asian-Australas J Anim Sci,* 27**,** 778-83.

WICKHAM, H., CHANG, W. & WICKHAM, M. H. 2016. Package 'ggplot2'. *Create elegant data visualisations using the grammar of graphics. Version,* 2**,** 1-189.

WIGGANS, G., VANRADEN, P. & COOPER, T. 2011. The genomic evaluation system in the United States: Past, present, future. *Journal of dairy science,* 94**,** 3202-3211.

WONG, A., RUHE, A., ROBERTSON, K., LOEW, E., WILLIAMS, D. & NEFF, M. 2013. A de novo mutation in KIT causes white spotting in a subpopulation of G erman S hepherd dogs. *Animal genetics,* 44**,** 305-310.

WRIGHT, D., BOIJE, H., MEADOWS, J. R., BED'HOM, B., GOURICHON, D., VIEAUD, A., TIXIER-BOICHARD, M., RUBIN, C. J., IMSLAND, F., HALLBÖÖK, F. & ANDERSSON, L. 2009. Copy number variation in intron 1 of SOX5 causes the Pea-comb phenotype in chickens. *PLoS Genet,* 5**,** e1000512.

YAMAGUCHI, Y., BRENNER, M. & HEARING, V. J. 2007. The regulation of skin pigmentation. *Journal of biological chemistry,* 282**,** 27557-27561.

YOUNG, A., DANDEKAR, U., PAN, C., SADER, A., ZHENG, J. J., LEWIS, R. A. & FARBER, D. B. 2016. GNAI3: another candidate gene to screen in persons with ocular albinism. *Plos one,* 11**,** e0162273.

ZARREI, M., MACDONALD, J. R., MERICO, D. & SCHERER, S. W. 2015. A copy number variation map of the human genome. *Nat Rev Genet,* 16**,** 172-83.

ZHOU, J., LIU, L., LOPDELL, T. J., GARRICK, D. J. & SHI, Y. 2021. HandyCNV: standardized summary, annotation, comparison, and visualization of copy number variant, copy number variation region, and runs of homozygosity. *Frontiers in Genetics,* 12**,** 731355.

ZHOU, J., REN, T., LI, Y., CHENG, A., XIE, W., XU, L., PENG, L., LIN, J., LIAN, L. & DIAO, Y. 2017. Oleoylethanolamide inhibits α-melanocyte stimulating hormone-stimulated melanogenesis via ERK, Akt and CREB signaling pathways in B16 melanoma cells. *Oncotarget,* 8**,** 56868.

ZHOU, Y., UTSUNOMIYA, Y. T., XU, L., HAY EL, H. A., BICKHART, D. M., ALEXANDRE, P. A., ROSEN, B. D., SCHROEDER, S. G., CARVALHEIRO, R., DE REZENDE NEVES, H. H., SONSTEGARD, T. S., VAN TASSELL, C. P., FERRAZ, J. B., FUKUMASU, H., GARCIA, J. F. & LIU, G. E. 2016. Genome-wide CNV analysis reveals variants associated with growth traits in Bos indicus. *BMC Genomics,* 17**,** 419.

Chapter 5: Genome-wide detection of signatures of selection for coat colour in South African indigenous meat-type goats with various coat colours and patterns.

Abstract

The detection of coat colour selection signatures within the South African goats offers an opportunity to unravel indigenous goat adaptation to adverse climatic and environmental conditions, which will facilitate the implementation of improved breeding schemes. However, for South African goats, such information is lacking. In this regard, the objective of this chapter was to detect signatures of selection for coat colour genes under selection within (iHS) and between (XP-EHH) South African commercial breeds such as the Boer (n = 97), Savanna (n = 38), and Kalahari Red (n = 52), as well as village goats with various coat colours and patterns including plain coats such as white (n = 25), red (n = 17), black (n = 19), grey (n = 10), and coat colour patterns such as speckled (n = 15), patchy (n = 15), belted (n = 16), white-sided (n = 14), and black-legs (n = 11). Candidate genes under selection were found to be involved in pathways associated with melanin production including the MAPK signalling pathway (e.g., *DUSP16, KDR, FGF10, MAP3K7*), the PI3K-Akt pathway (e.g., *CDKN1B, COL6A3, CRK, CREB5*), the Wnt signalling pathway (e.g., *SERPINF1, WNT2*), the Notch signalling pathway (e.g., *HEY2*), and Melanogenesis (e.g., *TYRP1*) suggesting that coat colour is under selection in communal goat production systems. Furthermore, genes associated with various traits important for goat productivity and adaptation were detected. For instance, within commercial breeds, genes involved in metabolism (*BAAT, GRIN3A*), and immune response (*IKZF2, PPP2R2B, TLERG1, ARAP3*) were detected suggesting that artificial selection pressures within intensive production systems target goat productivity. While, within, the ecotype population several genes involved in metabolism (e.g., *FXN, FTO, IRX3, EDNRA*), reproduction (e.g., *PGR*), growth and development (e.g., *SMARCAD1*), immune response (e.g., *ACAD8, DLG4, GPS2*), and environmental adaptation (e.g., *CNR1*) were detected, suggesting that natural selection pressures within extensive communal production systems target goat adaptability to adverse environmental conditions.

Keywords: melanogenesis, coat colour, goat adaptation, selection signatures, improved breeding schemes

5.1 Introduction

South Africa's indigenous goat populations found in extensive production systems, and defined by having heterogeneous phenotypes, including skin and coat coloration represent a precious genetic reservoir (Sevane et al., 2018). However, the prevalence of indiscriminate crossbreeding without proper breeding programmes with defined breeding goals presents a serious risk to these populations, by limiting their genetic potential for improvement (Rajaganapathy et al., 2018; Ojo et al., 2015). Because of this, the implementation of improved selection and breeding schemes is one of the main issues faced by farmers in communal production systems (Gandini et al., 2017). In some instances, farmers have explored alternative ways for the selection of replacement animals including using coat colour as a selection criterion (Adalsteinsson et al., 1994), as the trait is easy to recognise and record. Furthermore, the trait has historically been used as a phenotypical selection model (Akis et al., 2012) and as a key feature in breed identification (Henkel et al., 2019). However, the related information needed for the implementation of improved breeding schemes is still lacking in communal production systems (Becerril et al., 1996).

Coat colour has been demonstrated as an indicator trait for livestock adaptation to different climates (Onasanya et al., 2018). This is because a difference in thermoregulatory response has been reported between different coat colours, with darker coats reported to absorb more heat than lighter coats which in contrast reflect between 50-60% of direct solar radiation (Ferreira et al., 2021; Joy et al., 2020a,b). This difference suggests that goats with different coat colours possess different thermoregulatory abilities, such as heat tolerance (Joy et al., 2020a,b) with studies reporting an increased adaptation to cold weather conditions in black goats, compared to white goats (Ferreira et al., 2021; Chokoe et al., 2020a). Such differences in goat adaptation due to coat colour has led to differences in farmer preferences. For instance, some farmers choose to select for black coats over white because they believe that white goats are more susceptible to attack by predators and theft, having a higher risk when goats wander away from the flock (Torres-Hernández et al., 2022; Berhanu et al., 2012).

In village goat production systems in South Africa a similar case of selection based on farmer preference for coat colour has been reported with different coat colours having different socio-cultural and economic values (Rumosa Gwaze et al. 2008; Mahanjana & Cronje 2000). For instance, findings have suggested that the prevalence of white coat colour in the Eastern Cape is a result of selection for a white colour in these communities due to the higher demands for white goats for use in cultural ceremonies (Mahanjana and Cronje., 2000). While a prevalence

of black coat colour Limpopo and KwaZulu-Natal has been reported due to a higher demand for black goats for use in cultural ceremonies in the regions (Mahanjana and Cronje., 2000).

In this regard, the detection of coat colour selection signatures within the South African meat-type populations offers an opportunity to unravel the genetic mechanisms responsible for indigenous goat adaptation to the adverse climatic and environmental conditions of communal production systems, which will facilitate efforts to implement improved breeding schemes in the low scaled production systems of these populations. Furthermore, coat colour has also been linked to other economical traits important for goat productivity such as reproductive characteristics (Anzures-Olvera et al., 2019; Choudhury et al., 2013; Berhanu et al., 2012), milk and meat production (Mia et al., 2020; Archana et al., 2018), and growth traits (Hossain et al., 2020; Choudhury et al., 2013; Mabrouk et al., 2010; Daramola and Adeloye et al., 2009). For instance, Lee et al. (2016) found that coat colour directly affects Holstein cattle's longevity.

However, in goats, such information is lacking. Thus, the genetic characterisation of coat colour in the indigenous goat breeds also offers the opportunity to develop and employ breeding schemes that incorporate coat colour as an essential part of the selection criteria as the trait also impacts yield and goat productivity (Abraham et al., 2018; Berhanu et al., 2012; Odubote et al., 1994). Such breeding schemes would be advantageous in the communal production systems by offering easy identification of desirable animals for genetic improvement programs (Maldonado-Jáquez et al., 2023; Nguluma et al., 2020; Torres-Hernández et al., 2020; Abraham et al., 2018), limiting the costs of measuring and recording multiple economical traits. In this regard, the objective of this chapter was to detect signatures of selection for coat colour genes under positive selection in South African commercial and indigenous meat-type goats with various coat colours and coat patterns.

5.2 Methods and materials

5.2.1 Data quality control

The genotype data was filtered using PLINK v1.9 (Purcell et al., 2007) in R (R Core Team, 2023). Quality control was carried at the sample and SNP level to remove data with high errors rates. At the SNP level, QC removed SNPs with call rates lower than 0.05 (MIND = 0.05). At the sample level, QC filtered out animals with genotyping rates less than 0.05 (GENO = 0.05), and those with minor allele frequencies lower than 0.05 (MAF < 0.05). After QC, a total of 329 goats with 51 676 autosomal SNPs remained.

5.2.2 Data analysis

5.2.2.1 Within and between population signatures of analysis

The dataset used in this analysis contained 51 676 autosomal SNPs and 329 South African meat-type goats with various coat colours and patterns including plain coats such as white, red, black, grey, and coat colour patterns such as speckled, patchy, belted, white-sided, and black-legs (Table 3.2). The commercial populations, including the Savanna (n = 38), Kalahari Red, and Boer (n = 97) breeds were of their standard coat colours of white, red, and white body red head, respectively. Ecotypes were of various coat colours and patterns such as white (n = 25), red (n = 17), black (19), grey (n = 10), speckled (n = 15), patchy (n = 15), belted (n = 16), white-sided (n = 14), and blacklegs (11) as shown in Figure 3.1. The ecotypes were grouped according to their clustering as revealed by the population structure analysis of the previous chapter (Table 4.1). Signatures of selection analysis were carried out within and between goats with different coat colour and coat colour patterns using haplotype-based within-population iHS (integrated haplotype score) and XP-EHH (cross-population extended haplotype homozygosity) comparisons, respectively in R version 4.3 (R Core Team., 2023) using the rehh package (Gautier and Vitalis, 2012).

Table 5.1: Coat colour and coat colour pattern clusters for the ecotypes

Coat Colour/Pattern	Tswana-Xhosa cluster	Zulu-Venda cluster
White	25	-
Red	9	8
Black	-	19
Grey	-	10
Belted	7	9
Patchy	15	-
Speckled	8	7
White sided	-	14
Blacklegs	-	11
Total no. per cluster	**58**	**78**

5.2.2.2 Gene annotations

A SNP was considered associated with a gene if the chromosomal position of the marker was within the 250kb chromosomal position of the gene. Candidate genes were retrieved from the Ensembl genome browser using the goat reference genome assembly (ARS1.2:

GCA_001704415.2). Identified genes were analysed to find biological pathways in which they are involved using the Kyoto Encyclopaedia of Genes and Genomes (KEGG) mapper (Kanehisa and Sato, 2020).

5.3 Results

5.3.1 Signatures of selection (iHS) within commercial breeds

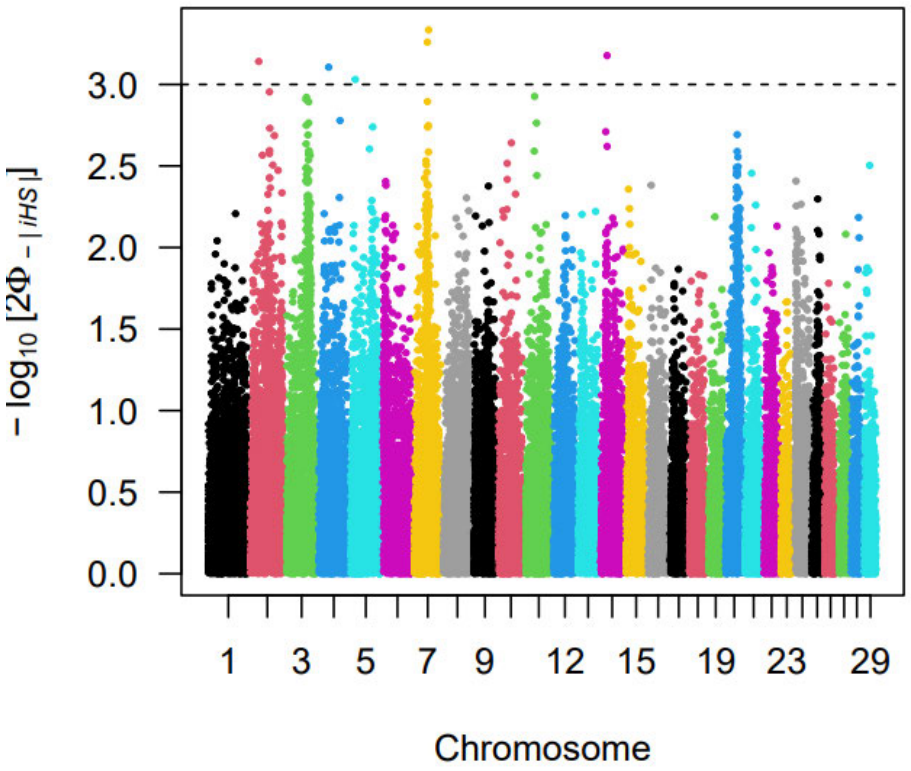

Figure 5.1: Selection signatures detected by an iHS analysis within the Savanna breed with a threshold of 3.

Markers within the Savanna breed found on chromosome 7 was linked to the *PPP2R2B* gene which was associated to several pathways included Adrenergic signalling in cardiomyocytes (chx04261), Hippo signalling pathway (chx04390), Tight junction (chx04530), T cell receptor signalling pathway (chx04660), Dopaminergic synapse (chx04728), Chagas disease (chx05142), Hepatitis C (chx05160), and Human papillomavirus infection (chx05165); *TCERG1* which is involved in Spliceosome (chx03040). *ARAP3* which is involved in pathways including Rap1 signalling pathway (chx04015), cAMP signalling pathway (chx04024), Endocytosis (chx04144); *Rell2* which is involved in the Cytokine-cytokine receptor interaction pathway (chx04060); *HDAC3* which is involved in Neutrophil extracellular trap formation (chx04613), Thyroid hormone signalling pathway (chx04919), Alcoholism (chx05034), and Viral carcinogenesis (chx05203); and *DIAPH1* which is involved in Focal adhesion (chx04510), Regulation of actin cytoskeleton (chx04810), and AGE-RAGE signalling pathway in diabetic complications (chx04933).

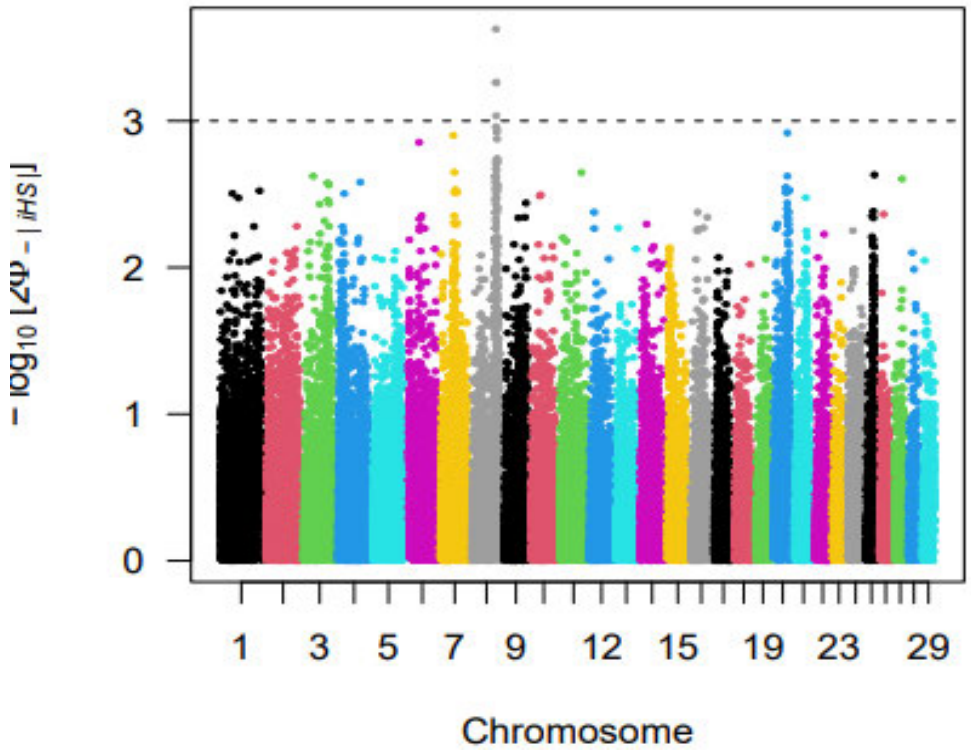

Figure 5.2: Selection signatures detected by an iHS analysis within the Kalahari Red breed with a threshold of 3.

Markers within the Kalahari Red breed found on chromosome 8 (i.e., Table 4.2) and were associated with the *BAAT* gene which is involved in several pathways including Primary bile acid biosynthesis (chx00120), Taurine and hypotaurine metabolism (chx00430), Biosynthesis of unsaturated fatty acids (chx01040), Metabolic pathways (chx01100), Peroxisome (chx04146), Bile secretion (chx04976); as well as the *GRIN3A* gene which is involved in Calcium signalling pathway (chx04020), cAMP signalling pathway (chx04024), Neuroactive ligand-receptor interaction (chx04080), Glutamatergic synapse (chx04724), Spinocerebellar ataxia (chx05017), Prion disease (chx05020), Cocaine addiction (chx05030), Amphetamine addiction (chx05031), Nicotine addiction (chx05033), and Alcoholism (chx05034). While no pathways were found for the *PLPPR1* and *RNF20* genes.

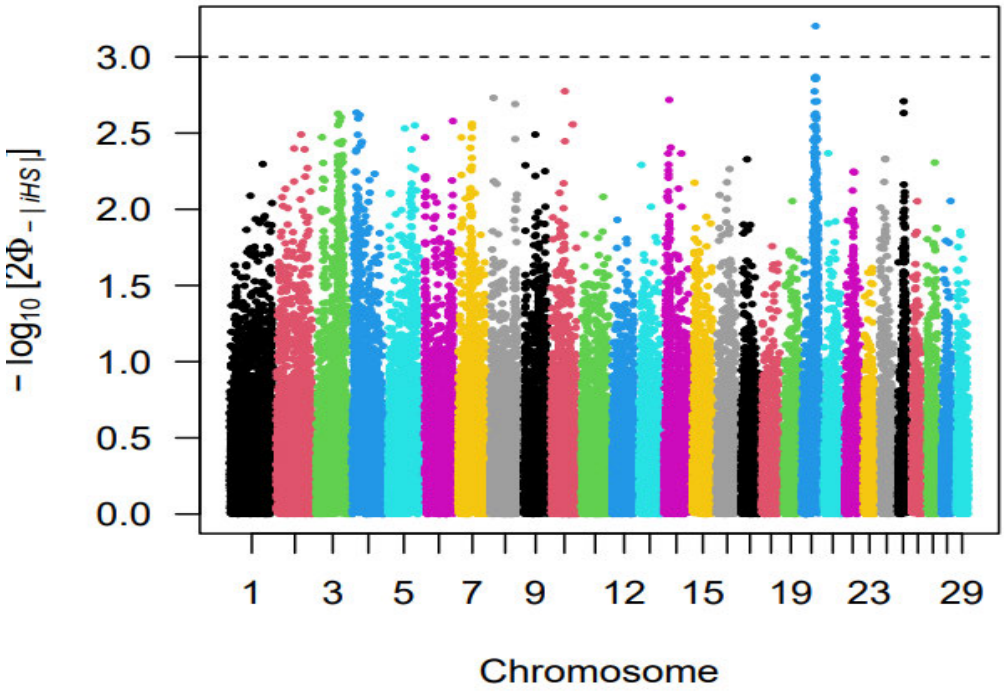

Figure 5.3: Selection signatures detected by an iHS analysis within the Boer breed with a threshold of 3.

The marker found on under selection on chromosome 20 in the Boer breed was associated with the *CDH12* gene which is involved with Cell adhesion molecules (BR:chx04515).

Table 5.2: Genes under selection within the commercial meat-type breeds with different coat colours and coat colour patterns and their associated pathways

Breed	Chr	Genes	Pathways
Boer	20	*CDH12*	
Kalahari Red	8	*PLPPR1; BAAT; RNF20; GRIN3A*	Primary bile acid biosynthesis; Taurine and hypotaurine metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Peroxisome; Bile secretion; Calcium signalling pathway; cAMP signalling pathway; Neuroactive ligand-receptor interaction; Glutamatergic synapse; Spinocerebellar ataxia; Prion-disease; Cocaine addiction; Amphetamine addiction; Nicotine addiction; Alcoholism
Savanna	2	*IKZF2*	
	7	*PPP2R2B; GPR151; TCERG1; PCDH12; DELE1; PCDH1; ARAP3; FCHSD1; RELL2; HDAC3; DIAPH1; PCDHGC4*	mRNA surveillance pathway; Sphingolipid signalling pathway; PI3K-Akt signalling pathway; AMPK signalling pathway; Adrenergic signalling in cardiomyocytes; Hippo signalling pathway; Tight junction; T cell receptor signalling pathway; Dopaminergic synapse; Chagas disease; Hepatitis C; Human papillomavirus infection; Spliceosome; Rap1 signalling pathway; cAMP signalling pathway; Endocytosis; Cytokine-cytokine receptor interaction; Neutrophil extracellular trap formation; Thyroid hormone signalling pathway; Alcoholism; Viral carcinogenesis; Focal adhesion; Regulation of actin cytoskeleton; AGE-RAGE signalling pathway in diabetic complications
	14	*NUDCD1; ENY2; PKHD1L1; SYBU*	

5.3.2 Signatures of selection (iHS) within ecotype populations with different coat colours and patterns

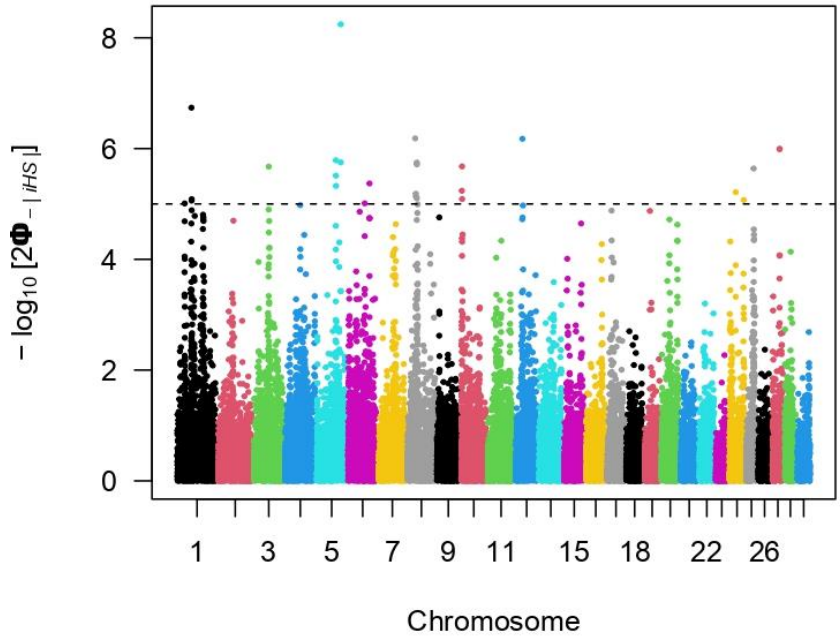

Figure 5.4: Selection signatures detected by an iHS analysis within white village goats with a threshold of 5.

Significant markers under selection detected within village goats with white coat colour were found on chromosome 1, 5, 6, 8, 10, and 24. Markers on chromosome 5 were linked to several genes including *DUSP16* which is involved in MAPK signalling pathway (chx04010); *BICD1* which is involved in the Viral life cycle - HIV-1 pathway (chx03250); *CDKN1B* which is involved in several pathways including Endocrine resistance (chx01522), ErbB signalling pathway (chx04012), HIF-1 signalling pathway (chx04066), FoxO signalling pathway (chx04068), Cell cycle (chx04110), PI3K-Akt signalling pathway (chx04151), AGE-RAGE signalling pathway in diabetic complications (chx04933), Cushing syndrome (chx04934), Measles (chx05162), Human papillomavirus infection (chx05165), Epstein-Barr virus infection (chx05169), Pathways in cancer (chx05200), Transcriptional misregulation in cancer (chx05202), Viral carcinogenesis (chx05203), MicroRNAs in cancer (chx05206), Prostate cancer (chx05215), Chronic myeloid leukemia (chx05220), Small cell lung cancer (chx05222), and Gastric cancer (chx05226).

Signals on chromosome 6 were linked to genes including *CNGA1* which is involved in pathways such as cGMP-PKG signalling pathway (chx04022), cAMP signalling pathway (chx04024), and Phototransduction (chx04744). Signals on chromosome 8 were linked to several genes including *RANBP6*; *MLANA*; *UHRF2*; *LURAP1L*; *IL33* which is involved in Cytokine-cytokine receptor interaction (chx04060), Necroptosis (chx04217), Cytosolic DNA-

sensing pathway (chx04623), and Influenza A (chx05164); *GLDC* which is involved in Glycine, serine and threonine metabolism (chx00260), Glyoxylate and dicarboxylate metabolism (chx00630), Lipoic acid metabolism (chx00785), Metabolic pathways (chx01100), Carbon metabolism (chx01200), *TYRP1* which is involved in Tyrosine metabolism (chx00350), Metabolic pathways (chx01100), Melanogenesis (chx04916). Signals on chromosome 10 were linked to genes including *KCNK10* which is involved in Gastric acid secretion (chx04971).

Signals on chromosome 24 were linked to genes including *B4GALT6* which is involved in Sphingolipid metabolism (chx00600), and Metabolic pathways (chx01100); *DSC1*; *ZNF532*; *TTR* which is involved in Thyroid hormone synthesis (chx04918); GRP which is involved in Neuroactive ligand-receptor interaction (chx04080); *DSG2* which is involved in Arrhythmogenic right ventricular cardiomyopathy (chx05412); *MALT1* NF-kappa B signalling pathway (chx04064), C-type lectin receptor signalling pathway (chx04625), T cell receptor signalling pathway (chx04660), B cell receptor signalling pathway (chx04662), and Tuberculosis (chx05152); and *DSG1* which is involved in Staphylococcus aureus infection (chx05150).

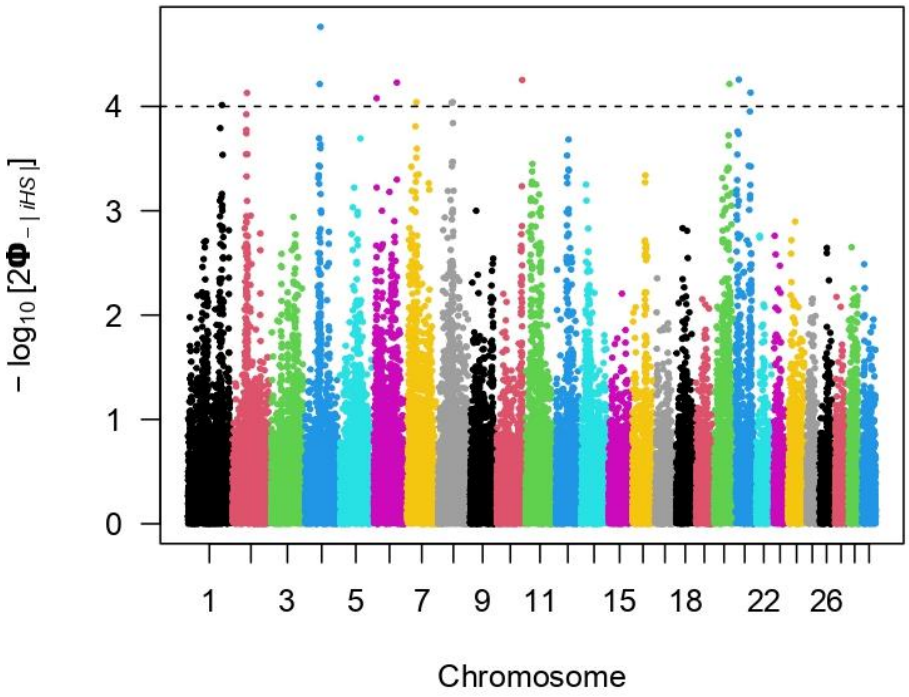

Figure 5.5: Selection signatures detected by an iHS analysis within red village goats in the TX cluster. The threshold for the detection of signatures was set to 4.

Within village goats with red coat colour in the TX cluster, significant signatures were found on chromosome 1, 4, 6, 8, 10, 20, and 21. Markers on chromosome 4 were linked to genes including *NT5C3A* which is involved in Transcriptional misregulation in cancer (chx05202). Markers on chromosome 6 were linked to genes including *UGT8* which is involved in Ether lipid metabolism (chx00565), Sphingolipid metabolism (chx00600), and Metabolic pathways (chx01100). Markers on chromosome 8 were linked to genes including *RORB* which is involved in Circadian rhythm (chx04710); *TRPM6* which is involved in Mineral absorption (chx04978); *PSAT1* which is involved in Glycine, serine and threonine metabolism (chx00260), Cysteine and methionine metabolism (chx00270), Vitamin B6 metabolism (chx00750), Metabolic pathways (chx01100), Carbon metabolism (chx01200), and Biosynthesis of amino acids (chx01230). Signals on chromosome 10 were linked to genes including *IQGAP2* which is involved in Regulation of actin cytoskeleton (chx04810); and *SV2C* which is involved in ECM-receptor interaction (chx04512). Signals on chromosome 20 were linked to genes including *MYO10* which is involved in Fc gamma R-mediated phagocytosis (chx04666), and Motor proteins (chx04814). Signals on chromosome 21 were linked to genes including *RGMA* which is involved in TGF-beta signalling pathway (chx04350), and Axon guidance (chx04360); and *CPSF2* which is involved in mRNA surveillance pathway (chx03015).

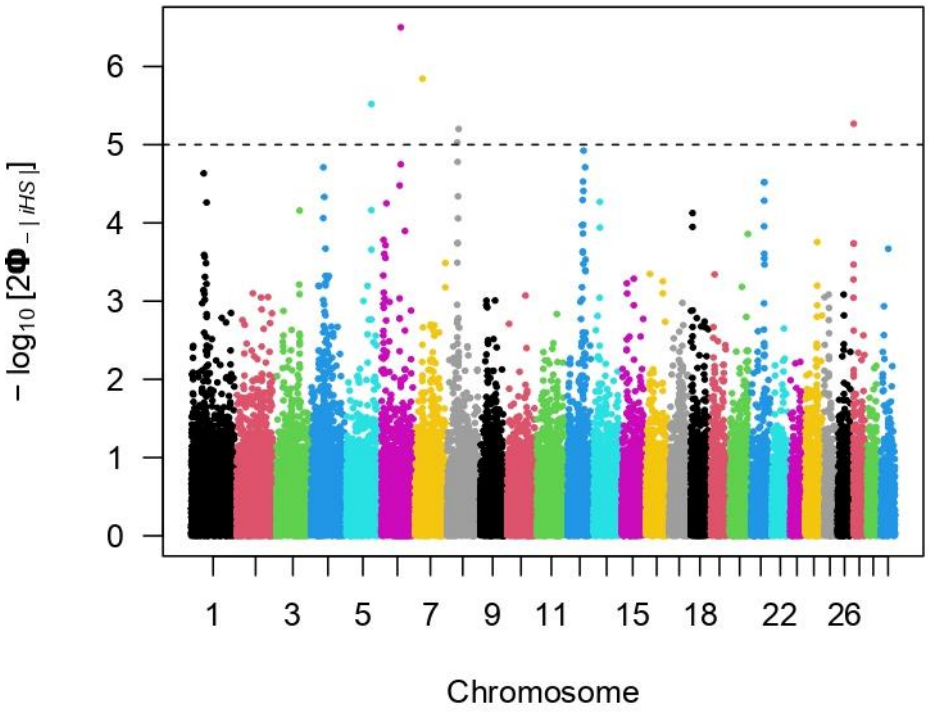

Figure 5.6: Selection signatures detected by an iHS analysis within red village goats in the ZV cluster. The threshold for the detection of signatures was set to 5.

With red village goats in the ZV cluster, significant signals were found on chromosome 5, 6, and 8. Signals on chromosome 5 were linked to the genes *CAPZA3* which is involved in Endocytosis (chx04144), and Motor proteins (chx04814); *PLCZ1* which is involved in Inositol phosphate metabolism (chx00562), Metabolic pathways (chx01100), Calcium signalling pathway (chx04020), Phosphatidylinositol signalling system (chx04070), Oocyte meiosis (chx04114), and Thyroid hormone signalling pathway (chx04919); and *PIK3C2G* which is involved in several pathways including Inositol phosphate metabolism (chx00562), Metabolic pathways (chx01100), Phosphatidylinositol signalling system (chx04070), and Salmonella infection (chx05132). Signals on chromosome 6 were linked to the genes *KDR* which is involved in several pathways including EGFR tyrosine kinase inhibitor resistance (chx01521), MAPK signalling pathway (chx04010), Ras signalling pathway (chx04014), Rap1 signalling pathway (chx04015), Calcium signalling pathway (chx04020), and PI3K-Akt signalling pathway (chx04151); *SRD5A3* which is involved in pathways such as Steroid hormone biosynthesis (chx00140), N-Glycan biosynthesis (chx00510), and Metabolic pathways (chx01100). Signatures found on chromosome 8 were linked to the genes including *VLDLR* which is involved in Spinocerebellar ataxia (chx05017), and Lipid and atherosclerosis (chx05417).

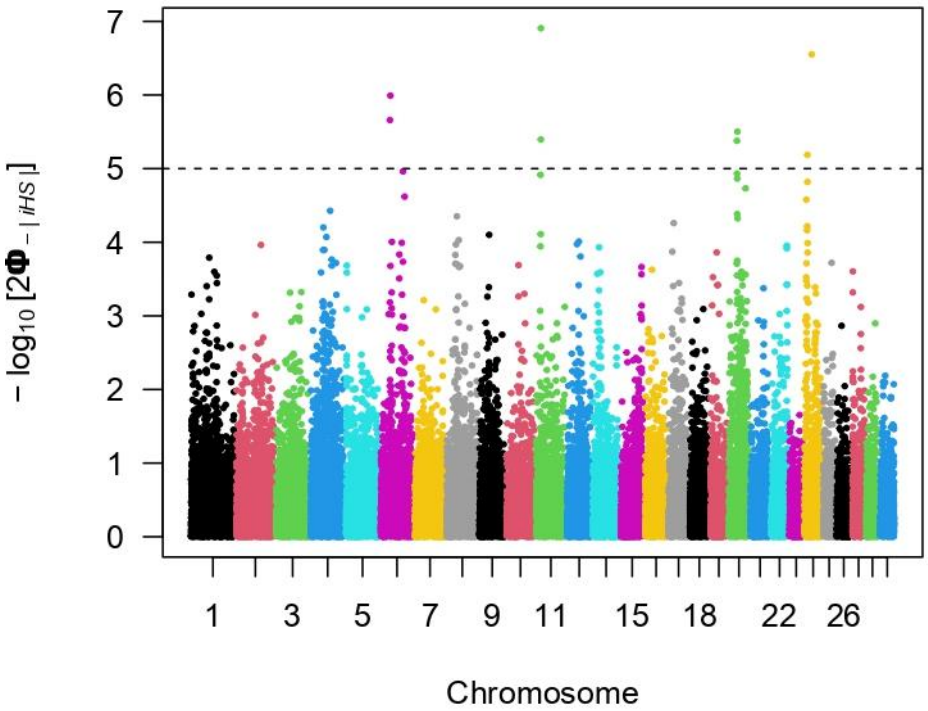

Figure 5.7: Selection signatures detected by an iHS analysis within black village goats with a threshold of 5

Within black village goats, markers under selection were found on chromosome 6, 20, and 24. Markers on chromosome 6 were linked to genes including *MMRN1*; *SNCA* which is involved in Alzheimer disease (chx05010), Parkinson disease (chx05012), and Pathways of neurodegeneration - multiple diseases (chx05022). Signals on chromosome 20 were linked to genes including *EMB*; and *FGF10* which is involved in several pathways including MAPK signalling pathway (chx04010), Ras signalling pathway (chx04014), Rap1 signalling pathway (chx04015), Calcium signalling pathway (chx04020), PI3K-Akt signalling pathway (chx04151), Regulation of actin cytoskeleton (chx04810), Pathways in cancer (chx05200), Chemical carcinogenesis - receptor activation (chx05207), Melanoma (chx05218), Breast cancer (chx05224), and Gastric cancer (chx05226). Signatures of selection on chromosome 24 were linked to several genes including *CDH7*; *B4GALT6* which is involved in Sphingolipid metabolism (chx00600), and Metabolic pathways (chx01100); *TTR* which is involved in Thyroid hormone synthesis (chx04918); *DSG2* which is involved in Arrhythmogenic right ventricular cardiomyopathy (chx05412); and *DSG1* which is involved in Staphylococcus aureus infection (chx05150).

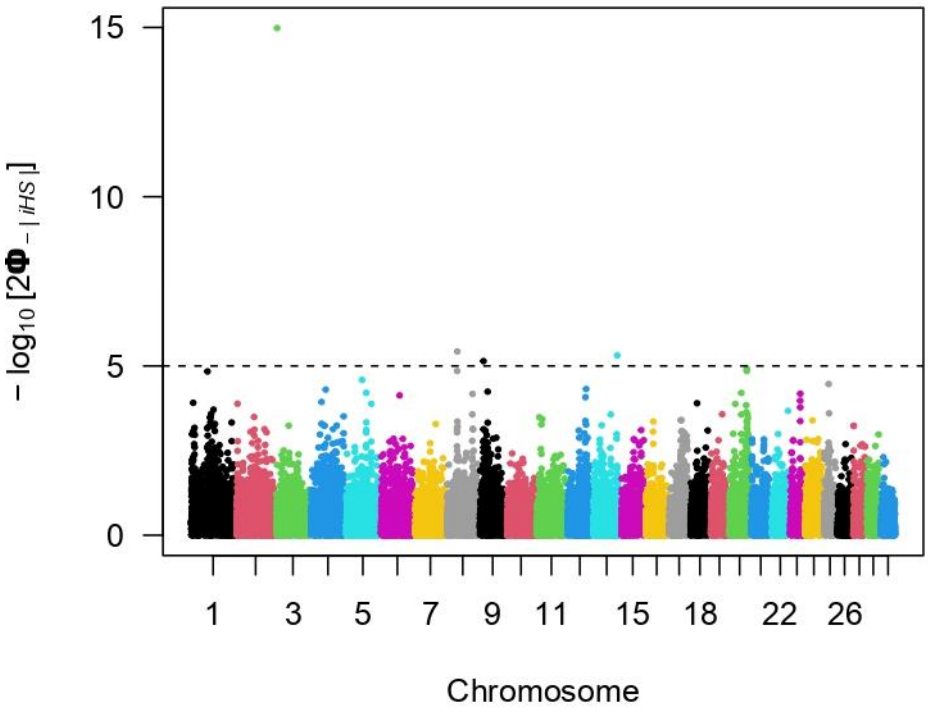

Figure 5.8: Selection signatures detected by an iHS analysis within grey village goats with a threshold of 5

Within grey village goats, markers under selection were found on chromosome 3, 8, and 9. Genes linked to signatures on chromosomes 3 included *LRRFIP1*; *RAB17*; *PRLH* which is involved in Butanoate metabolism (chx00650), and Metabolic pathways (chx01100); *MLPH*;

and *COL6A3* which is involved in PI3K-Akt signalling pathway (chx04151), Focal adhesion (chx04510), ECM-receptor interaction (chx04512), Protein digestion and absorption (chx04974), and Human papillomavirus infection (chx05165). Signals on chromosome 8 were linked to the gene *DMAC1*, while genes found on chromosome 9 were *NCOA7* which is involved in Lysosome (chx04142), and *HEY2* which is involved in several pathways including the Notch signalling pathway (chx04330), Human papillomavirus infection (chx05165), Pathways in cancer (chx05200), and Breast cancer (chx05224).

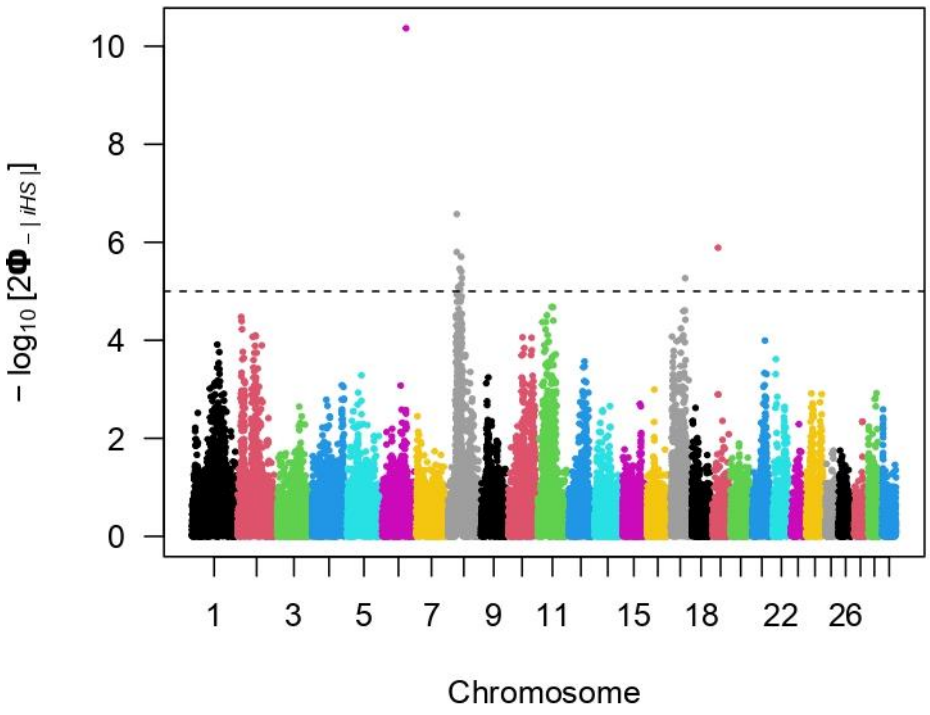

Figure 5.9: Selection signatures detected by an iHS analysis within speckled village goats in the TX cluster with a threshold of 5

Within speckled village goats in the TZ cluster, signatures of selection were found on chromosome 6, 8, 17, and 19. The signals found on chromosome 6 were linked to *ODAM*; *CABS1*; and *AMTN*. Genes detected to be under selection on chromosome 8 included *LURAP1L*; *TYRP1* which is involved in Tyrosine metabolism (chx00350), Metabolic pathways (chx01100), and Melanogenesis (chx04916); *PTPRD* which is involved with Cell adhesion molecules (chx04514); *KCNV2*; *VLDLR* which is involved in Spinocerebellar ataxia (chx05017), and Lipid and atherosclerosis (chx05417); *KANK1*; *DOCK8*; and *CEMIP2*. Signals on chromosome 17 were linked to the genes *SETD7* which is involved in Lysine degradation (chx00310), Metabolic pathways (chx01100), FoxO signalling pathway (chx04068); and *MGST2* which is involved in Glutathione metabolism (chx00480), Metabolism of xenobiotics by cytochrome P450 (chx00980), Drug metabolism - cytochrome

P450 (chx00982), Drug metabolism - other enzymes (chx00983), Metabolic pathways (chx01100), Platinum drug resistance (chx01524), Pathways in cancer (chx05200), Chemical carcinogenesis - DNA adducts (chx05204), Chemical carcinogenesis - receptor activation (chx05207), Chemical carcinogenesis - reactive oxygen species (chx05208), Hepatocellular carcinoma (chx05225), and Fluid shear stress and atherosclerosis (chx05418).

Markers on chromosome 19 were linked to several genes including *CRK* which is involved in MAPK signalling pathway (chx04010), ErbB signalling pathway (chx04012), Rap1 signalling pathway (chx04015), Chemokine signalling pathway (chx04062), Focal adhesion (chx04510), Fc gamma R-mediated phagocytosis (chx04666), Neurotrophin signalling pathway (chx04722), Regulation of actin cytoskeleton (chx04810), Insulin signalling pathway (chx04910), Growth hormone synthesis, secretion and action (chx04935), Bacterial invasion of epithelial cells (chx05100), Yersinia infection (chx05135), Human cytomegalovirus infection (chx05163), Human immunodeficiency virus 1 infection (chx05170), Pathways in cancer (chx05200), MicroRNAs in cancer (chx05206), Renal cell carcinoma (chx05211), and Chronic myeloid leukemia (chx05220); *MYO1C* which is involved with Motor proteins (chx04814); *INPP5K* which is involved in Inositol phosphate metabolism (chx00562), Metabolic pathways (chx01100); *PITPNA*; *SLC43A2*; *SCARF1*; *RILP* which is involved in Phagosome (chx04145), and Salmonella infection (chx05132); *PRPF8* which is involved in Spliceosome (chx03040); *TLCD2*; *WDR81*; *SERPINF2* which is involved in Complement and coagulation cascades (chx04610); *SERPINF1* which is involved in Wnt signalling pathway (chx04310); *SMYD4*; *RPA1* which is involved in DNA replication (chx03030), Nucleotide excision repair (chx03420), Mismatch repair (chx03430), Homologous recombination (chx03440), Fanconi anemia pathway (chx03460); *RTN4RL1*; and *DPH1*.

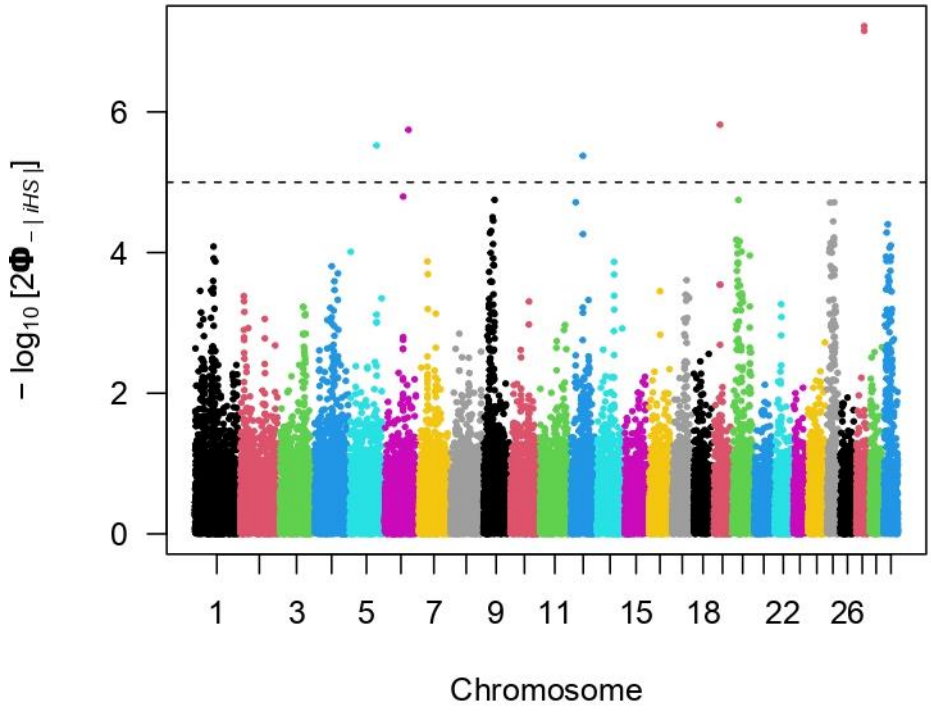

Figure 5.10: Selection signatures detected by an iHS analysis within speckled village goats in the ZV cluster with a threshold of 5

Within speckled goats in the ZV cluster, markers under selection were found on chromosome 5, 6, 19, and 27. Genes found to be under selection on chromosome 5 included *HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16* which is involved in MAPK signalling pathway (chx04010); and *BORCS5*. Signals on chromosome 6 were linked to the three genes *SULT1B1; SULT1E1*; and *ODAM*. Markers on chromosome 19 were linked to several genes including *CRK* which is involved in MAPK signalling pathway (chx04010), ErbB signalling pathway (chx04012), Rap1 signalling pathway (chx04015), Chemokine signalling pathway (chx04062), Focal adhesion (chx04510), Fc gamma R-mediated phagocytosis (chx04666), Neurotrophin signalling pathway (chx04722), Regulation of actin cytoskeleton (chx04810), Insulin signalling pathway (chx04910), Growth hormone synthesis, secretion and action (chx04935), Bacterial invasion of epithelial cells (chx05100), Yersinia infection (chx05135), Human cytomegalovirus infection (chx05163), Human immunodeficiency virus 1 infection (chx05170), Pathways in cancer (chx05200), MicroRNAs in cancer (chx05206), Renal cell carcinoma (chx05211), Chronic myeloid leukemia (chx05220); *MYO1C* which is involved in Motor proteins (chx04814); *INPP5K* which is involved in Inositol phosphate metabolism (chx00562) and Metabolic pathways (chx01100); *PITPNA; SLC43A2; SCARF1; RILP* which is involved in Phagosome (chx04145) and

Salmonella infection (chx05132); *PRPF8* which is involved in Spliceosome (chx03040); *TLCD2; WDR81; SERPINF2* which is involved in Complement and coagulation cascades (chx04610); SERPINF1 which is involved in Wnt signalling pathway (chx04310); *SMYD4; RPA1* which is involved in DNA replication (chx03030), Nucleotide excision repair (chx03420), Mismatch repair (chx03430), Homologous recombination (chx03440), Fanconi anemia pathway (chx03460); *RTN4RL1*; and *DPH1*.

Signatures on chromosome 27 were linked to the genes *FAT1*; *MTNR1A* which is involved Neuroactive ligand-receptor interaction (chx04080) and Circadian entrainment (chx04713); *F11; KLKB1* which is involved in Complement and coagulation cascades (chx04610); *CYP4V2;TLR3* which is involved in several pathways including Necroptosis (chx04217), Toll-like receptor signalling pathway (chx04620), Hepatitis C (chx05160), Hepatitis B (chx05161), Influenza A (chx05164), Human papillomavirus infection (chx05165), Kaposi sarcoma-associated herpesvirus infection (chx05167), Herpes simplex virus 1 infection (chx05168), and Coronavirus disease - COVID-19 (chx05171).

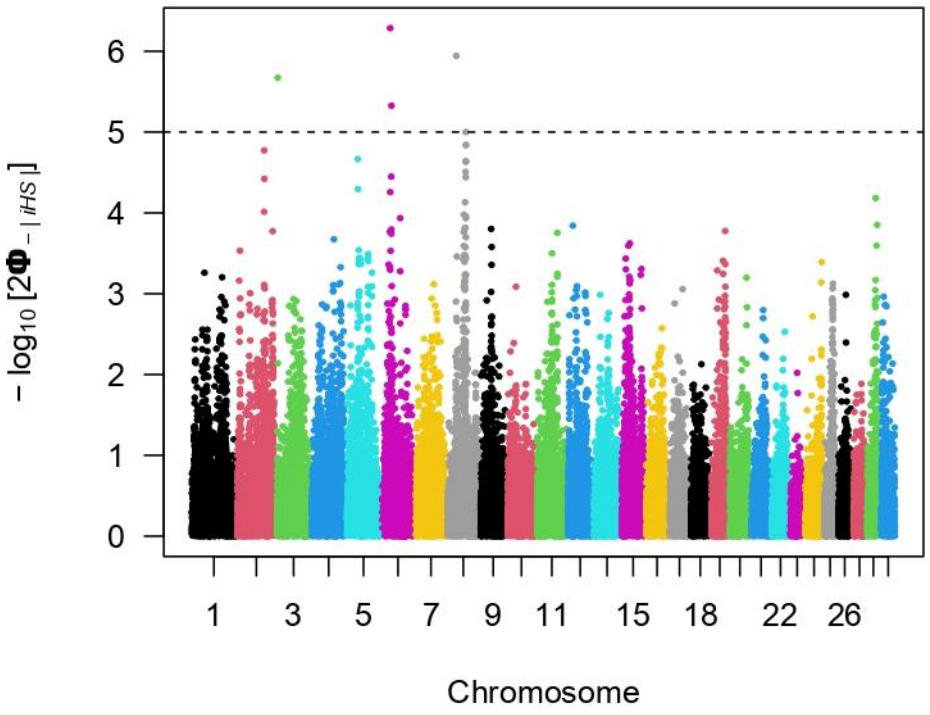

Figure 5.11: Selection signatures detected by an iHS analysis within patchy village goats with a threshold of 5.

Within patchy village goats, markers under selection were detected on chromosome 3, 6, and 8. Signals on chromosome 3 were linked to the genes *RBM44; LRRFIP1; RAB17; PRLH* which

is involved in Neuroactive ligand-receptor interaction (chx04080); *MLPH; COL6A3* PI3K-Akt signalling pathway (chx04151), Focal adhesion (chx04510), ECM-receptor interaction (chx04512), Protein digestion and absorption (chx04974), Human papillomavirus infection (chx05165). Signals on chromosome 6 were linked to the genes *SMARCAD1* which is involved in Signalling pathways regulating pluripotency of stem cells (chx04550); *ATOH1; GRID2* which is involved in Neuroactive ligand-receptor interaction (chx04080), and Long-term depression (chx04730); *MMRN1*; and *SNCA* which is involved in Alzheimer disease (chx05010), Parkinson disease (chx05012), and Pathways of neurodegeneration - multiple diseases (chx05022). Markers on chromosome 8 were linked to the two genes *LURAP1L*; and *TYRP1* which is involved in Tyrosine metabolism (chx00350), Metabolic pathways (chx01100), and Melanogenesis (chx04916).

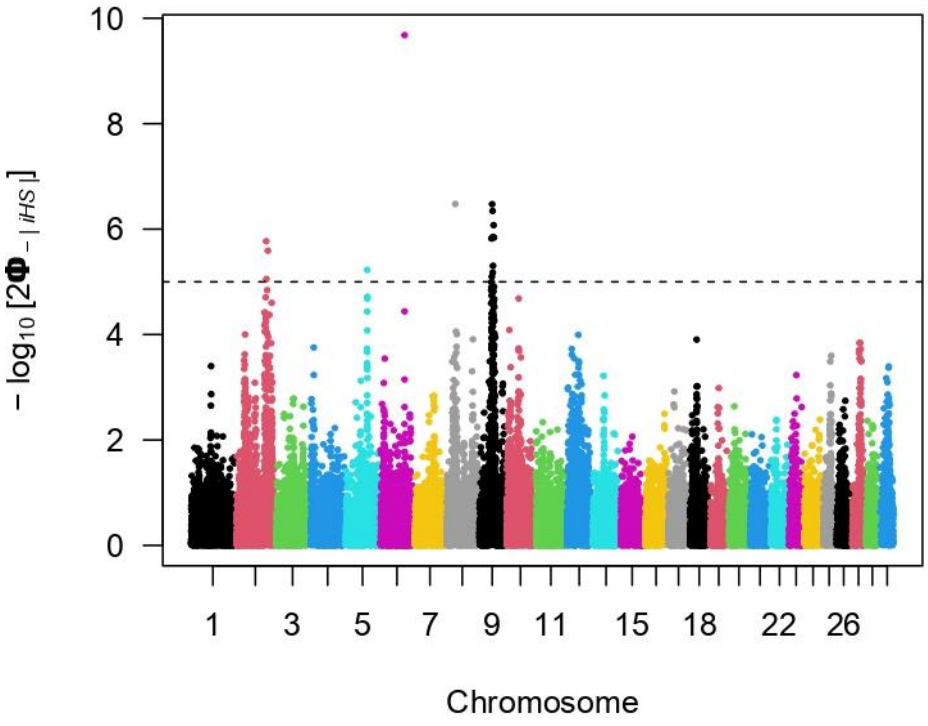

Figure 5.12: Selection signatures detected by an iHS analysis within belted village goats in the TX cluster with a threshold of 5.

Within belted village goats in the TX cluster, markers under selection were found on chromosome 2, 5, 6, 8, and 9. Signals on chromosome 2 were linked to the following genes *XIRP2; B3GALT1* which is involved in Glycosphingolipid biosynthesis - lacto and neolacto series (chx00601), and Metabolic pathways (chx01100); *STK39;* and *SP3*. Selection signatures on chromosome 5 were linked to the two genes *BICD1* which is involved in Viral life cycle - HIV-1 (chx03250); and *RESF1*. Markers on chromosome 6 were linked to the following genes *ODAM; CABS1; AMTN*. Signals on chromosome 8 were linked to the two genes *LURAP1L;*

and *TYRP1* which is involved in Tyrosine metabolism (chx00350), Metabolic pathways (chx01100), and Melanogenesis (chx04916).

Signatures on chromosome 9 were linked to several genes including *MAP3K7* which is involved in MAPK signalling pathway (chx04010), NF-kappa B signalling pathway (chx04064), Autophagy – animal (chx04140), AMPK signalling pathway (chx04152), Wnt signalling pathway (chx04310), Osteoclast differentiation (chx04380), Adherens junction (chx04520), Neutrophil extracellular trap formation (chx04613), Toll-like receptor signalling pathway (chx04620), NOD-like receptor signalling pathway (chx04621), RIG-I-like receptor signalling pathway (chx04622), IL-17 signalling pathway (chx04657), T cell receptor signalling pathway (chx04660), TNF signalling pathway (chx04668), Alcoholic liver disease (chx04936), Salmonella infection (chx05132), Yersinia infection (chx05135), Leishmaniasis (chx05140), Toxoplasmosis (chx05145), Hepatitis B (chx05161), Measles (chx05162), Herpes simplex virus 1 infection (chx05168), Epstein-Barr virus infection (chx05169), Human immunodeficiency virus 1 infection (chx05170), Coronavirus disease - COVID-19 (chx05171), Lipid and atherosclerosis (chx05417), and Fluid shear stress and atherosclerosis (chx05418); *GJA10; CASP8AP2; MDN1* which is involved in Ribosome biogenesis in eukaryotes (chx03008); *LYRM2; ANKRD6; RRAGD* which is involved in Autophagy – animal (chx04140), and mTOR signalling pathway (chx04150); *UBE2J1* which is involved in Ubiquitin mediated proteolysis (chx04120), Protein processing in endoplasmic reticulum (chx04141), Parkinson disease (chx05012), Pathways of neurodegeneration - multiple diseases (chx05022); *RNGTT* which is involved in mRNA surveillance pathway (chx03015); *CNR1* which is involved in Rap1 signalling pathway (chx04015), Neuroactive ligand-receptor interaction (chx04080), Thermogenesis (chx04714), and Retrograde endocannabinoid signalling (chx04723); *SYNCRIP; SNX14; NT5E* which is involved in Purine metabolism (chx00230), Pyrimidine metabolism (chx00240), Nicotinate and nicotinamide metabolism (chx00760), Metabolic pathways (chx01100), Nucleotide metabolism (chx01232); and *TBX18*.

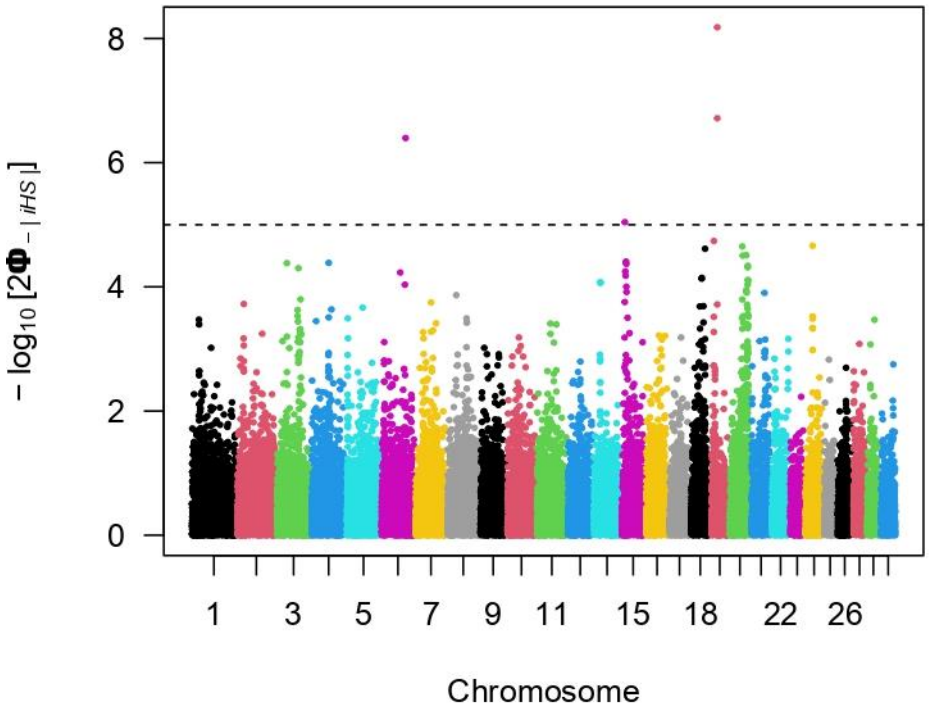

Figure 5.13: Selection signatures detected by an iHS analysis within belted village goats in the ZV cluster with a threshold of 5.

Within belted village goats in the ZV cluster, markers under selection were found on chromosome 6, and 19. The markers on chromosome 6 were linked to the genes *ODAM*; *CABS1*; and *AMTN*. Markers on chromosome 19 were linked to several genes including *CRK* which is involved in MAPK signalling pathway (chx04010), ErbB signalling pathway (chx04012), Rap1 signalling pathway (chx04015), Chemokine signalling pathway (chx04062), Focal adhesion (chx04510), Fc gamma R-mediated phagocytosis (chx04666), Neurotrophin signalling pathway (chx04722), Regulation of actin cytoskeleton (chx04810), Insulin signalling pathway (chx04910), Growth hormone synthesis, secretion and action (chx04935), Bacterial invasion of epithelial cells (chx05100), Yersinia infection (chx05135), Human cytomegalovirus infection (chx05163), Human immunodeficiency virus 1 infection (chx05170), Pathways in cancer (chx05200), MicroRNAs in cancer (chx05206), Renal cell carcinoma (chx05211), Chronic myeloid leukemia (chx05220); *MYO1C* which is involved in Motor proteins (chx04814); *INPP5K* which is involved in Inositol phosphate metabolism (chx00562) and Metabolic pathways (chx01100); *PITPNA; SLC43A2; SCARF1; RILP* which is involved in Phagosome (chx04145) and Salmonella infection (chx05132); *PRPF8* which is involved in Spliceosome (chx03040); *TLCD2; WDR81; SERPINF2* which is involved in Complement and coagulation cascades (chx04610); SERPINF1 which is involved in Wnt signalling pathway (chx04310); *SMYD4; RPA1* which is involved in DNA replication (chx03030), Nucleotide excision repair (chx03420), Mismatch repair (chx03430),

Homologous recombination (chx03440), Fanconi anemia pathway (chx03460); *RTN4RL1*; and *DPH1*.

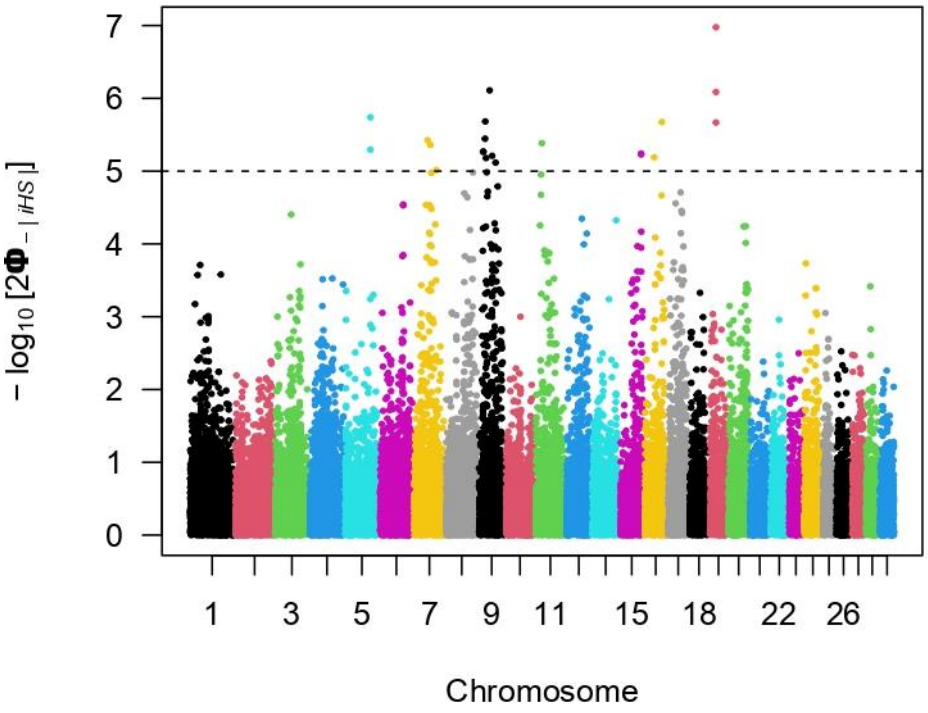

Figure 5.14: Selection signatures detected by an iHS analysis within white-sided village goats with a threshold of 5.

Within white sided village goats, markers under selection were found on chromosome 2, 4, 19, and 20. Signals on chromosome 2 were linked to the two genes *FTCDNL1*; and *SATB2*. Markers on chromosome 4 were linked to *CREB5* which is involved in several pathways including cGMP-PKG signalling pathway (chx04022), cAMP signalling pathway (chx04024), PI3K-Akt signalling pathway (chx04151), AMPK signalling pathway (chx04152), Longevity regulating pathway (chx04211), Adrenergic signalling in cardiomyocytes (chx04261), TNF signalling pathway (chx04668), Thermogenesis (chx04714), Cholinergic synapse (chx04725), Dopaminergic synapse (chx04728), Insulin secretion (chx04911), Estrogen signalling pathway (chx04915), Thyroid hormone synthesis (chx04918), Glucagon signalling pathway (chx04922), Aldosterone synthesis and secretion (chx04925), Relaxin signalling pathway (chx04926), Cortisol synthesis and secretion (chx04927), Parathyroid hormone synthesis, secretion and action (chx04928), Insulin resistance (chx04931), Cushing syndrome (chx04934), Growth hormone synthesis, secretion and action (chx04935), Vasopressin-regulated water reabsorption (chx04962), Huntington disease (chx05016), Prion disease (chx05020), Cocaine addiction (chx05030), Amphetamine addiction (chx05031), Alcoholism

(chx05034), Hepatitis B (chx05161), Human cytomegalovirus infection (chx05163), Human papillomavirus infection (chx05165), Human T-cell leukemia virus 1 infection (chx05166), Viral carcinogenesis (chx05203), Chemical carcinogenesis - receptor activation (chx05207), and Prostate cancer (chx05215); *TRIL; CPVL; CHN2; PDE1C; LSM5; AVL9; KBTBD2; FKBP9; NT5C3A* which is involved in Pyrimidine metabolism (chx00240), Metabolic pathways (chx01100), Nucleotide metabolism (chx01232); *NPSR1* which is involved in Neuroactive ligand-receptor interaction (chx04080); *DPY19L1; DPY19L2; TBX20; HERPUD2; SEPTIN7; EEPD1; AOAH; ELMO1* which is involved in Chemokine signalling pathway (chx04062), Bacterial invasion of epithelial cells (chx05100), Salmonella infection (chx05132), Yersinia infection (chx05135); *WNT2* which is involved in several pathways including the mTOR signalling pathway (chx04150), Wnt signalling pathway (chx04310), Hippo signalling pathway (chx04390), Signalling pathways regulating pluripotency of stem cells (chx04550), Melanogenesis (chx04916), Cushing syndrome (chx04934), Alzheimer disease (chx05010), Pathways of neurodegeneration - multiple diseases (chx05022), Human papillomavirus infection (chx05165), Pathways in cancer (chx05200), Proteoglycans in cancer (chx05205), Basal cell carcinoma (chx05217), Breast cancer (chx05224), Hepatocellular carcinoma (chx05225), and Gastric cancer (chx05226); *ASZ1; CFTR* which is involved in several pathways including ABC transporters (chx02010), cAMP signalling pathway (chx04024), AMPK signalling pathway (chx04152), Tight junction (chx04530), Gastric acid secretion (chx04971), Pancreatic secretion (chx04972), and Bile secretion (chx04976); *CTTNBP2; SRPK2; KMT2E* which is involved in Lysine degradation (chx00310), and Metabolic pathways (chx01100); *ORC5* which is involved in Cell cycle (chx04110); *RELN* which is involved in PI3K-Akt signalling pathway (chx04151), Focal adhesion (chx04510), ECM-receptor interaction (chx04512), Spinocerebellar ataxia (chx05017), and Human papillomavirus infection (chx05165); *SEMA3E; SEMA3A; SEMA3D* which are involved in Axon guidance (chx04360); and *AGMO*.

Markers on chromosome 19 were linked to several genes including *CRK* which is involved in MAPK signalling pathway (chx04010), ErbB signalling pathway (chx04012), Rap1 signalling pathway (chx04015), Chemokine signalling pathway (chx04062), Focal adhesion (chx04510), Fc gamma R-mediated phagocytosis (chx04666), Neurotrophin signalling pathway (chx04722), Regulation of actin cytoskeleton (chx04810), Insulin signalling pathway (chx04910), Growth hormone synthesis, secretion and action (chx04935), Bacterial invasion of epithelial cells (chx05100), Yersinia infection (chx05135), Human cytomegalovirus

infection (chx05163), Human immunodeficiency virus 1 infection (chx05170), Pathways in cancer (chx05200), MicroRNAs in cancer (chx05206), Renal cell carcinoma (chx05211), Chronic myeloid leukemia (chx05220); *MYO1C* which is involved in Motor proteins (chx04814); *INPP5K* which is involved in Inositol phosphate metabolism (chx00562) and Metabolic pathways (chx01100); *PITPNA; SLC43A2; SCARF1; RILP* which is involved in Phagosome (chx04145) and Salmonella infection (chx05132); *PRPF8* which is involved in Spliceosome (chx03040); *TLCD2; WDR81; SERPINF2* which is involved in Complement and coagulation cascades (chx04610); *SERPINF1* which is involved in Wnt signalling pathway (chx04310); *SMYD4; RPA1* which is involved in DNA replication (chx03030), Nucleotide excision repair (chx03420), Mismatch repair (chx03430), Homologous recombination (chx03440), Fanconi anemia pathway (chx03460); *RTN4RL1*; and *DPH1*. Markers on chromosome 20 is linked to the *CDH9* gene.

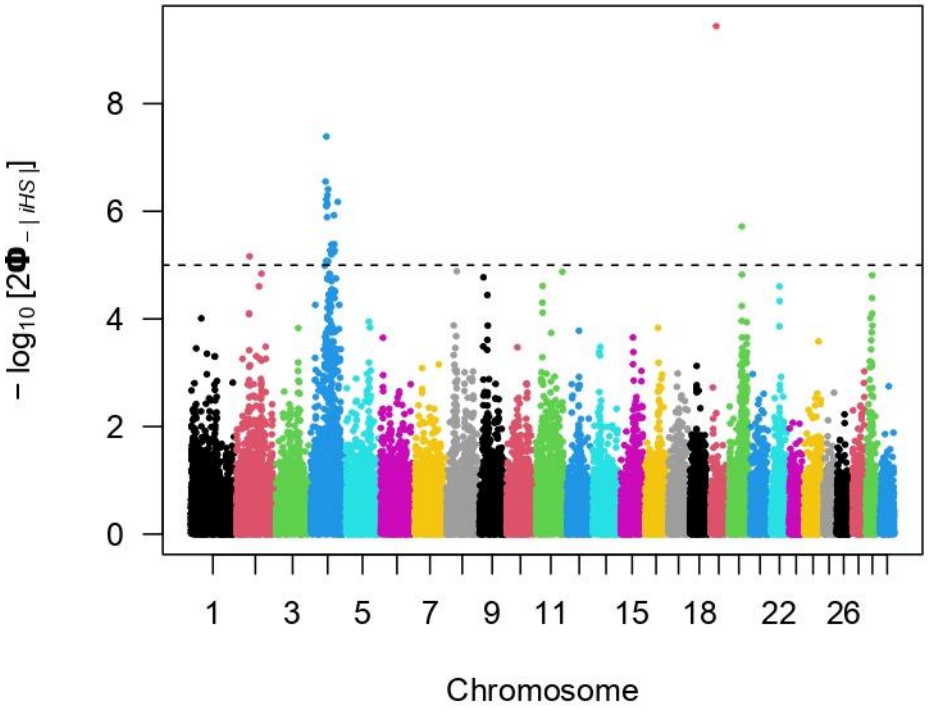

Figure 5.15: Selection signatures detected by an iHS analysis within village goats with black legs with a threshold of 5.

Within blacklegged village goats, markers under selection were detected on chromosome 5, 7, 9, 15, 16, and 19. Markers on chromosome 5 were linked to the genes *CAPZA3* which is involved in Endocytosis (chx04144), and Motor proteins (chx04814); *PLCZ1* which is involved in Inositol phosphate metabolism (chx00562), Metabolic pathways (chx01100),

Calcium signalling pathway (chx04020), Phosphatidylinositol signalling system (chx04070), Oocyte meiosis (chx04114), Thyroid hormone signalling pathway (chx04919); *PIK3C2G* which is involved in Inositol phosphate metabolism (chx00562), Metabolic pathways (chx01100), Phosphatidylinositol signalling system (chx04070), and Salmonella infection (chx05132). Signals on chromosome 7 were linked to the genes *SLC36A2* which is involved in Protein digestion and absorption (chx04974); *SLC36A3; GM2A* which is involved in Lysosome (chx04142); *CCDC69; TNIP1; GPX3* which is involved in Glutathione metabolism (chx00480), Metabolic pathways (chx01100), Thyroid hormone synthesis (chx04918), Amyotrophic lateral sclerosis (chx05014), Huntington disease (chx05016), and Pathways of neurodegeneration - multiple diseases (chx05022); *ZNF300* which is involved in Herpes simplex virus 1 infection (chx05168); *SMIM3; PCDH12; DELE1; PCDH1; ARAP3* which is involved in Rap1 signalling pathway (chx04015), cAMP signalling pathway (chx04024), and Endocytosis (chx04144); *FCHSD1; RELL2* which is involved in Cytokine-cytokine receptor interaction (chx04060); *HDAC3* which is involved in Neutrophil extracellular trap formation (chx04613), Thyroid hormone signalling pathway (chx04919), Alcoholism (chx05034), and Viral carcinogenesis (chx05203); *DIAPH1* which is involved in Focal adhesion (chx04510), Regulation of actin cytoskeleton (chx04810), AGE-RAGE signalling pathway in diabetic complications (chx04933); *PCDHGC4*; *SNCAIP* which is involved in Parkinson disease (chx05012), and Pathways of neurodegeneration - multiple diseases (chx05022); and *SNX2* which is involved in Endocytosis (chx04144).

Markers on chromosome 9 were linked to several genes including *CLVS2*; *FAM229B*; *TUBE1* which is involved in Motor proteins (chx04814); *CCN6*; *FYN* which is involved in Sphingolipid signalling pathway (chx04071), Phospholipase D signalling pathway (chx04072), Axon guidance (chx04360), Osteoclast differentiation (chx04380), Focal adhesion (chx04510), Adherens junction (chx04520), Platelet activation (chx04611), Natural killer cell mediated cytotoxicity (chx04650), T cell receptor signalling pathway (chx04660), Fc epsilon RI signalling pathway (chx04664), Cholinergic synapse (chx04725), Prion disease (chx05020), Viral myocarditis (chx05416); *TRAF3IP2* which is involved in Cellular senescence (chx04218), and IL-17 signalling pathway (chx04657); *BACH2; GJA10; CASP8AP2; MDN1* which is involved in Ribosome biogenesis in eukaryotes (chx03008); *LYRM2; ANKRD6; ALDH8A1* which is involved in Tryptophan metabolism (chx00380), and Metabolic pathways (chx01100); *HBS1L* which is involved in mRNA surveillance pathway (chx03015), and Legionellosis (chx05134).

Markers on chromosome 15 were linked to the genes *ARHGAP42; PGR* which is involved in Oocyte meiosis (chx04114), Progesterone-mediated oocyte maturation (chx04914), Estrogen signalling pathway (chx04915), Chemical carcinogenesis - receptor activation (chx05207), Breast cancer (chx05224). Markers on chromosome 16 were linked to the two genes *CACNA1E* which is involved in Glycerophospholipid metabolism (chx00564); and *ZNF648*. Markers on chromosome 19 were linked to several genes including *CRK* which is involved in MAPK signalling pathway (chx04010), ErbB signalling pathway (chx04012), Rap1 signalling pathway (chx04015), Chemokine signalling pathway (chx04062), Focal adhesion (chx04510), Fc gamma R-mediated phagocytosis (chx04666), Neurotrophin signalling pathway (chx04722), Regulation of actin cytoskeleton (chx04810), Insulin signalling pathway (chx04910), Growth hormone synthesis, secretion and action (chx04935), Bacterial invasion of epithelial cells (chx05100), Yersinia infection (chx05135), Human cytomegalovirus infection (chx05163), Human immunodeficiency virus 1 infection (chx05170), Pathways in cancer (chx05200), MicroRNAs in cancer (chx05206), Renal cell carcinoma (chx05211), Chronic myeloid leukemia (chx05220); *MYO1C* which is involved in Motor proteins (chx04814); *INPP5K* which is involved in Inositol phosphate metabolism (chx00562) and Metabolic pathways (chx01100); *PITPNA; SLC43A2; SCARF1; RILP* which is involved in Phagosome (chx04145) and Salmonella infection (chx05132); *PRPF8* which is involved in Spliceosome (chx03040); *TLCD2; WDR81; SERPINF2* which is involved in Complement and coagulation cascades (chx04610); *SERPINF1* which is involved in Wnt signalling pathway (chx04310); *SMYD4; RPA1* which is involved in DNA replication (chx03030), Nucleotide excision repair (chx03420), Mismatch repair (chx03430), Homologous recombination (chx03440), Fanconi anemia pathway (chx03460); *RTN4RL1*; and *DPH1*.

Table 5.3: Genes under selection within village meat-type goats with different coat colours and coat colour patterns and their associated pathways

Coat colour/pattern	Chr	Genes	Pathways
White	1	*ZBED2; CCDC80; PLCXD2; NEPRO; ABHD10; PHLDB2; SLC35A5; SLC9C1; BTLA; GTPBP8; CD96; CD200*	
	5	*HEBP1; GPR19; APOLD1; GPRC5A; MANSC1; RESF1; DUSP16; BICD1; CDKN1B; AMN1; BORCS5; CREBL2; GPRC5D; DDX47*	MAPK signalling pathway; Viral life cycle - HIV-1; Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer;
	6	*SULT1B1; CORIN; CNGA1; ODAM; ATP10D; UGT2A2; NFXL1; SULT1E1*	cGMP-PKG signalling pathway; cAMP signalling pathway; Phototransduction;
	8	*RANBP6; MLANA; UHRF2; LURAP1L; IL33; GLDC; KIAA2026; TYRP1; KDM4C*	Cytokine-cytokine receptor interaction; Necroptosis; Cytosolic DNA-sensing pathway; Influenza A; Glycine, serine and threonine metabolism; Glyoxylate and dicarboxylate metabolism; Lipoic acid metabolism; Metabolic pathways; Carbon metabolism; Tyrosine metabolism; Metabolic pathways; Melanogenesis
	10	*PTPN21; KCNK10; EML5; SPATA7*	Gastric acid secretion

	24	*B4GALT6; DSC1; ZNF532; TTR; GRP; DSG2; MALT1; DSG1; DSG3; ALPK2; DSG4*	Sphingolipid metabolism; Metabolic pathways; Thyroid hormone synthesis; Neuroactive ligand-receptor interaction; Arrhythmogenic right ventricular cardiomyopathy; NF-kappa B signalling pathway; C-type lectin receptor signalling pathway; T cell receptor signalling pathway; B cell receptor signalling pathway; Tuberculosis; Staphylococcus aureus infection;
Red-TX	1	*DIPK2A; SLC9A9*	
	4	*TRIL; CPVL; CHN2; PDE1C; LSM5; AVL9; KBTBD2; FKBP9; NT5C3A*	Transcriptional misregulation in cancer
	6	*UGT8; SULT1B1; SULT1E1; ODAM*	Ether lipid metabolism; Sphingolipid metabolism; Metabolic pathways
	8	*RORB; TRPM6; GNAQ; CEP78; PSAT1*	Circadian rhythm; Mineral absorption; Glycine, serine and threonine metabolism; Cysteine and methionine metabolism; Vitamin B6 metabolism; Metabolic pathways; Carbon metabolism; Biosynthesis of amino acids; Biosynthesis of cofactors
	10	*IQGAP2; SV2C*	Regulation of actin cytoskeleton; ECM-receptor interaction
	20	*BASP1; MYO10*	Fc gamma R-mediated phagocytosis; Motor proteins
	21	*RGMA; CHD2; FAM174B; TC2N; FBLN5; TRIP11; CPSF2*	TGF-beta signalling pathway; Axon guidance; mRNA surveillance pathway
Red-ZV	5	*CAPZA3; PLCZ1; PIK3C2G;*	Endocytosis; Motor proteins; Metabolic pathways; Calcium signalling pathway; Phosphatidylinositol signalling system; Oocyte meiosis; Thyroid hormone signalling pathway; Inositol phosphate metabolism; Metabolic pathways; Phosphatidylinositol signalling system; Salmonella infection
	6	*KDR; SRD5A3*	EGFR tyrosine kinase inhibitor resistance; MAPK signalling pathway; Ras signalling pathway; Rap1 signalling pathway; Calcium signalling pathway; PI3K-Akt signalling pathway; VEGF signalling pathway; Focal adhesion; Proteoglycans in cancer; Fluid shear stress and atherosclerosis; Steroid hormone biosynthesis; N-Glycan biosynthesis; Metabolic pathways
	8	*DMAC1; PUM3; KCNV2; VLDLR*	Spinocerebellar ataxia; Lipid and atherosclerosis
Black	6	*MMRN1; SNCA*	Alzheimer disease; Parkinson disease; Pathways of neurodegeneration - multiple diseases
	20	*EMB; FGF10*	MAPK signalling pathway; Ras signalling pathway; Rap1 signalling pathway; Calcium signalling pathway; PI3K-Akt signalling pathway; Regulation of actin cytoskeleton; Pathways in cancer; Chemical carcinogenesis - receptor activation; Melanoma; Breast cancer; Gastric cancer

	24	*CDH7; B4GALT6; TTR; DSG2; DSG3; DSG4; DSG1; DSC1*	Sphingolipid metabolism; Metabolic pathways; Thyroid hormone synthesis; Arrhythmogenic right ventricular cardiomyopathy
Grey	3	*LRRFIP1; RAB17; PRLH; MLPH; COL6A3*	Butanoate metabolism; Metabolic pathways; PI3K-Akt signalling pathway; Focal adhesion; ECM-receptor interaction; Protein digestion and absorption; Human papillomavirus infection
	8	*DMAC1*	
	9	*NCOA7; HEY2*	Lysosome; Notch signalling pathway; Human papillomavirus infection; Pathways in cancer; Breast cancer
Speckled-TX	6	*ODAM; CABS1; AMTN*	
	8	*LURAP1L; TYRP1; PTPRD; KCNV2; VLDLR; KANK1; DOCK8; CEMIP2*	Tyrosine metabolism; Metabolic pathways; Melanogenesis; Cell adhesion molecules; Spinocerebellar ataxia; Lipid and atherosclerosis;
	17	*SETD7; MGST2*	Lysine degradation; Metabolic pathways; FoxO signalling pathway; Glutathione metabolism; Metabolism of xenobiotics by cytochrome P450; Drug metabolism - cytochrome P450; Drug metabolism - other enzymes; Metabolic pathways; Platinum drug resistance; Pathways in cancer; Chemical carcinogenesis - DNA adducts; Chemical carcinogenesis - receptor activation; Chemical carcinogenesis - reactive oxygen species; Hepatocellular carcinoma; Fluid shear stress and atherosclerosis
	19	*CRK; MYO1C; INPP5K; PITPNA; SLC43A2; SCARF1; RILP; PRPF8; TLCD2; WDR81; SERPINF2; SERPINF1; SMYD4; RPA1; RTN4RL1; DPH1*	MAPK signalling pathway; ErbB signalling pathway; Rap1 signalling pathway; Chemokine signalling pathway; Focal adhesion; Fc gamma R-mediated phagocytosis; Neurotrophin signalling pathway; Regulation of actin cytoskeleton; Insulin signalling pathway; Growth hormone synthesis, secretion and action; Bacterial invasion of epithelial cells; Yersinia infection; Human cytomegalovirus infection; Human immunodeficiency virus 1 infection; Pathways in cancer; MicroRNAs in cancer; Renal cell carcinoma; Chronic myeloid leukemia; Motor proteins; Inositol phosphate metabolism; Metabolic pathways; Phagosome; Salmonella infection; Spliceosome; Complement and coagulation cascades; Wnt signalling pathway; DNA replication; Nucleotide excision repair; Mismatch repair; Homologous recombination; Fanconi anemia pathway
Speckled-ZV	5	*HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer; MAPK signalling pathway

	6	*SULT1B1; SULT1E1; ODAM*	
	19	*CRK; MYO1C; INPP5K; PITPNA; SLC43A2; SCARF1; RILP; PRPF8; TLCD2; WDR81; SERPINF2; SERPINF1; SMYD4; RPA1; RTN4RL1; DPH1*	MAPK signalling pathway; ErbB signalling pathway; Rap1 signalling pathway; Chemokine signalling pathway; Focal adhesion; Fc gamma R-mediated phagocytosis; Neurotrophin signalling pathway; Regulation of actin cytoskeleton; Insulin signalling pathway; Growth hormone synthesis, secretion and action; Bacterial invasion of epithelial cells; Yersinia infection; Human cytomegalovirus infection; Human immunodeficiency virus 1 infection; Pathways in cancer; MicroRNAs in cancer; Renal cell carcinoma; Chronic myeloid leukemia; Motor proteins; Inositol phosphate metabolism; Metabolic pathways; Phagosome; Salmonella infection; Spliceosome; Complement and coagulation cascades; Wnt signalling pathway; DNA replication; Nucleotide excision repair; Mismatch repair; Homologous recombination; Fanconi anemia pathway
	27	*FAT1; MTNR1A; F11; KLKB1; CYP4V2; TLR3*	Neuroactive ligand-receptor interaction; Circadian entrainment; Complement and coagulation cascades; Necroptosis; Toll-like receptor signalling pathway; Hepatitis C; Hepatitis B; Influenza A; Human papillomavirus infection; Kaposi sarcoma-associated herpesvirus infection; Herpes simplex virus 1 infection; Coronavirus disease - COVID-19
Patchy	3	*RBM44; LRRFIP1; RAB17; PRLH; MLPH; COL6A3*	Neuroactive ligand-receptor interaction; PI3K-Akt signalling pathway; Focal adhesion; ECM-receptor interaction; Protein digestion and absorption; Human papillomavirus infection
	6	*SMARCAD1; ATOH1; GRID2; MMRN1; SNCA*	Signalling pathways regulating pluripotency of stem cells; Neuroactive ligand-receptor interaction; Long-term depression; Alzheimer disease; Parkinson disease; Pathways of neurodegeneration - multiple diseases
	8	*LURAP1L; TYRP1*	Tyrosine metabolism; Metabolic pathways; Melanogenesis
Belted-TX	2	*XIRP2; B3GALT1; STK39; SP3*	Glycosphingolipid biosynthesis - lacto and neolacto series; Metabolic pathways;
	5	*BICD1; RESF1*	Viral life cycle - HIV-1;
	6	*ODAM; CABS1; AMTN*	
	8	*LURAP1L; TYRP1*	Tyrosine metabolism; Metabolic pathways; Melanogenesis
	9	*MAP3K7; GJA10; CASP8AP2; MDN1; LYRM2; ANKRD6; RRAGD; UBE2J1; RNGTT; CNR1; SYNCRIP; SNX14; NT5E; TBX18*	MAPK signalling pathway; NF-kappa B signalling pathway; Autophagy – animal; AMPK signalling pathway; Wnt signalling pathway; Osteoclast differentiation; Adherens junction; Neutrophil extracellular trap formation; Toll-like receptor signalling pathway; NOD-like receptor signalling pathway; RIG-I-like receptor signalling pathway; IL-17 signalling pathway; T cell receptor signalling pathway; TNF signalling pathway; Alcoholic liver disease; Salmonella infection; Yersinia infection;

			Leishmaniasis; Toxoplasmosis; Hepatitis B; Measles; Herpes simplex virus 1 infection; Epstein-Barr virus infection; Ribosome biogenesis in eukaryotes; Autophagy – animal; mTOR signalling pathway; Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum; Parkinson disease; Pathways of neurodegeneration - multiple diseases; mRNA surveillance pathway; Rap1 signalling pathway; Neuroactive ligand-receptor interaction; Thermogenesis; Retrograde endocannabinoid signalling; Purine metabolism; Pyrimidine metabolism; Nicotinate and nicotinamide metabolism; Metabolic pathways; Nucleotide metabolism
Belted-ZV	6	*ODAM; CABS1; AMTN*	
	19	*CRK; MYO1C; INPP5K; PITPNA; SLC43A2; SCARF1; RILP; PRPF8; TLCD2; WDR81; SERPINF2; SERPINF1; SMYD4; RPA1; RTN4RL1; DPH1*	MAPK signalling pathway; ErbB signalling pathway; Rap1 signalling pathway; Chemokine signalling pathway; Focal adhesion; Fc gamma R-mediated phagocytosis; Neurotrophin signalling pathway; Regulation of actin cytoskeleton; Insulin signalling pathway; Growth hormone synthesis, secretion and action; Bacterial invasion of epithelial cells; Motor proteins; Inositol phosphate metabolism; Metabolic pathways; Phagosome; Salmonella infection; Spliceosome; Complement and coagulation cascades; Wnt signalling pathway; DNA replication; Nucleotide excision repair; Mismatch repair; Homologous recombination; Fanconi anemia pathway
White sided	2	*FTCDNL1; SATB2*	
	4	*CREB5; TRIL; CPVL; CHN2; PDE1C; LSM5; AVL9; KBTBD2; FKBP9; NT5C3A; NPSR1; DPY19L1; DPY19L2; TBX20; HERPUD2; SEPTIN7; EEPD1; AOAH; ELMO1; WNT2; ASZ1; CFTR; CTTNBP2; SRPK2; KMT2E; ORC5; RELN; SEMA3E; SEMA3A; SEMA3D; AGMO*	cGMP-PKG signalling pathway; cAMP signalling pathway; PI3K-Akt signalling pathway; AMPK signalling pathway; Longevity regulating pathway; Adrenergic signalling in cardiomyocytes; TNF signalling pathway; Thermogenesis; Cholinergic synapse; Dopaminergic synapse; Insulin secretion; Estrogen signalling pathway; Thyroid hormone synthesis; Glucagon signalling pathway; Aldosterone synthesis and secretion; Relaxin signalling pathway; Cortisol synthesis and secretion; Parathyroid hormone synthesis, secretion and action; Insulin resistance; Cushing syndrome; Growth hormone synthesis, secretion and action; Vasopressin-regulated water reabsorption; Huntington disease; Prion disease; Cocaine addiction; Amphetamine addiction; Alcoholism; Hepatitis B; Human cytomegalovirus infection; Human papillomavirus infection; Human T-cell leukemia virus 1 infection; Viral carcinogenesis; Chemical carcinogenesis - receptor activation; Prostate cancer; Pyrimidine metabolism; Metabolic pathways; Nucleotide metabolism; Neuroactive ligand-receptor interaction; Chemokine signalling pathway; Bacterial invasion of epithelial cells; Salmonella infection; Yersinia infection; mTOR signalling pathway; Wnt signalling pathway; Hippo signalling pathway; Signalling pathways regulating pluripotency of stem cells; Melanogenesis; Cushing syndrome; Alzheimer disease; Pathways of neurodegeneration - multiple diseases; Human papillomavirus infection; Pathways in cancer; Proteoglycans in cancer; Basal cell carcinoma; Breast cancer; Hepatocellular carcinoma; Gastric cancer; ABC transporters; cAMP signalling pathway; AMPK signalling pathway; Tight junction; Gastric acid secretion; Pancreatic secretion; Bile secretion; Lysine degradation; Metabolic pathways; Cell cycle; PI3K-Akt signalling pathway; Focal adhesion; ECM-receptor interaction; Spinocerebellar ataxia;

			Human papillomavirus infection; PI3K-Akt signalling pathway; Focal adhesion; ECM-receptor interaction; Spinocerebellar ataxia; Human papillomavirus infection; Axon guidance;
	19	*CRK; MYO1C; INPP5K; PITPNA; SLC43A2; SCARF1; RILP; PRPF8; TLCD2; WDR81; SERPINF2; SERPINF1; SMYD4; RPA1; RTN4RL1; DPH1*	Motor proteins; Inositol phosphate metabolism; Metabolic pathways; Phagosome; Salmonella infection; Spliceosome; Complement and coagulation cascades; Wnt signalling pathway; DNA replication; Nucleotide excision repair; Mismatch repair; Homologous recombination; Fanconi anemia pathway
	20	*CDH9*	
Blacklegged	5	*CAPZA3; PLCZ1; PIK3C2G*	Endocytosis; Motor proteins; Inositol phosphate metabolism; Metabolic pathways; Calcium signalling pathway; Phosphatidylinositol signalling system; Oocyte meiosis; Thyroid hormone signalling pathway; Inositol phosphate metabolism; Metabolic pathways; Phosphatidylinositol signalling system; Salmonella infection
	7	*SLC36A2; SLC36A3; GM2A; CCDC69; TNIP1; GPX3; ZNF300; SMIM3; PCDH12; DELE1; PCDH1; ARAP3; FCHSD1; RELL2; HDAC3; DIAPH1; PCDHGC4; SNCAIP; SNX2*	Protein digestion and absorption; Lysosome; Glutathione metabolism; Metabolic pathways; Thyroid hormone synthesis; Amyotrophic lateral sclerosis; Huntington disease; Pathways of neurodegeneration - multiple diseases; Herpes simplex virus 1 infection; Rap1 signalling pathway; cAMP signalling pathway; Endocytosis; Cytokine-cytokine receptor interaction; Neutrophil extracellular trap formation; Thyroid hormone signalling pathway; Alcoholism; Viral carcinogenesis; Focal adhesion; Regulation of actin cytoskeleton; AGE-RAGE signalling pathway in diabetic complications; Parkinson disease; Pathways of neurodegeneration - multiple diseases; Endocytosis
	9	*CLVS2; FAM229B; TUBE1; CCN6; FYN; TRAF3IP2; BACH2; GJA10; CASP8AP2; MDN1; LYRM2; ANKRD6; ALDH8A1; HBS1L*	Motor proteins; Sphingolipid signalling pathway; Phospholipase D signalling pathway; Axon guidance; Osteoclast differentiation; Focal adhesion; Adherens junction; Platelet activation; Natural killer cell mediated cytotoxicity; T cell receptor signalling pathway; Fc epsilon RI signalling pathway; Cholinergic synapse; Prion disease; Viral myocarditis; Cellular senescence; IL-17 signalling pathway; Ribosome biogenesis in eukaryotes; Tryptophan metabolism; Metabolic pathways; mRNA surveillance pathway; Legionellosis;
	15	*ARHGAP42; PGR*	Oocyte meiosis; Progesterone-mediated oocyte maturation; Estrogen signalling pathway; Chemical carcinogenesis - receptor activation; Breast cancer
	16	*CACNA1E; ZNF648*	Glycerophospholipid metabolism
	19	*CRK; MYO1C; INPP5K; PITPNA; SLC43A2; SCARF1;*	MAPK signalling pathway; ErbB signalling pathway; Rap1 signalling pathway; Chemokine signalling pathway; Focal adhesion; Fc gamma R-mediated phagocytosis; Neurotrophin signalling pathway;

	RILP; PRPF8; TLCD2; WDR81; SERPINF2; SERPINF1; SMYD4; RPA1; RTN4RL1; DPH1	Regulation of actin cytoskeleton; Insulin signalling pathway; Growth hormone synthesis, secretion and action; Bacterial invasion of epithelial cells; Yersinia infection; Human cytomegalovirus infection; Human immunodeficiency virus 1 infection; Pathways in cancer; MicroRNAs in cancer; Renal cell carcinoma; Chronic myeloid leukemia; Motor proteins; Inositol phosphate metabolism; Metabolic pathways; Phagosome; Salmonella infection; Spliceosome; Complement and coagulation cascades; Wnt signalling pathway; DNA replication; Nucleotide excision repair; Mismatch repair; Homologous recombination; Fanconi anemia pathway

5.3.3 Signatures of selection (XP-EHH) between commercial breeds and village meat-type goats.

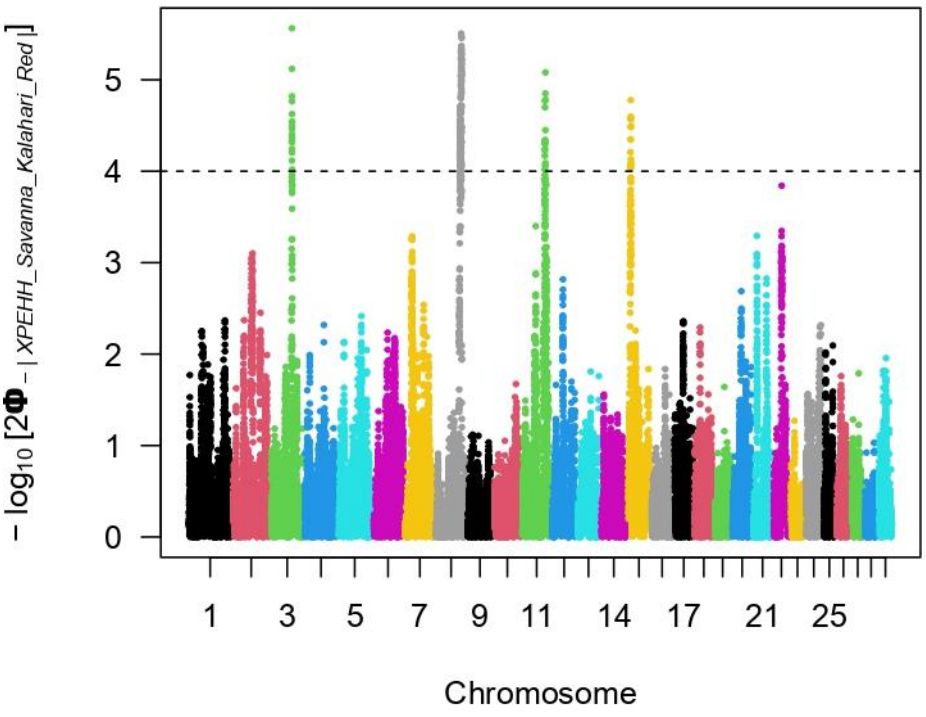

Figure 5.16: Selection signatures detected by an XP-EHH analysis between Savanna and Kalahari Red goats. The threshold for the detection of signatures was set to 4.

Between the Savanna and the Kalahari Red breeds signatures were found under selection on chromosome 8 and 15 in the Kalahari Red but not in the Savanna. The signals on chromosome 8 were associated to the following genes, *SHC3* which is involved in EGFR tyrosine kinase inhibitor resistance (chx01521), Endocrine resistance (chx01522), ErbB signalling pathway (chx04012), Ras signalling pathway (chx04014), Chemokine signalling pathway (chx04062), Phospholipase D signalling pathway (chx04072), Focal adhesion (chx04510), Natural killer cell mediated cytotoxicity (chx04650), Neurotrophin signalling pathway (chx04722), Insulin signalling pathway (chx04910), Estrogen signalling pathway (chx04915), Prolactin signalling pathway (chx04917), Relaxin signalling pathway (chx04926), Growth hormone synthesis, secretion and action (chx04935), Alcoholism (chx05034), Bacterial invasion of epithelial cells (chx05100), Glioma (chx05214), Chronic myeloid leukemia (chx05220), Breast cancer (chx05224), Hepatocellular carcinoma (chx05225), Gastric cancer (chx05226)*; S1PR3* which is involved in Sphingolipid signalling pathway (chx04071), and Neuroactive ligand-receptor interaction (chx04080)*; NXNL2; CDK20; MSANTD3; TMEFF1; CAVIN4; PLPPR1; BAAT* Primary bile acid biosynthesis (chx00120), Taurine and hypotaurine metabolism (chx00430), Biosynthesis of unsaturated fatty acids (chx01040), Metabolic pathways (chx01100),

Peroxisome (chx04146), and Bile secretion (chx04976)*; MRPL50; ZNF189* which is involved in Herpes simplex virus 1 infection (chx05168)*; ALDOB* which is involved in Glycolysis / Gluconeogenesis (chx00010), Pentose phosphate pathway (chx00030), Fructose and mannose metabolism (chx00051), Metabolic pathways (chx01100), Carbon metabolism (chx01200), Biosynthesis of amino acids (chx01230), and HIF-1 signalling pathway (chx04066)*; PGAP4; RNF20; GRIN3A* which is involved in Calcium signalling pathway (chx04020), cAMP signalling pathway (chx04024), Neuroactive ligand-receptor interaction (chx04080), Glutamatergic synapse (chx04724), Spinocerebellar ataxia (chx05017), Prion disease (chx05020), Cocaine addiction (chx05030), Amphetamine addiction (chx05031), Nicotine addiction (chx05033), and Alcoholism (chx05034)*; SMC2; OR13C4; OR13C3; OR13D1; NIPSNAP3A; ABCA1* which is involved in ABC transporters (chx02010), Fat digestion and absorption (chx04975), Cholesterol metabolism (chx04979), Lipid and atherosclerosis (chx05417)*; SLC44A1* which is involved in Choline metabolism in cancer (chx05231)*; FSD1L; FKTN* which is involved in Mannose type O-glycan biosynthesis (chx00515), and Metabolic pathways (chx01100)*; TAL2;* and *TMEM38B*. While signals on chromosome 15 were linked to the *TTC17*; and *API5* genes.

Signatures under selection in the Savanna but not in the Kalahari Red were found on chromosome 3 and 11. Signals on chromosome 3 were linked to the following genes, *SASS6; TRMT13; LRRC39; RTCA; CDC14A* which is involved in Cell cycle (chx04110)*; GPR88; EXTL2* which is involved in Glycosaminoglycan biosynthesis - heparan sulfate / heparin (chx00534), and Metabolic pathways (chx01100)*; SLC30A7; DPH5; S1PR1* which is involved in FoxO signalling pathway (chx04068), Sphingolipid signalling pathway (chx04071), and Neuroactive ligand-receptor interaction (chx04080). Signals on chromosome 11 were linked to the following genes, *TRIB2; LPIN1* which is involved in Glycerolipid metabolism (chx00561), Glycerophospholipid metabolism (chx00564), Metabolic pathways (chx01100), mTOR signalling pathway (chx04150), and Alcoholic liver disease (chx04936)*; NTSR2* which is involved in Neuroactive ligand-receptor interaction (chx04080)*; GREB1; E2F6* which is involved in Polycomb repressive complex (chx03083)*;*and *ROCK2* which is involved in several pathways including cGMP-PKG signalling pathway (chx04022), cAMP signalling pathway (chx04024), Chemokine signalling pathway (chx04062), Sphingolipid signalling pathway (chx04071), Vascular smooth muscle contraction (chx04270), Wnt signalling pathway (chx04310), Axon guidance (chx04360), Focal adhesion (chx04510), Adherens junction (chx04520), Tight junction (chx04530), Platelet activation (chx04611), Leukocyte

transendothelial migration (chx04670), Regulation of actin cytoskeleton (chx04810), Oxytocin signalling pathway (chx04921), Salmonella infection (chx05132), Yersinia infection (chx05135), Human cytomegalovirus infection (chx05163), Pathways in cancer (chx05200), Proteoglycans in cancer (chx05205), and Lipid and atherosclerosis (chx05417).

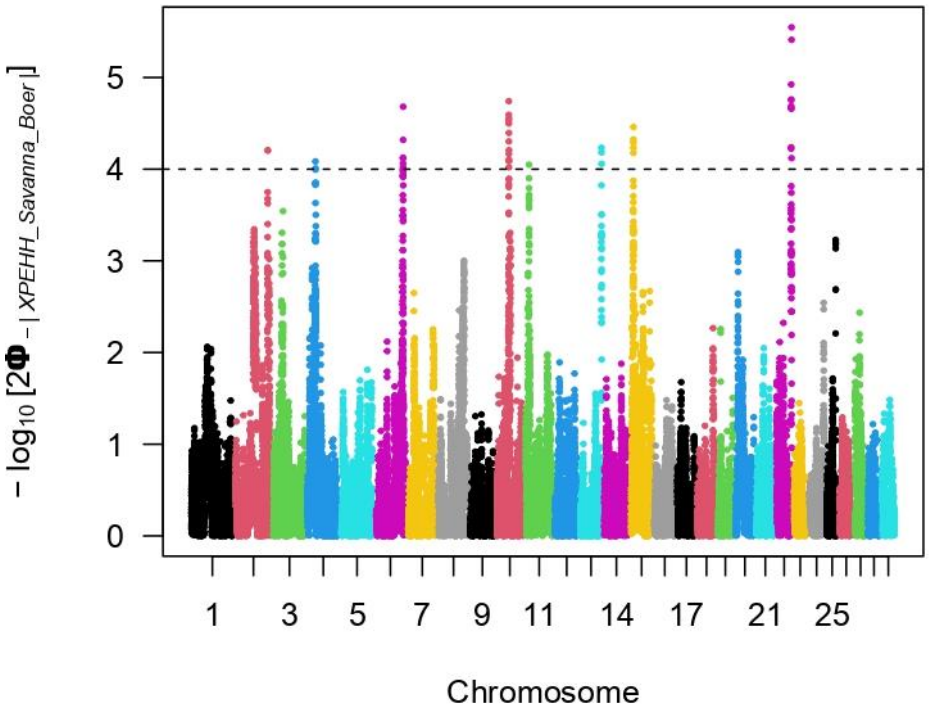

Figure 5.17: Selection signatures detected by an XP-EHH analysis between Savanna and Boer goats. The threshold for the detection of signatures was set to 4.

Between the Savanna and the Boer breed, signatures under selection within the Boer breed but not the Savanna were found on chromosome 4, 6, 10, and 11. Signals on chromosome 4 were linked to the *TMEM229A* gene. Signals on chromosome 6 were linked to the three genes *SLC2A9; WDR1*; and *ZNF518B*. Signals on chromosome 10 were linked to the *WDR72* gene. Signals on chromosome 11 were linked to the *LTBP1* gene which is involved in the TGF-beta signalling pathway. Signatures under selection in the Savanna but not the Boer breed were found in chromosome 2, 13, and 22. Significant signals on chromosome 2 were linked to three genes, *ITGA4* which is involved in PI3K-Akt signalling pathway (chx04151), Focal adhesion (chx04510), ECM-receptor interaction (chx04512), Cell adhesion molecules (chx04514), Hematopoietic cell lineage (chx04640), Leukocyte transendothelial migration (chx04670), Intestinal immune network for IgA production (chx04672), Regulation of actin cytoskeleton (chx04810), Yersinia infection (chx05135), Leishmaniasis (chx05140), Human papillomavirus infection (chx05165), Hypertrophic cardiomyopathy (chx05410), Arrhythmogenic right ventricular cardiomyopathy (chx05412), and Dilated cardiomyopathy (chx05414); *CERKL;*

and *NEUROD1* which is involved in Maturity onset diabetes of the young (chx04950). Signals on chromosome 13 were linked to the three genes, *TSHZ2; ZNF217*; and *BCAS1*. Signatures on chromosome 22 were linked to several genes including *PLXND1; H1-8; RHO; IFT122; MBD4* which is involved in Base excision repair (chx03410); *EFCAB12; RPL32* which is involved in Ribosome (chx03010), and Coronavirus disease - COVID-19 (chx05171); *CAND2; RAF1* which is involved in numerous pathways including MAPK signalling pathway (chx04010), Rap1 signalling pathway (chx04015), PI3K-Akt signalling pathway (chx04151), Estrogen signalling pathway (chx04915), and Melanogenesis (chx04916); *MKRN2; MKRN2OS; TSEN2; PPARG* which is involved in PPAR signalling pathway (chx03320), AMPK signalling pathway (chx04152), Longevity regulating pathway (chx04211), Osteoclast differentiation (chx04380), Thermogenesis (chx04714), Non-alcoholic fatty liver disease (chx04932), Huntington disease (chx05016), Pathways in cancer (chx05200), Transcriptional misregulation in cancer (chx05202), Thyroid cancer (chx05216), Lipid and atherosclerosis (chx05417); *FGD5*; *C3orf20*; *CCDC174*; *GRIP2*; and *SLC6A6*.

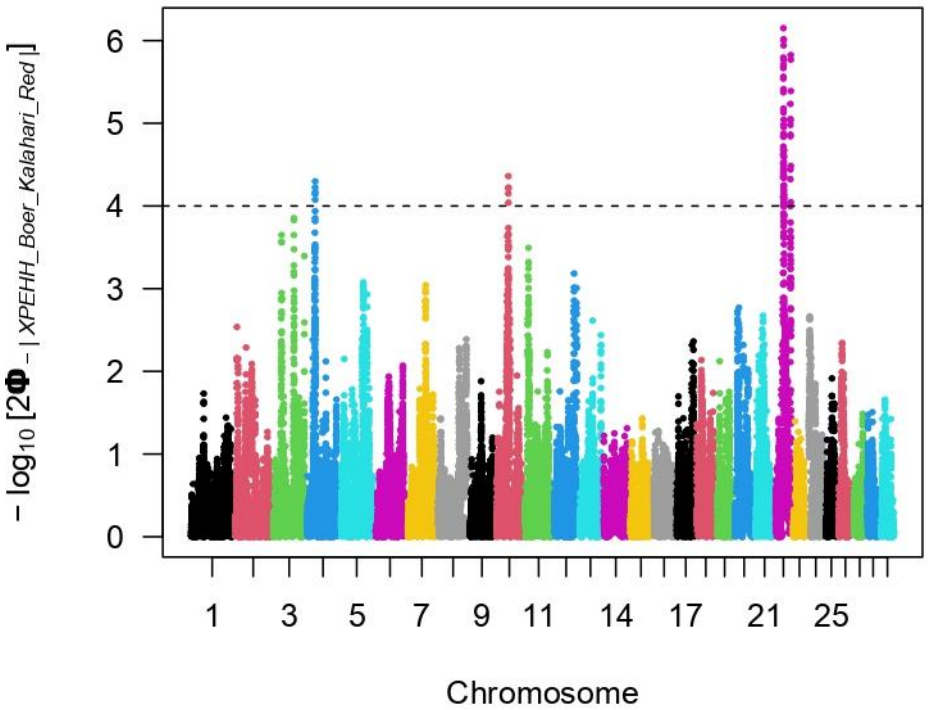

Figure 5.18: Selection signatures detected by an XP-EHH analysis between Kalahari Red and Boer goats. The threshold for the detection of signatures was set to 4.

Between the Kalahari Red and the Boer, signatures under selection in the Kalahari Red but not in the Boer were found on chromosome 22. These signatures were linked to several genes including *PLXND1; H1-8; RHO; IFT122; MBD4; EFCAB12; RPL32; CAND2; RAF1; MKRN2; MKRN2OS; TSEN2; PPARG; FGD5; C3orf20; CCDC174; GRIP2*; and *SLC6A6*. Signatures under selection in the Boer but not in the Kalahari Red breed were found on chromosome 4, 10, and 22. Signals on chromosome 4 were linked to the *TMEM229A* gene. Significant signatures on chromosome 22 were linked to numerous genes including, *PROK2*; *GPR27; EIF4E3; FOXP1* which is involved in MicroRNAs in cancer (chx05206); *MDFIC2; MITF* which is involved in Mitophagy - animal (chx04137), Osteoclast differentiation (chx04380), Melanogenesis (chx04916), Pathways in cancer (chx05200), Transcriptional misregulation in cancer (chx05202), and Melanoma (chx05218); *FRMD4B; LMOD3; ARL6IP5; UBA3* which is involved in Ubiquitin mediated proteolysis (chx04120); *TMF1*; *EOGT* which is involved in Other types of O-glycan biosynthesis (chx00514); and *TAFA4*.

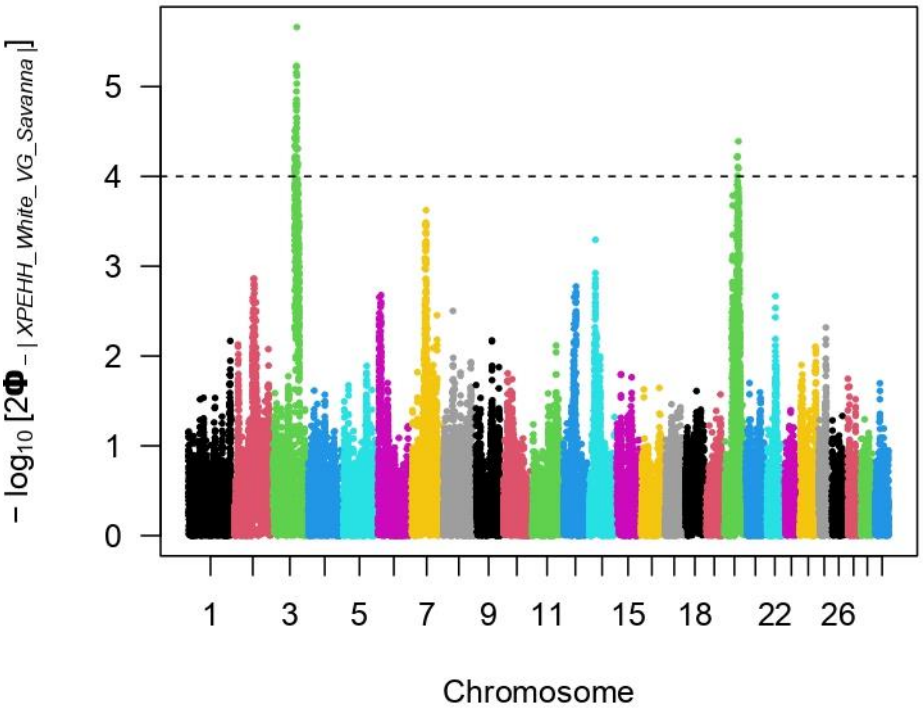

Figure 5.19: Selection signatures detected by an XP-EHH analysis between Savanna and white village goats. The threshold for the detection of signatures was set to 4.

Between the Savanna and the white village goats, signatures under selection in the Savanna but not in the white village goats were found on chromosome 3 and 20. Signals on chromosome 3 were linked to numerous genes including *RTCA; CDC14A* which is involved in Cell cycle (chx04110); *GPR88; EXTL2* which is involved in Glycosaminoglycan biosynthesis - heparan sulfate / heparin (chx00534), Metabolic pathways (chx01100); *SLC30A7; DPH5; S1PR1*

which is involved in FoxO signalling pathway (chx04068), Sphingolipid signalling pathway (chx04071), Neuroactive ligand-receptor interaction (chx04080); *OLFM3; COL11A1* which is involved in Protein digestion and absorption (chx04974); *RNPC3; PRMT6; NTNG1* which is involved in Axon guidance (chx04360), and Cell adhesion molecules (chx04514); *VAV3* which is involved in Rap1 signalling pathway (chx04015), cAMP signalling pathway (chx04024), Chemokine signalling pathway (chx04062), Focal adhesion (chx04510), Natural killer cell mediated cytotoxicity (chx04650), T cell receptor signalling pathway (chx04660), B cell receptor signalling pathway (chx04662), Fc epsilon RI signalling pathway (chx04664), Fc gamma R-mediated phagocytosis (chx04666), Leukocyte transendothelial migration (chx04670), Regulation of actin cytoskeleton (chx04810), Yersinia infection (chx05135), Proteoglycans in cancer (chx05205), and Lipid and atherosclerosis (chx05417); *SLC25A24; PROK1; CYM; KCNA10; KCNA2; KCNA3; CD53; OVN; WDR77; ATP5PB; C1orf162; TMIGD3; ADORA3* which is involved in cGMP-PKG signalling pathway (chx04022), Sphingolipid signalling pathway (chx04071), Neuroactive ligand-receptor interaction (chx04080); *RAP1A* which is involved in MAPK signalling pathway (chx04010), Ras signalling pathway (chx04014), Rap1 signalling pathway (chx04015), cAMP signalling pathway (chx04024), Chemokine signalling pathway (chx04062), Focal adhesion (chx04510), Adherens junction (chx04520), Tight junction (chx04530), Platelet activation (chx04611), Leukocyte transendothelial migration (chx04670), Long-term potentiation (chx04720), Neurotrophin signalling pathway (chx04722), Cushing syndrome (chx04934), Pancreatic secretion (chx04972), Renal cell carcinoma (chx05211), and Lipid and atherosclerosis (chx05417); *INKA2; DDX20*; and *KCND3* which is involved in Spinocerebellar ataxia (chx05017). Signatures on chromosome 20 were linked to the *CDH9* gene.

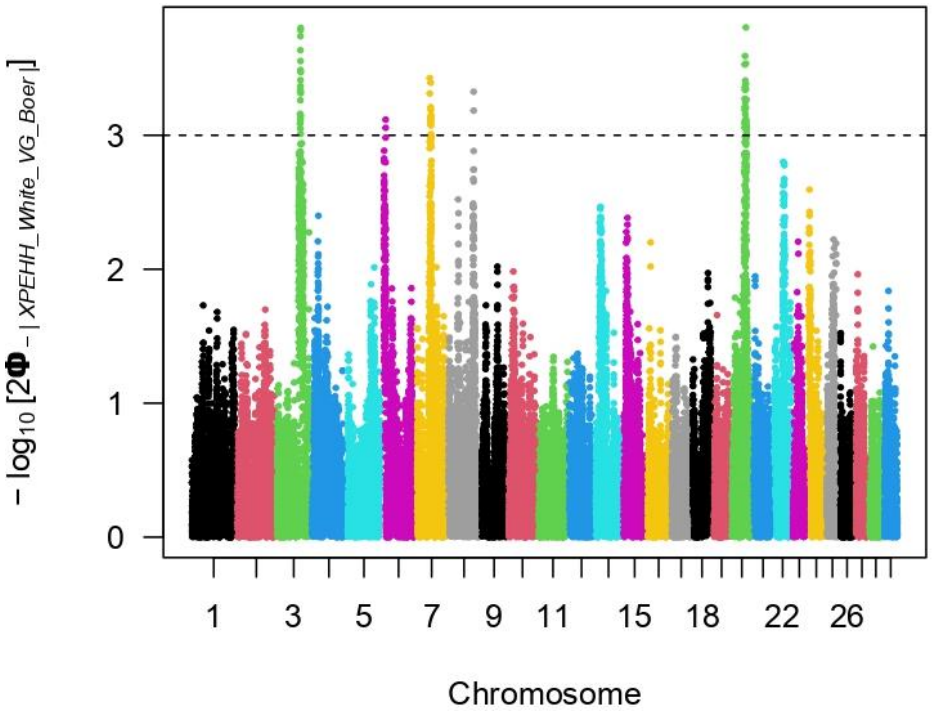

Figure 5.20: Selection signatures detected by an XP-EHH analysis between Boer and white village goats. The threshold for the detection of signatures was set to 4.

Between the Boer breed and the white village goats, signatures of selection were found in the Boer breed but not in the white village goats. These signals were found on chromosome 3, 6, 7, 8, and 20. Signals on chromosome 3 were linked to the following genes *PRMT6; NTNG1*, and *VAV3*. Signals on chromosome 6 were linked to the *ANK2* gene. Signals on chromosome 7 were linked to the following genes including *GLRA1* which is involved in Neuroactive ligand-receptor interaction (chx04080); *G3BP1; SPARC; FAT2; SLC36A1; SLC36A2* which is involved in Protein digestion and absorption (chx04974); *SLC36A3; SPINK6; SPINK5; SCGB3A2; SPINK1; JAKMIP2; DPYSL3; STK32A*; and *PPP2R2B* which is involved in mRNA surveillance pathway (chx03015), Sphingolipid signalling pathway (chx04071), PI3K-Akt signalling pathway (chx04151), AMPK signalling pathway (chx04152), Adrenergic signalling in cardiomyocytes (chx04261), Hippo signalling pathway (chx04390), Tight junction (chx04530), T cell receptor signalling pathway (chx04660), Dopaminergic synapse (chx04728), Chagas disease (chx05142), Hepatitis C (chx05160), and Human papillomavirus infection (chx05165). Signals on chromosomes 8 were linked to the *PLPPR1*; and *BAAT* genes. While signals on chromosome 20 were linked to the *CDH9*; and *CDH12* genes.

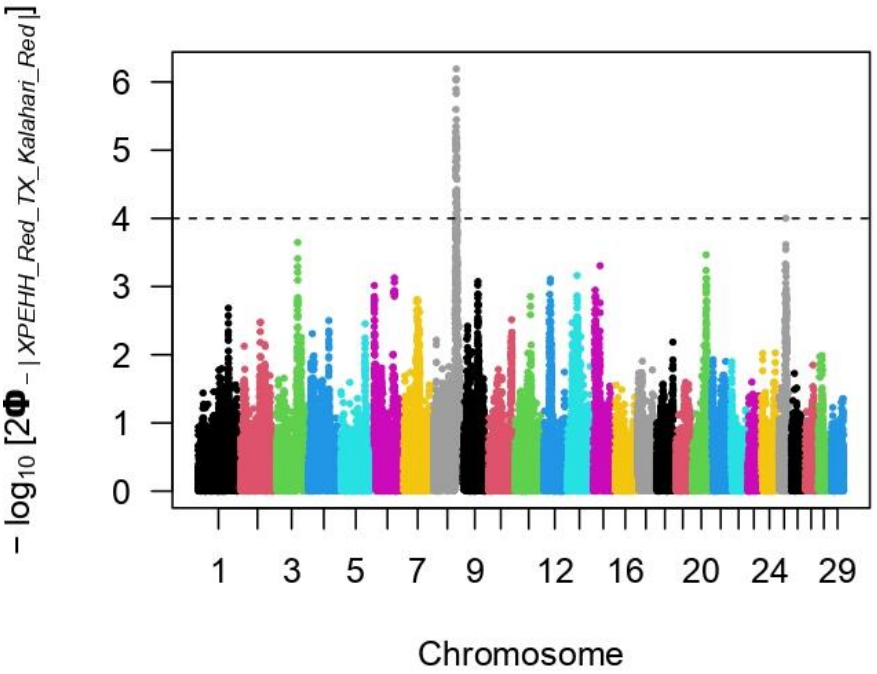

Figure 5.21: Selection signatures detected by an XP-EHH analysis between Kalahari Red and red village goats within the TX cluster. The threshold for the detection of signatures was set to 4.

Between the Kalahari Red breed and red village goats within the TX cluster, Significant signals under selection were found in the Kalahari Red breed on chromosome 8. Signatures on chromosome 8 were linked to several genes including *GADD45G* which is involved in numerous pathways including MAPK signalling pathway (chx04010), NF-kappa B signalling pathway (chx04064), FoxO signalling pathway (chx04068), Cell cycle (chx04110), p53 signalling pathway (chx04115), Apoptosis (chx04210), Cellular senescence (chx04218), Epstein-Barr virus infection (chx05169), Pathways in cancer (chx05200), Transcriptional misregulation in cancer (chx05202), Colorectal cancer (chx05210), Pancreatic cancer (chx05212), Endometrial cancer (chx05213), Glioma (chx05214), Thyroid cancer (chx05216), Basal cell carcinoma (chx05217), Melanoma (chx05218), Chronic myeloid leukemia (chx05220), Small cell lung cancer (chx05222), Non-small cell lung cancer (chx05223), Breast cancer (chx05224), Hepatocellular carcinoma (chx05225), Gastric cancer (chx05226), *SEMA4D* which is involved in Axon guidance (chx04360), *SECISBP2*, *CKS2* which is involved in numerous pathways including EGFR tyrosine kinase inhibitor resistance (chx01521), Endocrine resistance (chx01522), ErbB signalling pathway (chx04012), Ras signalling pathway (chx04014), Chemokine signalling pathway (chx04062), Phospholipase D signalling pathway (chx04072), Focal adhesion (chx04510), Natural killer cell mediated cytotoxicity (chx04650), Neurotrophin signalling pathway (chx04722), Insulin signalling pathway (chx04910), Estrogen signalling pathway (chx04915), Prolactin signalling pathway

(chx04917), Relaxin signalling pathway (chx04926), Growth hormone synthesis, secretion and action (chx04935), Alcoholism (chx05034), Bacterial invasion of epithelial cells (chx05100), Glioma (chx05214), Chronic myeloid leukemia (chx05220), Breast cancer (chx05224), Hepatocellular carcinoma (chx05225), Gastric cancer (chx05226); *SHC3* which is involved in several pathways including EGFR tyrosine kinase inhibitor resistance (chx01521), Endocrine resistance (chx01522), ErbB signalling pathway (chx04012), Ras signalling pathway (chx04014), Chemokine signalling pathway (chx04062), Phospholipase D signalling pathway (chx04072), Focal adhesion (chx04510), Natural killer cell mediated cytotoxicity (chx04650), Neurotrophin signalling pathway (chx04722), Insulin signalling pathway (chx04910), Estrogen signalling pathway (chx04915), Prolactin signalling pathway (chx04917), Relaxin signalling pathway (chx04926), Growth hormone synthesis, secretion and action (chx04935), Alcoholism (chx05034), Bacterial invasion of epithelial cells (chx05100), Glioma (chx05214), Chronic myeloid leukemia (chx05220), Breast cancer (chx05224), Hepatocellular carcinoma (chx05225), and Gastric cancer (chx05226); *S1PR3* which is involved in Sphingolipid signalling pathway (chx04071), and Neuroactive ligand-receptor interaction (chx04080); *NXNL2; CDK20; MSANTD3; TMEFF1; CAVIN4; PLPPR1; BAAT; MRPL50; ZNF189* which is involved in Herpes simplex virus 1 infection (chx05168); *ALDOB* which is involved in Glycolysis / Gluconeogenesis (chx00010), Pentose phosphate pathway (chx00030), Fructose and mannose metabolism (chx00051), Metabolic pathways (chx01100), Carbon metabolism (chx01200), Biosynthesis of amino acids (chx01230), and HIF-1 signalling pathway (chx04066); *PGAP4; RNF20; GRIN3A* which is involved in Calcium signalling pathway (chx04020), cAMP signalling pathway (chx04024), Neuroactive ligand-receptor interaction (chx04080), Glutamatergic synapse (chx04724), Spinocerebellar ataxia (chx05017), Prion disease (chx05020), Cocaine addiction (chx05030), Amphetamine addiction (chx05031), Nicotine addiction (chx05033), Alcoholism (chx05034); and *SMC2*.

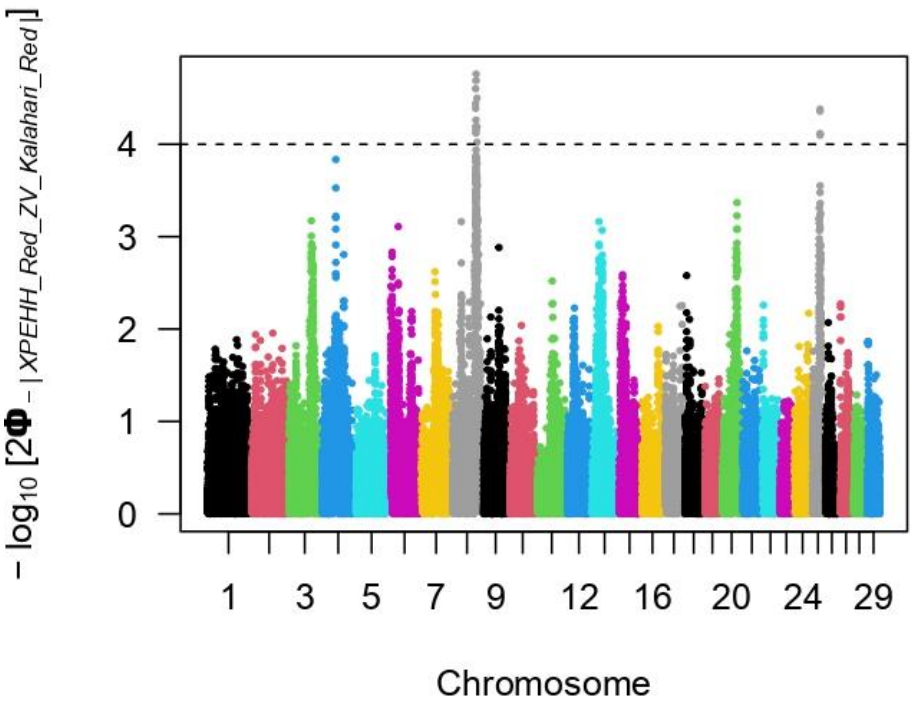

Figure 5.22: Selection signatures detected by an XP-EHH analysis between Kalahari Red and red village goats within the ZV cluster. The threshold for the detection of signatures was set to 4.

Between the Kalahari Red breed and red village goats within the ZV cluster, Significant signals under selection were found in the Kalahari Red breed on chromosome 8 and 25. Markers on chromosome 8 were similar to those found between Kalahari Red breed and red village goats in the TX cluster. Whereas, signals on chromosome 25 were linked to numerous genes, including *MRPS17*; *PSPH* which is involved in Glycine, serine and threonine metabolism (chx00260), Metabolic pathways (chx01100), Carbon metabolism (chx01200), and Biosynthesis of amino acids (chx01230); *SUMF2; PHKG1* which is involved in Calcium signalling pathway (chx04020), Insulin signalling pathway (chx04910), and Glucagon signalling pathway (chx04922); *NUPR2; VKORC1L1* which is involved in Ubiquinone and other terpenoid-quinone biosynthesis (chx00130), Metabolic pathways (chx01100), Biosynthesis of cofactors (chx01240); *ASL* which is involved in Arginine biosynthesis (chx00220), Alanine, aspartate and glutamate metabolism (chx00250), Metabolic pathways (chx01100), and Biosynthesis of amino acids (chx01230); *CRCP* which is involved in RNA polymerase (chx03020); *TPST1; CALN1; GALNT17*; and *AUTS2* which is involved in Polycomb repressive complex (chx03083).

Table 5.4: Genes under selection between commercial and village meat-type goats with different coat colours and coat colour patterns and their associated pathways

Breed		Chr	Genes	Pathways
Savanna vs Kalahari Red	Kalahari Red	8	*SHC3; S1PR3; NXNL2; CDK20; MSANTD3; TMEFF1; CAVIN4; PLPPR1; BAAT; MRPL50; ZNF189; ALDOB; PGAP4; RNF20; GRIN3A; SMC2; OR13C4; OR13C3; OR13D1; NIPSNAP3A; ABCA1; SLC44A1; FSD1L; FKTN; TAL2; TMEM38B*	EGFR tyrosine kinase inhibitor resistance; Endocrine resistance; ErbB signalling pathway; Ras signalling pathway; Chemokine signalling pathway; Phospholipase D signalling pathway; Focal adhesion; Natural killer cell mediated cytotoxicity; Neurotrophin signalling pathway; Insulin signalling pathway; Estrogen signalling pathway; Prolactin signalling pathway; Relaxin signalling pathway; Growth hormone synthesis, secretion and action; Alcoholism; Bacterial invasion of epithelial cells; Glioma; Chronic myeloid leukemia; Breast cancer; Hepatocellular carcinoma; Gastric cancer; Sphingolipid signalling pathway; Neuroactive ligand-receptor interaction; Primary bile acid biosynthesis; Taurine and hypotaurine metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Peroxisome; Bile secretion; Herpes simplex virus 1 infection; Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Fructose and mannose metabolism; Metabolic pathways; Carbon metabolism; Biosynthesis of amino acids; HIF-1 signalling pathway; Calcium signalling pathway; cAMP signalling pathway; Neuroactive ligand-receptor interaction; Glutamatergic synapse; Spinocerebellar ataxia; Prion disease; Cocaine addiction; Amphetamine addiction; Nicotine addiction; Alcoholism; ABC transporters; Fat digestion and absorption; Cholesterol metabolism; Lipid and atherosclerosis; Choline metabolism in cancer; Mannose type O-glycan biosynthesis; Metabolic pathways
		15	*TTC17; API5*	
	Savanna	3	*SASS6; TRMT13; LRRC39; RTCA; CDC14A; GPR88; EXTL2; SLC30A7; DPH5; S1PR1*	Cell cycle; Glycosaminoglycan biosynthesis - heparan sulfate / heparin; Metabolic pathways; FoxO signalling pathway; Sphingolipid signalling pathway; Neuroactive ligand-receptor interaction
		11	*TRIB2; LPIN1; NTSR2; GREB1; E2F6; ROCK2*	Glycerolipid metabolism; Glycerophospholipid metabolism; Metabolic pathways; mTOR signalling pathway; Alcoholic liver disease; Neuroactive ligand-receptor interaction; Polycomb repressive complex; cGMP-PKG signalling pathway; cAMP signalling pathway; Chemokine signalling pathway; Sphingolipid signalling pathway; Vascular smooth muscle contraction; Wnt signalling pathway; Axon guidance; Focal adhesion; Adherens junction; Tight junction; Platelet activation; Leukocyte transendothelial migration; Regulation of actin cytoskeleton; Oxytocin signalling pathway; Salmonella infection; Yersinia infection; Human cytomegalovirus infection; Pathways in cancer; Proteoglycans in cancer; Lipid and atherosclerosis
Savanna vs Boer	Boer	4	*TMEM229A*	
		6	*SLC2A9; WDR1; ZNF518B*	
		10	*WDR72*	
		11	*LTBP1*	TGF-beta signalling pathway
		15		

	Savanna	2	*ITGA4; CERKL; NEUROD1*	PI3K-Akt signalling pathway; Focal adhesion; ECM-receptor interaction; Cell adhesion molecules; Hematopoietic cell lineage; Leukocyte transendothelial migration; Intestinal immune network for IgA production; Regulation of actin cytoskeleton; Yersinia infection; Leishmaniasis; Human papillomavirus infection; Hypertrophic cardiomyopathy; Arrhythmogenic right ventricular cardiomyopathy; Dilated cardiomyopathy; Maturity onset diabetes of the young
		13	*TSHZ2; ZNF217; BCAS1*	
		22	*PLXND1; H1-8; RHO; IFT122; MBD4; EFCAB12; RPL32; CAND2; RAF1; MKRN2; MKRN2OS; TSEN2; PPARG; FGD5; C3orf20; CCDC174; GRIP2; SLC6A6*	Base excision repair; Ribosome; Coronavirus disease - COVID-19; EGFR tyrosine kinase inhibitor resistance; Endocrine resistance; MAPK signalling pathway; ErbB signalling pathway; Ras signalling pathway; Rap1 signalling pathway; cGMP-PKG signalling pathway; cAMP signalling pathway; Chemokine signalling pathway; FoxO signalling pathway; Sphingolipid signalling pathway; Phospholipase D signalling pathway; Autophagy – animal; mTOR signalling pathway; PI3K-Akt signalling pathway; Apoptosis; Cellular senescence; Vascular smooth muscle contraction; Axon guidance; VEGF signalling pathway; Apelin signalling pathway; Focal adhesion; Gap junction; Signalling pathways regulating pluripotency of stem cells; Neutrophil extracellular trap formation; C-type lectin receptor signalling pathway; JAK-STAT signalling pathway; Natural killer cell mediated cytotoxicity; T cell receptor signalling pathway; B cell receptor signalling pathway; Fc epsilon RI signalling pathway; Fc gamma R-mediated phagocytosis; Long-term potentiation; Neurotrophin signalling pathway; Serotonergic synapse; Long-term depression; Regulation of actin cytoskeleton; Insulin signalling pathway; GnRH signalling pathway; Progesterone-mediated oocyte maturation; Estrogen signalling pathway; Melanogenesis; Prolactin signalling pathway; Thyroid hormone signalling pathway; Oxytocin signalling pathway; Relaxin signalling pathway; Parathyroid hormone synthesis, secretion and action; GnRH secretion; Growth hormone synthesis, secretion and action; Alzheimer disease; Pathways of neurodegeneration - multiple diseases; Alcoholism; Salmonella infection; Tuberculosis; Hepatitis C; Hepatitis B; Human cytomegalovirus infection; Influenza A; Human papillomavirus infection; Kaposi sarcoma-associated herpesvirus infection; Human immunodeficiency virus 1 infection; Pathways in cancer; Proteoglycans in cancer; MicroRNAs in cancer; Chemical carcinogenesis - receptor activation; Chemical carcinogenesis - reactive oxygen species; Colorectal cancer; Renal cell carcinoma; Pancreatic cancer; Endometrial cancer; Glioma; Prostate cancer; Melanoma; Bladder cancer; Chronic myeloid leukemia; Acute myeloid leukemia; Non-small cell lung cancer; Breast cancer; Hepatocellular carcinoma; Gastric cancer; Central carbon metabolism in cancer; Choline metabolism in cancer; PD-L1 expression and PD-1 checkpoint pathway in cancer; PPAR signalling pathway; AMPK signalling pathway; Longevity regulating pathway; Osteoclast differentiation; Thermogenesis; Non-alcoholic fatty liver disease; Huntington disease; Pathways in cancer; Transcriptional misregulation in cancer; Thyroid cancer; Lipid and atherosclerosis
Boer vs Kalahari Red	Kalahari Red	22	*PLXND1; H1-8; RHO; IFT122; MBD4; EFCAB12; RPL32; CAND2; RAF1; MKRN2; MKRN2OS; TSEN2; PPARG; FGD5; C3orf20; CCDC174; GRIP2; SLC6A6*	Base excision repair; Ribosome; Coronavirus disease - COVID-19; EGFR tyrosine kinase inhibitor resistance; Endocrine resistance; MAPK signalling pathway; ErbB signalling pathway; Ras signalling pathway; Rap1 signalling pathway; cGMP-PKG signalling pathway; cAMP signalling pathway; Chemokine signalling pathway; FoxO signalling pathway; Sphingolipid signalling pathway; Phospholipase D signalling pathway; Autophagy – animal; mTOR signalling pathway; PI3K-Akt signalling pathway; Apoptosis; Cellular senescence; Vascular smooth muscle contraction; Axon guidance; VEGF signalling

				pathway; Apelin signalling pathway; Focal adhesion; Gap junction; Signalling pathways regulating pluripotency of stem cells; Neutrophil extracellular trap formation; C-type lectin receptor signalling pathway; JAK-STAT signalling pathway; Natural killer cell mediated cytotoxicity; T cell receptor signalling pathway; B cell receptor signalling pathway; Fc epsilon RI signalling pathway; Fc gamma R-mediated phagocytosis; Long-term potentiation; Neurotrophin signalling pathway; Serotonergic synapse; Long-term depression; Regulation of actin cytoskeleton; Insulin signalling pathway; GnRH signalling pathway; Progesterone-mediated oocyte maturation; Estrogen signalling pathway; Melanogenesis; Prolactin signalling pathway; Thyroid hormone signalling pathway; Oxytocin signalling pathway; Relaxin signalling pathway; Parathyroid hormone synthesis, secretion and action; GnRH secretion; Growth hormone synthesis, secretion and action; Alzheimer disease; Pathways of neurodegeneration - multiple diseases; Alcoholism; Salmonella infection; Tuberculosis; Hepatitis C; Hepatitis B; Human cytomegalovirus infection; Influenza A; Human papillomavirus infection; Kaposi sarcoma-associated herpesvirus infection; Human immunodeficiency virus 1 infection; Pathways in cancer; Proteoglycans in cancer; MicroRNAs in cancer; Chemical carcinogenesis - receptor activation; Chemical carcinogenesis - reactive oxygen species; Colorectal cancer; Renal cell carcinoma; Pancreatic cancer; Endometrial cancer; Glioma; Prostate cancer; Melanoma; Bladder cancer; Chronic myeloid leukemia; Acute myeloid leukemia; Non-small cell lung cancer; Breast cancer; Hepatocellular carcinoma; Gastric cancer; Central carbon metabolism in cancer; Choline metabolism in cancer; PD-L1 expression and PD-1 checkpoint pathway in cancer; PPAR signalling pathway; AMPK signalling pathway; Longevity regulating pathway; Osteoclast differentiation; Thermogenesis; Non-alcoholic fatty liver disease; Huntington disease; Pathways in cancer; Transcriptional misregulation in cancer; Thyroid cancer; Lipid and atherosclerosis
	Boer	4	*TMEM229A*	
		10		
		22	*PROK2; GPR27; EIF4E3; FOXP1; MDFIC2; MITF; FRMD4B; LMOD3; ARL6IP5; UBA3; TMF1; EOGT; TAFA4*	MicroRNAs in cancer; Mitophagy – animal; Osteoclast differentiation; Melanogenesis; Pathways in cancer; Transcriptional misregulation in cancer; Melanoma; Ubiquitin mediated proteolysis; Other types of O-glycan biosynthesis;
White vs Savanna	White			
	Savanna	3	*RTCA; CDC14A; GPR88; EXTL2; SLC30A7; DPH5; S1PR1; OLFM3; COL11A1; RNPC3; PRMT6; NTNG1; VAV3; SLC25A24; PROK1; CYM; KCNA10; KCNA2; KCNA3; CD53; OVN; WDR77; ATP5PB; C1orf162;*	Cell cycle; Glycosaminoglycan biosynthesis - heparan sulfate / heparin; Metabolic pathways; FoxO signalling pathway; Sphingolipid signalling pathway; Neuroactive ligand-receptor interaction; Protein digestion and absorption; Axon guidance; Cell adhesion molecules; Rap1 signalling pathway; cAMP signalling pathway; Chemokine signalling pathway; Focal adhesion; Natural killer cell mediated cytotoxicity; T cell receptor signalling pathway; B cell receptor signalling pathway; Fc epsilon RI signalling pathway; Fc gamma R-mediated phagocytosis; Leukocyte transendothelial migration; Regulation of actin cytoskeleton; Yersinia infection; Proteoglycans in cancer; Lipid and atherosclerosis; cGMP-PKG signalling pathway; Sphingolipid signalling pathway; Neuroactive ligand-receptor interaction; MAPK signalling

			TMIGD3; ADORA3; RAP1A; INKA2; DDX20; KCND3	pathway; Ras signalling pathway; Rap1 signalling pathway; cAMP signalling pathway; Chemokine signalling pathway; Focal adhesion; Adherens junction; Tight junction; Platelet activation; Leukocyte transendothelial migration; Long-term potentiation; Neurotrophin signalling pathway; Cushing syndrome; Pancreatic secretion; Renal cell carcinoma; Lipid and atherosclerosis; Spinocerebellar ataxia
		20	*CDH9*	
White vs Boer	White			
	Boer	3	*PRMT6; NTNG1; VAV3*	Axon guidance; Cell adhesion molecules; Rap1 signalling pathway; cAMP signalling pathway; Chemokine signalling pathway; Focal adhesion; Natural killer cell mediated cytotoxicity; T cell receptor signalling pathway; B cell receptor signalling pathway; Fc epsilon RI signalling pathway; Fc gamma R-mediated phagocytosis; Leukocyte transendothelial migration; Regulation of actin cytoskeleton; Yersinia infection; Proteoglycans in cancer; Lipid and atherosclerosis
		6	*ANK2*	Proteoglycans in cancer
		7	*GLRA1; G3BP1; SPARC; FAT2; SLC36A1; SLC36A2; SLC36A3; SPINK6; SPINK5; SCGB3A2; SPINK1; JAKMIP2; DPYSL3; STK32A; PPP2R2B*	Neuroactive ligand-receptor interaction; Protein digestion and absorption; mRNA surveillance pathway; Sphingolipid signalling pathway; PI3K-Akt signalling pathway; AMPK signalling pathway; Adrenergic signalling in cardiomyocytes; Hippo signalling pathway; Tight junction; T cell receptor signalling pathway; Dopaminergic synapse; Chagas disease; Hepatitis C; Human papillomavirus infection
		8	*PLPPR1; BAAT*	Primary bile acid biosynthesis; Taurine and hypotaurine metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Peroxisome; Bile secretion
		20	*CDH9; CDH12*	
Red-TX vs Kalahari Red	Red-TX			
	Kalahari Red	8	*GADD45G; SEMA4D; SECISBP2; CKS2; SHC3; S1PR3; NXNL2; CDK20; MSANTD3; TMEFF1; CAVIN4; PLPPR1; BAAT; MRPL50; ZNF189; ALDOB; PGAP4; RNF20; GRIN3A; SMC2*	MAPK signalling pathway; NF-kappa B signalling pathway; FoxO signalling pathway; Cell cyclel p53 signalling pathway; Apoptosis; Cellular senescence; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Colorectal cancer; Pancreatic cancer; Endometrial cancer; Glioma; Thyroid cancer; Basal cell carcinoma; Melanoma; Chronic myeloid leukemia; Small cell lung cancer; Non-small cell lung cancer; Breast cancer; Hepatocellular carcinoma; Gastric cancer; Axon guidance; Pathways in cancer; Small cell lung cancer; Primary bile acid biosynthesis; Taurine and hypotaurine metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Peroxisome; Bile secretion; Herpes simplex virus 1 infection; Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Fructose and mannose metabolism; Metabolic pathways; Carbon metabolism; Biosynthesis of amino acids; HIF-1 signalling pathway; Calcium signalling pathway; cAMP signalling pathway; Neuroactive ligand-receptor interaction; Glutamatergic synapse; Spinocerebellar ataxia; Prion disease; Cocaine addiction; Amphetamine addiction; Nicotine addiction; Alcoholism

		25	*MRPS17; PSPH; SUMF2; PHKG1; NUPR2; VKORC1L1; ASL; CRCP; TPST1*	Glycine, serine and threonine metabolism; Metabolic pathways; Carbon metabolism; Biosynthesis of amino acids; Calcium signalling pathway; Insulin signalling pathway; Glucagon signalling pathway; terpenoid-quinone biosynthesis; Metabolic pathways; Biosynthesis of cofactors; Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Metabolic pathways; Biosynthesis of amino acids; RNA polymerase
Red-ZV vs Kalahari Red	Red-ZV			
	Kalahari Red	8	*SEMA4D; SECISBP2; CKS2; SHC3; S1PR3; MSANTD3; TMEFF1; CAVIN4; PLPPR1; BAAT; MRPL50; SMC2*	Axon guidance; Pathways in cancer; Small cell lung cancer; EGFR tyrosine kinase inhibitor resistance; Endocrine resistance; ErbB signalling pathway; Ras signalling pathway; Chemokine signalling pathway; Phospholipase D signalling pathway; Focal adhesion; Natural killer cell mediated cytotoxicity; Neurotrophin signalling pathway; Insulin signalling pathway; Estrogen signalling pathway; Prolactin signalling pathway; Relaxin signalling pathway; Growth hormone synthesis, secretion and action; Alcoholism; Bacterial invasion of epithelial cells; Glioma; Chronic myeloid leukemia; Breast cancer; Hepatocellular carcinoma; Gastric cancer; Sphingolipid signalling pathway; Neuroactive ligand-receptor interaction; Primary bile acid biosynthesis; Taurine and hypotaurine metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Peroxisome; Bile secretion
		25	*CALN1; GALNT17; AUTS2*	Polycomb repressive complex

5.3.4 Signatures of selection (XP-EHH) between ecotype populations with different coat colours and coat colour patterns.

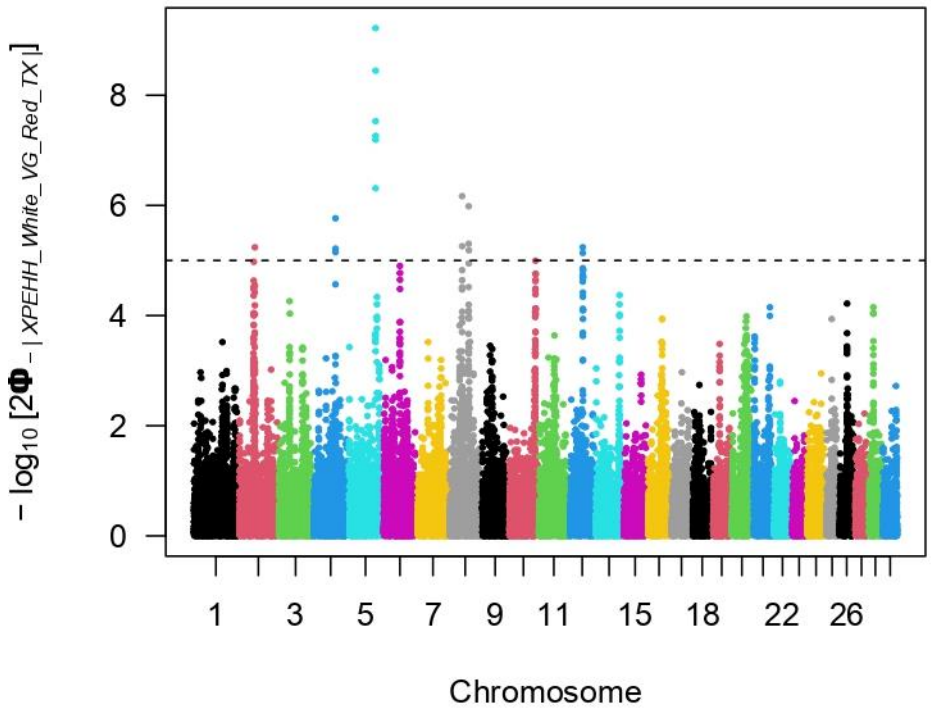

Figure 5.23: Selection signatures detected by an XP-EHH analysis between white and red goats within the TX cluster. The threshold for the detection of signatures was set to 5.

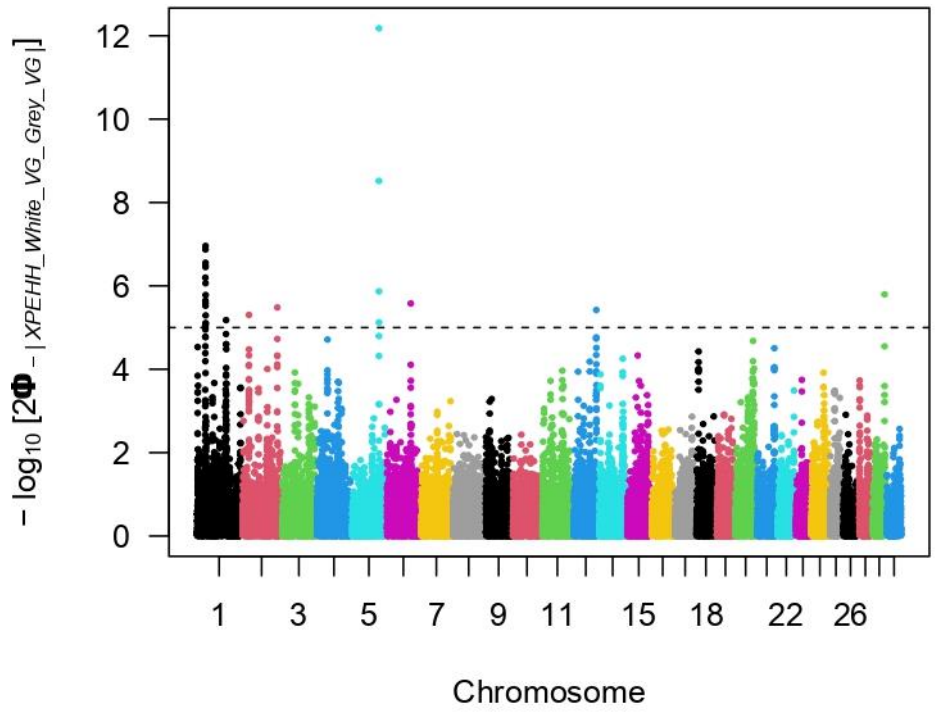

Figure 5.24: Selection signatures detected by an XP-EHH analysis between white and red goats within the ZV cluster. The threshold for the detection of signatures was set to 5.

Between white and red village goats, significant signatures were found on chromosome 4 and 5 in the white goats and on chromosome 2, 8, 4, 16, 20, and 25 in the red goats. Markers on chromosome 4 in white goats were linked to the *MAGI2* gene which is involved in Rap1 signalling pathway (chx04015), and PI3K-Akt signalling pathway (chx04151). Markers on chromosome 5 were linked to several genes including *HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B* which is involved in numerous pathways including Endocrine resistance (chx01522), ErbB signalling pathway (chx04012), HIF-1 signalling pathway (chx04066), FoxO signalling pathway (chx04068), Cell cycle (chx04110), PI3K-Akt signalling pathway (chx04151), AGE-RAGE signalling pathway in diabetic complications (chx04933), Cushing syndrome (chx04934), Measles (chx05162), Human papillomavirus infection (chx05165), Epstein-Barr virus infection (chx05169), Pathways in cancer (chx05200), Transcriptional misregulation in cancer (chx05202), Viral carcinogenesis (chx05203), MicroRNAs in cancer (chx05206), Prostate cancer (chx05215), Chronic myeloid leukemia (chx05220), Small cell lung cancer (chx05222), Gastric cancer (chx05226); *GPR19; CREBL2; DUSP16* which is involved in the MAPK signalling pathway (chx04010); *BORCS5; MANSC1*, and *FAM234B.*

Markers on chromosome 2 in the red goats were linked to genes *ACVR2A; MBD5; EPC2; KIF5C; LYPD6B; LYPD6; NABP1*; and *MYO1B* which is involved in Motor proteins (chx04814). Markers on chromosome 4 were linked to several genes including *HOXA1; HOXA2; HOXA3; HOXA4; HOXA5; HOXA6; HOXA7; HOXA9; HOXA10; HOXA11; HOXA13* which are involved in Transcriptional misregulation in cancer (chx05202); and *EVX1*. Signals on chromosome 8 were linked to the following genes, FXN which is involved in Porphyrin metabolism (chx00860); *TJP2* which is involved in Tight junction (chx04530); *APBA1; GFRA2; DOK2; XPO7* which is involved in Nucleocytoplasmic transport (chx03013); *NPM2*; *FGF17* which is involved in MAPK signalling pathway (chx04010), Ras signalling pathway (chx04014), Rap1 signalling pathway (chx04015), Calcium signalling pathway (chx04020), PI3K-Akt signalling pathway (chx04151), Regulation of actin cytoskeleton (chx04810), Pathways in cancer (chx05200), Chemical carcinogenesis - receptor activation (chx05207), Melanoma (chx05218), Breast cancer (chx05224), Gastric cancer (chx05226); *BMP1; PHYHIP; POLR3D* which is involved in RNA polymerase (chx03020), and Cytosolic DNA-sensing pathway (chx04623); *PIWIL2; SLC39A14* which is involved in Ferroptosis (chx04216), Alzheimer disease (chx05010), Parkinson disease (chx05012); *PPP3CC* which is involved in numerous pathways MAPK signalling pathway (chx04010), Calcium signalling pathway (chx04020), cGMP-PKG signalling pathway (chx04022), Oocyte meiosis

(chx04114), Cellular senescence (chx04218), Wnt signalling pathway (chx04310), Axon guidance (chx04360), VEGF signalling pathway (chx04370), Osteoclast differentiation (chx04380), C-type lectin receptor signalling pathway (chx04625), Natural killer cell mediated cytotoxicity (chx04650), Th1 and Th2 cell differentiation (chx04658), Th17 cell differentiation (chx04659), T cell receptor signalling pathway (chx04660), B cell receptor signalling pathway (chx04662), Long-term potentiation (chx04720), Glutamatergic synapse (chx04724), Dopaminergic synapse (chx04728), Oxytocin signalling pathway (chx04921), Glucagon signalling pathway (chx04922), Renin secretion (chx04924), Alzheimer disease (chx05010), Amyotrophic lateral sclerosis (chx05014), Prion disease (chx05020), Pathways of neurodegeneration - multiple diseases (chx05022), Amphetamine addiction (chx05031), Tuberculosis (chx05152), Human cytomegalovirus infection (chx05163), Human T-cell leukemia virus 1 infection (chx05166), Kaposi sarcoma-associated herpesvirus infection (chx05167), Human immunodeficiency virus 1 infection (chx05170), PD-L1 expression and PD-1 checkpoint pathway in cancer (chx05235), and Lipid and atherosclerosis (chx05417); *SORBS3; PDLIM2; CCAR2; BIN3; EGR3* which is involved in C-type lectin receptor signalling pathway (chx04625), Hepatitis B (chx05161), Viral carcinogenesis (chx05203); and *PEBP4*. Markers on chromosome 16 were linked to the *CACNA1E*; and *ZNF648* genes. Markers on chromosome 20 were linked to the *CTNND2* gene which is involved in Wnt signalling pathway (chx04310); and the *DAP* gene which is involved in Autophagy - animal (chx04140), Pathways in cancer (chx05200), and Bladder cancer (chx05219). Markers on chromosome 25 were linked to the following genes *TMC7; TMC5; GDE1; CCP110*; *VPS35L*; and *KNOP1*.

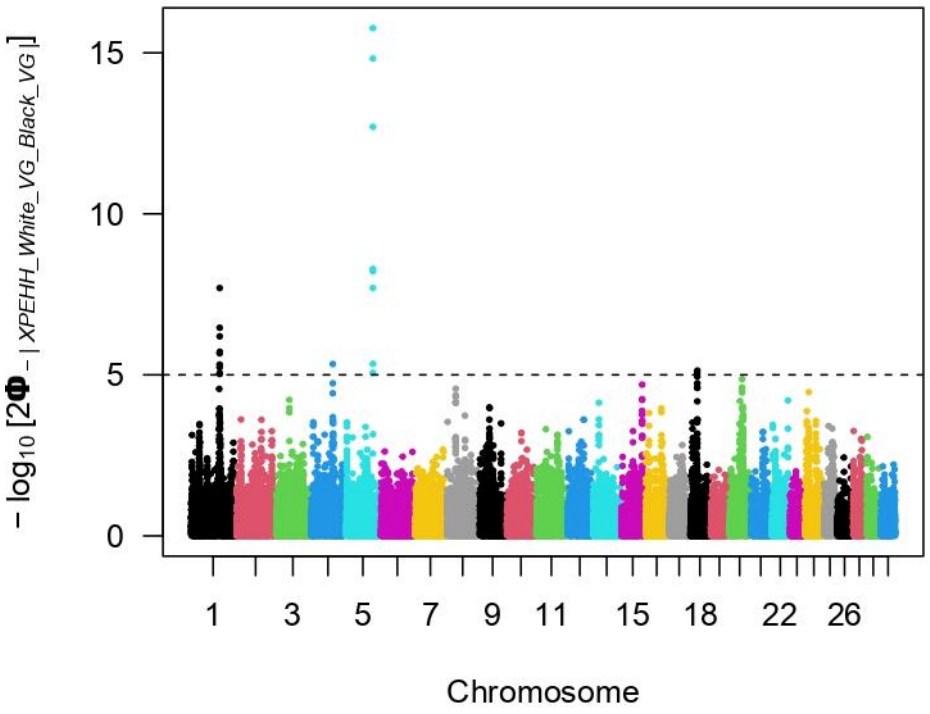

Figure 5.25: Selection signatures detected by an XP-EHH analysis between white and black goats. The threshold for the detection of signatures was set to 5

Between white and black village goats, significant markers under selection in white goats were found on chromosome 1, 4, and 5; while in black goats, significant markers were found on chromosome 18. In white goats, markers on chromosome 1 were linked to the *SI* gene. Markers on chromosome 4 were linked to the *MAGI2* gene which is involved in Rap1 signalling pathway (chx04015), PI3K-Akt signalling pathway (chx04151). Markers on chromosome 5 were linked to several gene including *GSG1; FAM234B; HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5*; and *MANSC1*. Significant markers under selection in the black goats found on chromosome 18 were linked to the *FTO; IRX3*; and *IRX5* genes.

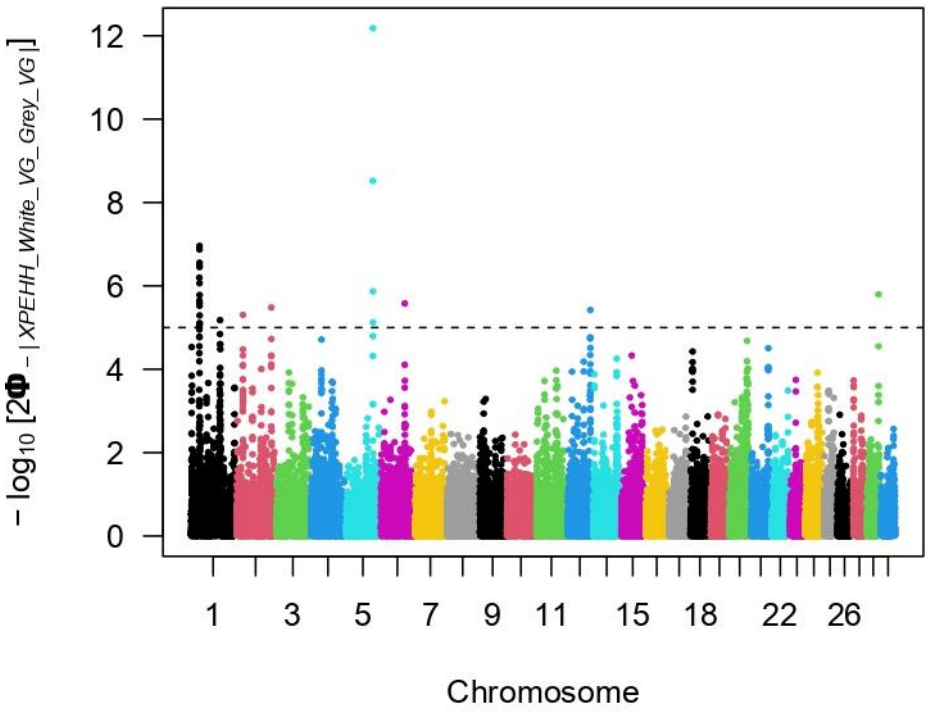

Figure 5.26: Selection signatures detected by an XP-EHH analysis between white and grey goats. The threshold for the detection of signatures was set to 5.

Between white and grey village goats, significant markers under selection in white goats were found on chromosome 5; while in grey goats, significant markers were found on chromosome 2, 6, 12, 28. Markers found in the white goats on chromosome 5 were linked to the following genes *HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5*; and *MANSC1*. Markers found in grey goats in chromosome 2 were linked to *CUL3* which is involved in Ubiquitin mediated proteolysis (chx04120), and Hedgehog signalling pathway (chx04340); and *FAM124B*. Markers on chromosome 6 were linked to the *UGT2A2* gene. Signals on chromosome 12 were linked to the *PCDH17* gene. Markers on chromosome 28 were linked to several genes including *GNPAT* which is involved in Glycerophospholipid metabolism (chx00564) and Peroxisome (chx04146); *C1orf131; TRIM67; FAM89A; ARV1; TTC13*; and *C1orf198*.

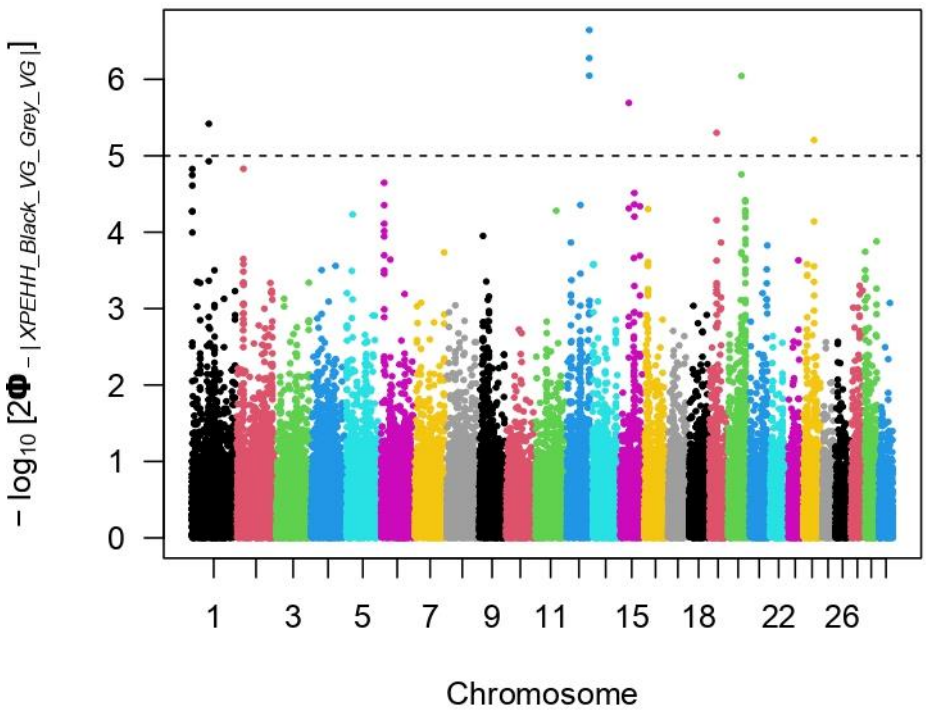

Figure 5.27: Selection signatures detected by an XP-EHH analysis between black and grey goats. The threshold for the detection of signatures was set to 5.

Between black and grey village goats, significant signals were found under selection in chromosome 20, and 24 in the black goats, and on chromosome 15, and 19. In the grey goats, markers on chromosome 15 were link to several genes including *ARAP1* which is involved in Endocytosis (chx04144); *VPS26B; THYN1; ACAD8* which is involved in Valine, leucine and isoleucine degradation (chx00280), Metabolic pathways (chx01100); *GLB1L3*. Markers on chromosomes 19 were linked to several genes including *PLD2* which is involved in Glycerophospholipid metabolism (chx00564), Ether lipid metabolism (chx00565), Metabolic pathways (chx01100), Ras signalling pathway (chx04014), cAMP signalling pathway (chx04024), Sphingolipid signalling pathway (chx04071), Phospholipase D signalling pathway (chx04072), Endocytosis (chx04144), Fc gamma R-mediated phagocytosis (chx04666), Glutamatergic synapse (chx04724), GnRH signalling pathway (chx04912), Parathyroid hormone synthesis, secretion and action (chx04928), Pathways in cancer (chx05200), Chemical carcinogenesis - reactive oxygen species (chx05208), Pancreatic cancer (chx05212), and Choline metabolism in cancer (chx05231); *GLTPD2; TM4SF5; VMO1; ZMYND15; MED11; ARRB2* which is involved in MAPK signalling pathway (chx04010), Chemokine signalling pathway (chx04062), Endocytosis (chx04144), Hedgehog signalling pathway (chx04340), Dopaminergic synapse (chx04728), Olfactory transduction (chx04740), Relaxin signalling pathway (chx04926), Parathyroid hormone synthesis, secretion and action

(chx04928), GnRH secretion (chx04929), Morphine addiction (chx05032), Chemical carcinogenesis - receptor activation (chx05207); *PELP1; ALOX15; ALOX12* which is involved in Arachidonic acid metabolism (chx00590), Metabolic pathways (chx01100), Antifolate resistance (chx01523), Serotonergic synapse (chx04726), Inflammatory mediator regulation of TRP channels (chx04750); *RNASEK; C17orf49; BCL6B; SLC16A13; SLC16A11; ASGR1* and *ASGR2* which are both involved in Thyroid hormone synthesis (chx04918); *DLG4* which is involved in several pathways including Hippo signalling pathway (chx04390), Glutamatergic synapse (chx04724), Huntington disease (chx05016), Pathways of neurodegeneration - multiple diseases (chx05022), and Cocaine addiction (chx05030); *ACADVL* which is involved in Fatty acid degradation (chx00071), Metabolic pathways (chx01100), Fatty acid metabolism (chx01212), Alcoholic liver disease (chx04936); *DVL2* which is involved in mTOR signalling pathway (chx04150), Wnt signalling pathway (chx04310), Notch signalling pathway (chx04330), Hippo signalling pathway (chx04390), Signalling pathways regulating pluripotency of stem cells (chx04550), Melanogenesis (chx04916), Cushing syndrome (chx04934), Alzheimer disease (chx05010), Pathways of neurodegeneration - multiple diseases (chx05022), Human papillomavirus infection (chx05165), Pathways in cancer (chx05200), Basal cell carcinoma (chx05217), Breast cancer (chx05224), Hepatocellular carcinoma (chx05225), and Gastric cancer (chx05226); *PHF23*; *GABARAP* which is involved in FoxO signalling pathway (chx04068), Autophagy - other (chx04136), Mitophagy - animal (chx04137), Autophagy - animal (chx04140), NOD-like receptor signalling pathway (chx04621), GABAergic synapse (chx04727); *CTDNEP1; ELP5; CLDN7* which is involved in Cell adhesion molecules (chx04514), Tight junction (chx04530), Leukocyte transendothelial migration (chx04670), Hepatitis C (chx05160); *SLC2A4; YBX2; GPS2* which is involved in Human T-cell leukemia virus 1 infection (chx05166); *NEURL4; ACAP1* which is involved in Endocytosis (chx04144).

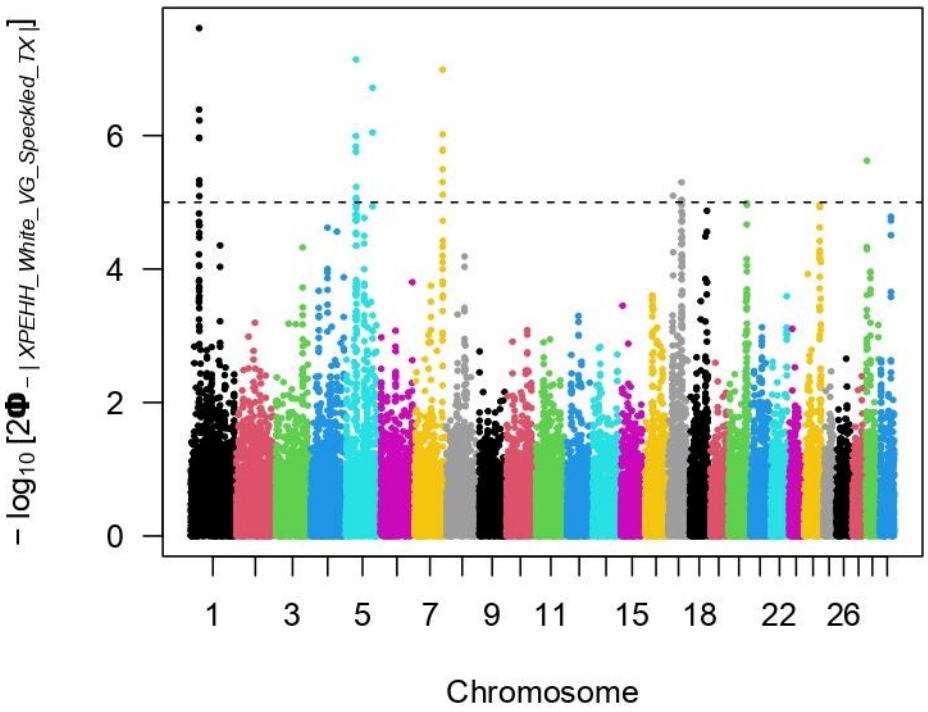

Figure 5.28: Selection signatures detected by an XP-EHH analysis between white and speckled goats within the TX cluster. The threshold for the detection of signatures was set to 5.

Between white and speckled goats, significant markers under selection were found on chromosome 5 and 7 in white goats, and on chromosome 5, 17, and 28 in speckled goats within the TX cluster. In white goats, markers found on chromosome 5 were linked to the following genes, *FAM234B; HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16*; and *BORCS5*. Markers under selection in the speckled goats on chromosome 5 were linked to the following genes *TMEM117; TWF1; IRAK4* which is involved in MAPK signalling pathway (chx04010), NF-kappa B signalling pathway (chx04064), Toll-like receptor signalling pathway (chx04620), NOD-like receptor signalling pathway (chx04621), Neurotrophin signalling pathway (chx04722), Alcoholic liver disease (chx04936), Salmonella infection (chx05132), Pertussis (chx05133), Yersinia infection (chx05135), Leishmaniasis (chx05140), Chagas disease (chx05142), Toxoplasmosis (chx05145), Tuberculosis (chx05152), Hepatitis B (chx05161), Measles (chx05162), Influenza A (chx05164), Herpes simplex virus 1 infection (chx05168), Epstein-Barr virus infection (chx05169), Human immunodeficiency virus 1 infection (chx05170), Coronavirus disease - COVID-19 (chx05171), Lipid and atherosclerosis (chx05417); *PUS7L; ADAMTS20; PRICKLE1* which is involved in the Wnt signalling pathway (chx04310); and *PPHLN1*. Significant markers on chromosome 17 were linked to the following genes, *CUX2; MYL2* which is involved in numerous pathways including Cardiac muscle contraction (chx04260),

Adrenergic signalling in cardiomyocytes (chx04261), Apelin signalling pathway (chx04371), Focal adhesion (chx04510), Adherens junction (chx04520), Tight junction (chx04530), Leukocyte transendothelial migration (chx04670), Regulation of actin cytoskeleton (chx04810), Motor proteins (chx04814), Salmonella infection (chx05132), Hypertrophic cardiomyopathy (chx05410), and Dilated cardiomyopathy (chx05414); *CCDC63; PPP1CC* which is involved in mRNA surveillance pathway (chx03015), cGMP-PKG signalling pathway (chx04022), cAMP signalling pathway (chx04024), Oocyte meiosis (chx04114), Cellular senescence (chx04218), Adrenergic signalling in cardiomyocytes (chx04261), Vascular smooth muscle contraction (chx04270), Hippo signalling pathway (chx04390), Focal adhesion (chx04510), Platelet activation (chx04611), Long-term potentiation (chx04720), Dopaminergic synapse (chx04728), Inflammatory mediator regulation of TRP channels (chx04750), Regulation of actin cytoskeleton (chx04810), Insulin signalling pathway (chx04910), Oxytocin signalling pathway (chx04921), Insulin resistance (chx04931), Amphetamine addiction (chx05031), Alcoholism (chx05034), Herpes simplex virus 1 infection (chx05168), Proteoglycans in cancer (chx05205), Diabetic cardiomyopathy (chx05415); and *PABPC4L* which is involved in mRNA surveillance pathway (chx03015), RNA degradation (chx03018). While markers on chromosome 28 were linked to the genes *VSTM4; LRRC18*; and *ARHGAP22*.

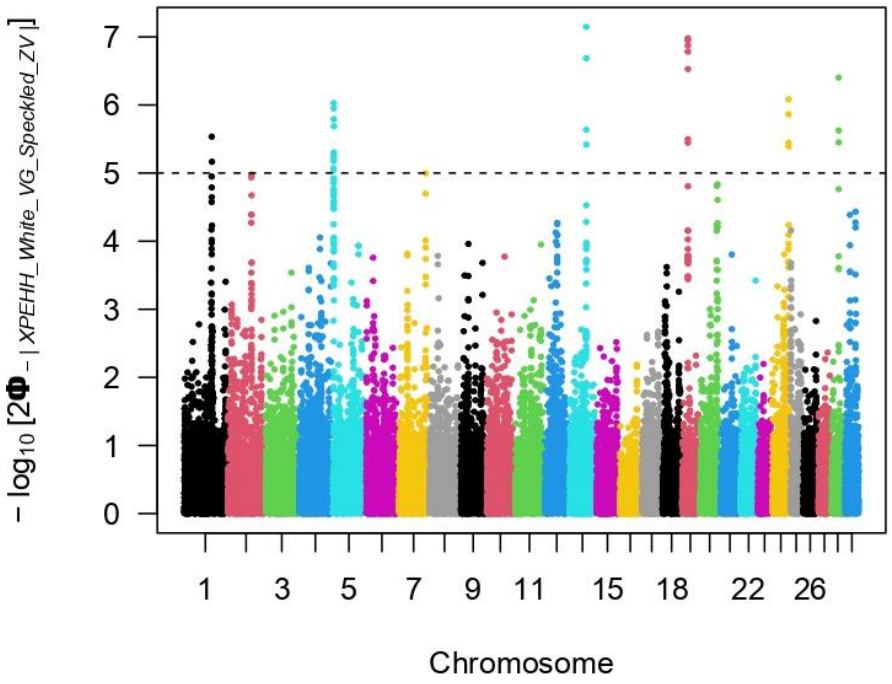

Figure 5.29: Selection signatures detected by an XP-EHH analysis between white and speckled goats within the ZV cluster. The threshold for the detection of signatures was set to 5.

Between white goats and speckled goats within the ZV cluster, significant markers under selection were found on chromosome 19, and 28 in the white goats, and on chromosome 5, 14, and 24 in speckled goats. Markers on chromosome 19 in white goats were linked to several genes including *CRK* which is involved in MAPK signalling pathway (chx04010), ErbB signalling pathway (chx04012), Rap1 signalling pathway (chx04015), Chemokine signalling pathway (chx04062), Focal adhesion (chx04510), Fc gamma R-mediated phagocytosis (chx04666), Neurotrophin signalling pathway (chx04722), Regulation of actin cytoskeleton (chx04810), Insulin signalling pathway (chx04910), Growth hormone synthesis, secretion and action (chx04935), Bacterial invasion of epithelial cells (chx05100), Yersinia infection (chx05135), Human cytomegalovirus infection (chx05163), Human immunodeficiency virus 1 infection (chx05170), Pathways in cancer (chx05200), MicroRNAs in cancer (chx05206), Renal cell carcinoma (chx05211), Chronic myeloid leukemia (chx05220); *MYO1C* which is in Motor proteins (chx04814); *INPP5K* which is in involved in Inositol phosphate metabolism (chx00562), Metabolic pathways (chx01100); *PITPNA; SLC43A2; SCARF1; RILP* which is involved in Phagosome (chx04145), Salmonella infection (chx05132); *PRPF8* which is involved in Spliceosome (chx03040); *TLCD2; WDR81; SERPINF2* which is involved in Complement and coagulation cascades (chx04610); *SERPINF1* which is in involved in Wnt signalling pathway (chx04310); *SMYD4; RPA1* which is involved in RNA polymerase (chx03020); *RTN4RL1*; and *DPH1*. Signals on chromosome 28 were linked to the two genes, *REEP3*; and *JMJD1C* which is involved in Transcriptional misregulation in cancer (chx05202). Markers under selection on chromosome 5 in speckled goats within the ZV cluster were linked to the *TRHDE*; and *ATXN7L3B* genes. Markers on chromosome 14 were linked to the *SNTG1*; *PPDPFL*; and *SNAI2* gene which is in involved in Hippo signalling pathway (chx04390), Adherens junction (chx04520). Signals on chromosome 24 were linked to the *RNF152* which is involved in mTOR signalling pathway (chx04150); *PIGN* which is involved in Glycosylphosphatidylinositol (GPI)-anchor biosynthesis (chx00563), Metabolic pathways (chx01100); and *RELCH* genes.

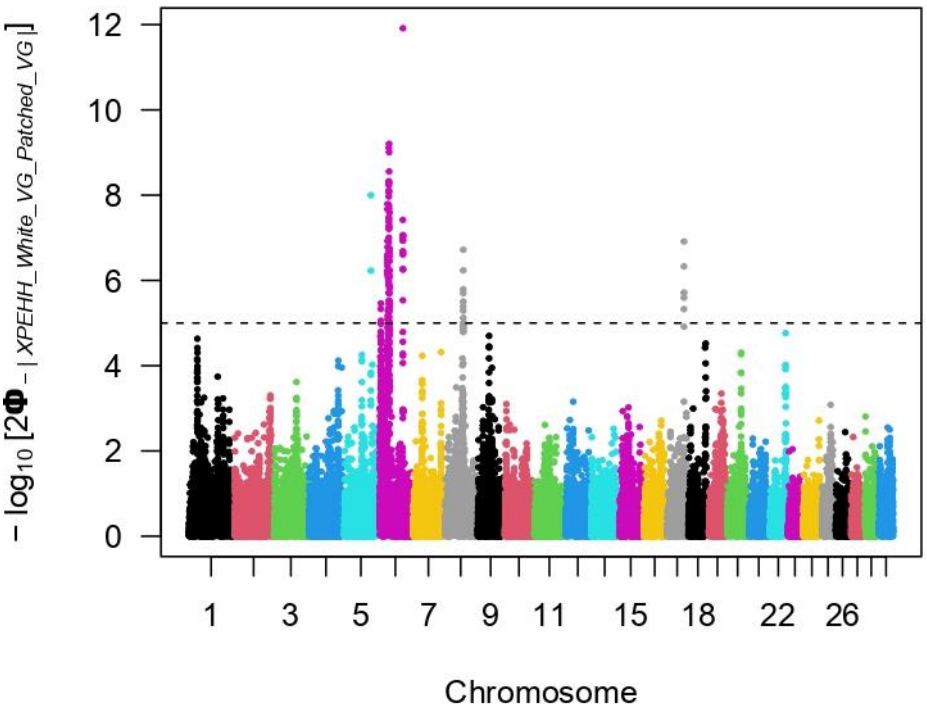

Figure 5.30: Selection signatures detected by an XP-EHH analysis between white and patchy goats. The threshold for the detection of signatures was set to 5.

Between white and patchy village goats, significant markers under selection in white goats but not in patchy goats were detected on chromosome 5, 6, and 17. Markers on chromosome 5 were linked to the following genes, *HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5*. Markers on chromosome 6 were linked to the following genes, *SULT1B1; SULT1E1; ODAM*. Markers on chromosome 17 were linked to the *EDNRA* gene which is in involved in Calcium signalling pathway (chx04020), cGMP-PKG signalling pathway (chx04022), cAMP signalling pathway (chx04024), Neuroactive ligand-receptor interaction (chx04080), Vascular smooth muscle contraction (chx04270), Renin secretion (chx04924), Pathways in cancer (chx05200); and the *PRMT9* gene. Significant markers under selection in patchy village goats were found on chromosome 6. These markers were linked to several genes including *MYOZ2; SYNPO2; SEC24D* which is involved in Protein processing in endoplasmic reticulum (chx04141); *METTL14; PRSS12; NDST3* which is involved in Glycosaminoglycan biosynthesis - heparan sulfate / heparin (chx00534), Metabolic pathways (chx01100); *HPGDS* which is involved in Glutathione metabolism (chx00480), Arachidonic acid metabolism (chx00590), Metabolism of xenobiotics by cytochrome P450 (chx00980), Drug metabolism - cytochrome P450 (chx00982), Metabolic pathways (chx01100), and Chemical carcinogenesis - DNA adducts (chx05204); *SMARCAD1* which is involved in Signalling pathways regulating pluripotency of stem cells (chx04550); *ATOH1; GRID2* which is involved in Neuroactive ligand-receptor interaction (chx04080), and

Long-term depression (chx04730); *CCSER1; MMRN1; SNCA* which is involved in Alzheimer disease (chx05010), Parkinson disease (chx05012), Pathways of neurodegeneration - multiple diseases (chx05022); *GPRIN3; TIGD2; FAM13A; HERC3* which is involved in Ubiquitin mediated proteolysis (chx04120); *NAP1L5; HERC5; HERC6; PPM1K; ABCG2* which is involved in Antifolate resistance (chx01523), ABC transporters (chx02010), Bile secretion (chx04976); *PKD2; SPP1* which is involved in PI3K-Akt signalling pathway (chx04151), Apelin signalling pathway (chx04371), Focal adhesion (chx04510), ECM-receptor interaction (chx04512), Toll-like receptor signalling pathway (chx04620), GnRH secretion (chx04929), Human papillomavirus infection (chx05165); *NCAPG*; and *LCORL*.

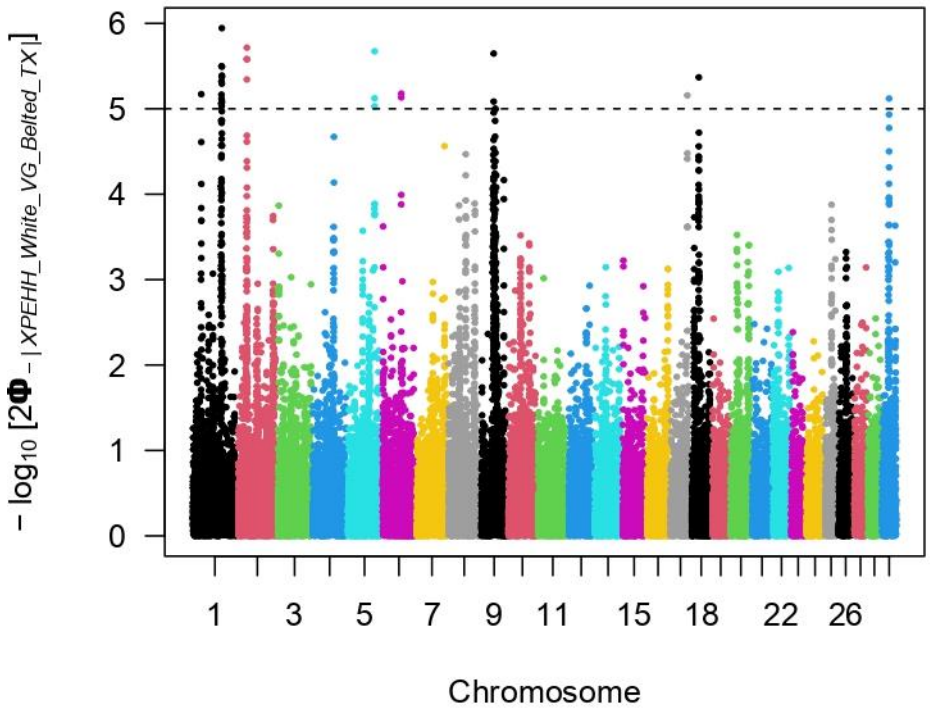

Figure 5.31: Selection signatures detected by an XP-EHH analysis between white and belted goats within the TX cluster. The threshold for the detection of signatures was set to 5.

Between white and belted village goats, candidate signatures under selection in white goats but not in belted goats in the TX cluster were found on chromosome 1, 5, 6, and 17. Markers on chromosome 1 were linked to the *SLITRK3*; and *SI* genes. The *SLITRK3* gene is linked to Cell adhesion molecules (chx04514). Markers on chromosome 5 is linked to several genes including *HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5*; and *MANSC1*. Markers on chromosome 6 were linked to the following genes, *ATP10D; CORIN; NFXL1; CNGA1* which is involved in cGMP-PKG signalling pathway (chx04022), cAMP signalling pathway (chx04024), Phototransduction (chx04744); *NIPAL1*; and *TXK* which is involved in Leukocyte transendothelial migration (chx04670). Markers on

chromosome 17 were linked to *EDNRA* which is involved in Calcium signalling pathway (chx04020), cGMP-PKG signalling pathway (chx04022), cAMP signalling pathway (chx04024), Neuroactive ligand-receptor interaction (chx04080), Vascular smooth muscle contraction (chx04270), Renin secretion (chx04924), and Pathways in cancer (chx05200); and *PRMT9*.

Significant candidate signatures under selection in the belted goats in the TX cluster but not in white goats were found on chromosome 2, 9, 18, and 20. Markers on chromosome 2 were linked to *TNP1*; and *IGFBP5*. Markers on chromosome 9 were linked to *EPHA7* which is involved in Axon guidance (chx04360); *CNR1* which is involved in Rap1 signalling pathway (chx04015), Neuroactive ligand-receptor interaction (chx04080), Thermogenesis (chx04714), Retrograde endocannabinoid signalling (chx04723); *SPACA1*; and *AKIRIN2*. Markers on chromosome 18 were linked to several genes including *IRX6*; *MMP2* which is involved in Endocrine resistance (chx01522), Leukocyte transendothelial migration (chx04670), GnRH signalling pathway (chx04912), Estrogen signalling pathway (chx04915), Relaxin signalling pathway (chx04926), AGE-RAGE signalling pathway in diabetic complications (chx04933), Pathways in cancer (chx05200), Proteoglycans in cancer (chx05205), Bladder cancer (chx05219), Diabetic cardiomyopathy (chx05415), Fluid shear stress and atherosclerosis (chx05418); *LPCAT2* which is involved in Glycerophospholipid metabolism (chx00564), Ether lipid metabolism (chx00565), Metabolic pathways (chx01100); and *SLC6A2* which is involved in Synaptic vesicle cycle (chx04721). Markers on chromosome 29 were linked to several genes including *E2F8; CSRP3; ZDHHC13; PTPN5* which is involved in MAPK signalling pathway (chx04010); *TMEM86A; SPTY2D1*; and *MISFA*.

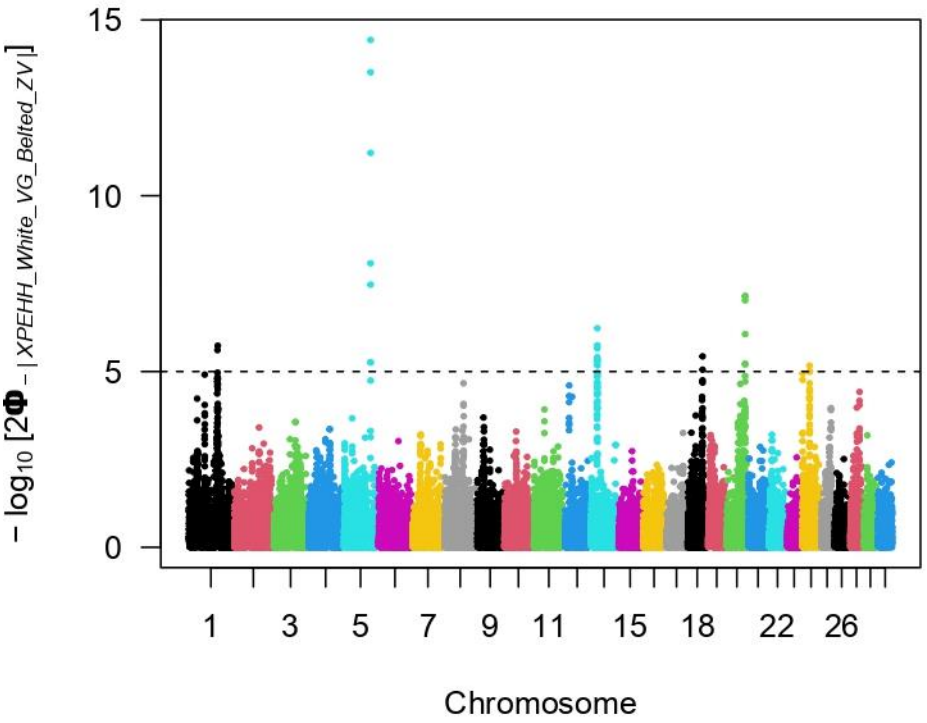

Figure 5.32: Selection signatures detected by an XP-EHH analysis between white and belted goats within the ZV cluster. The threshold for the detection of signatures was set to 5.

Candidate signatures under selection in white goats but not in belted goats in the ZV cluster were found on chromosome 5, and 24. Signatures on chromosome 5 were linked to several genes including *GSG1*; *FAM234B*; *HEBP1*; *GPRC5D*; *GPRC5A*; *DDX47; APOLD1*; *CDKN1B*; *GPR19*; *CREBL2*; *DUSP16*; *BORCS5*; and *MANSC1*. Markers on chromosome 24 were linked to the following genes, *B4GALT6* which is involved in Purine metabolism (chx00230), Metabolic pathways (chx01100), Endocrine resistance (chx01522), Rap1 signalling pathway (chx04015), Calcium signalling pathway (chx04020), cGMP-PKG signalling pathway (chx04022), cAMP signalling pathway (chx04024), Chemokine signalling pathway (chx04062), Phospholipase D signalling pathway (chx04072), Oocyte meiosis (chx04114), Longevity regulating pathway (chx04211), Longevity regulating pathway - multiple species (chx04213), Adrenergic signalling in cardiomyocytes (chx04261), Vascular smooth muscle contraction (chx04270), Apelin signalling pathway (chx04371), Gap junction (chx04540), Platelet activation (chx04611), Circadian entrainment (chx04713), Thermogenesis (chx04714), Retrograde endocannabinoid signalling (chx04723), Glutamatergic synapse (chx04724), Cholinergic synapse (chx04725), GABAergic synapse (chx04727), Inflammatory mediator regulation of TRP channels (chx04750), Insulin secretion (chx04911), GnRH signalling pathway (chx04912), Ovarian steroidogenesis (chx04913), Progesterone-mediated oocyte maturation (chx04914), Estrogen signalling pathway (chx04915), Melanogenesis

(chx04916), Thyroid hormone synthesis (chx04918), Oxytocin signalling pathway (chx04921), Glucagon signalling pathway (chx04922), Regulation of lipolysis in adipocytes (chx04923), Aldosterone synthesis and secretion (chx04925), Relaxin signalling pathway (chx04926), Cortisol synthesis and secretion (chx04927), Parathyroid hormone synthesis, secretion and action (chx04928), Cushing syndrome (chx04934), Growth hormone synthesis, secretion and action (chx04935), Salivary secretion (chx04970), Gastric acid secretion (chx04971), Pancreatic secretion (chx04972), Bile secretion (chx04976), Morphine addiction (chx05032), Human cytomegalovirus infection (chx05163), Human T-cell leukemia virus 1 infection (chx05166), Pathways in cancer (chx05200), Chemical carcinogenesis - receptor activation (chx05207), Dilated cardiomyopathy (chx05414); *TTR* which is involved in Thyroid hormone synthesis (chx04918); *DSG2* which is involved in Arrhythmogenic right ventricular cardiomyopathy (chx05412); *DSG3; DSG4; DSG1* which is involved in Staphylococcus aureus infection (chx05150); and *DSC1*.

Significant candidate signatures under selection in the belted goats in the ZV cluster but not in white goats were found on chromosome 14, 18, and 20. Makers on chromosome 14 were linked to the *ZFPM2* gene which is involved in MicroRNAs in cancer (chx05206). Markers on chromosome 18 were linked to several genes including *PHLDB3; ETHE1* which is involved in Sulfur metabolism (chx00920), and Metabolic pathways (chx01100); *ZNF575; XRCC1* which is involved in Base excision repair (chx03410); *IRGQ; ZNF576; ZNF428; CADM4; PLAUR* which is involved in Complement and coagulation cascades (chx04610), and Proteoglycans in cancer (chx05205); *IRGC; SMG9; KCNN4* which is involved in Insulin secretion (chx04911), GnRH secretion (chx04929), Salivary secretion (chx04970), Protein digestion and absorption (chx04974); *LYPD5; ZNF404* which is involved in Herpes simplex virus 1 infection (chx05168); *ZNF45; ZNF234*; and *ZNF226*. Candidate markers on chromosome 20 were linked to the *ADCY2* gene which is involved in Purine metabolism (chx00230), Metabolic pathways (chx01100), Endocrine resistance (chx01522), Rap1 signalling pathway (chx04015), Calcium signalling pathway (chx04020), cGMP-PKG signalling pathway (chx04022), cAMP signalling pathway (chx04024), Chemokine signalling pathway (chx04062), Phospholipase D signalling pathway (chx04072), Oocyte meiosis (chx04114), Longevity regulating pathway (chx04211), Longevity regulating pathway - multiple species (chx04213), Adrenergic signalling in cardiomyocytes (chx04261), Vascular smooth muscle contraction (chx04270), Apelin signalling pathway (chx04371), Gap junction (chx04540), Platelet activation (chx04611), Circadian entrainment (chx04713), Thermogenesis (chx04714), Retrograde

endocannabinoid signalling (chx04723), Glutamatergic synapse (chx04724), Cholinergic synapse (chx04725), GABAergic synapse (chx04727), Inflammatory mediator regulation of TRP channels (chx04750), Insulin secretion (chx04911), GnRH signalling pathway (chx04912), Ovarian steroidogenesis (chx04913), Progesterone-mediated oocyte maturation (chx04914), Estrogen signalling pathway (chx04915), Melanogenesis (chx04916), Thyroid hormone synthesis (chx04918), Oxytocin signalling pathway (chx04921), Glucagon signalling pathway (chx04922), Regulation of lipolysis in adipocytes (chx04923), Aldosterone synthesis and secretion (chx04925), Relaxin signalling pathway (chx04926), Cortisol synthesis and secretion (chx04927), Parathyroid hormone synthesis, secretion and action (chx04928), Cushing syndrome (chx04934), Growth hormone synthesis, secretion and action (chx04935), Salivary secretion (chx04970), Gastric acid secretion (chx04971), Pancreatic secretion (chx04972), Bile secretion (chx04976), Morphine addiction (chx05032), Human cytomegalovirus infection (chx05163), Human T-cell leukemia virus 1 infection (chx05166), Pathways in cancer (chx05200), Chemical carcinogenesis - receptor activation (chx05207), and Dilated cardiomyopathy (chx05414).

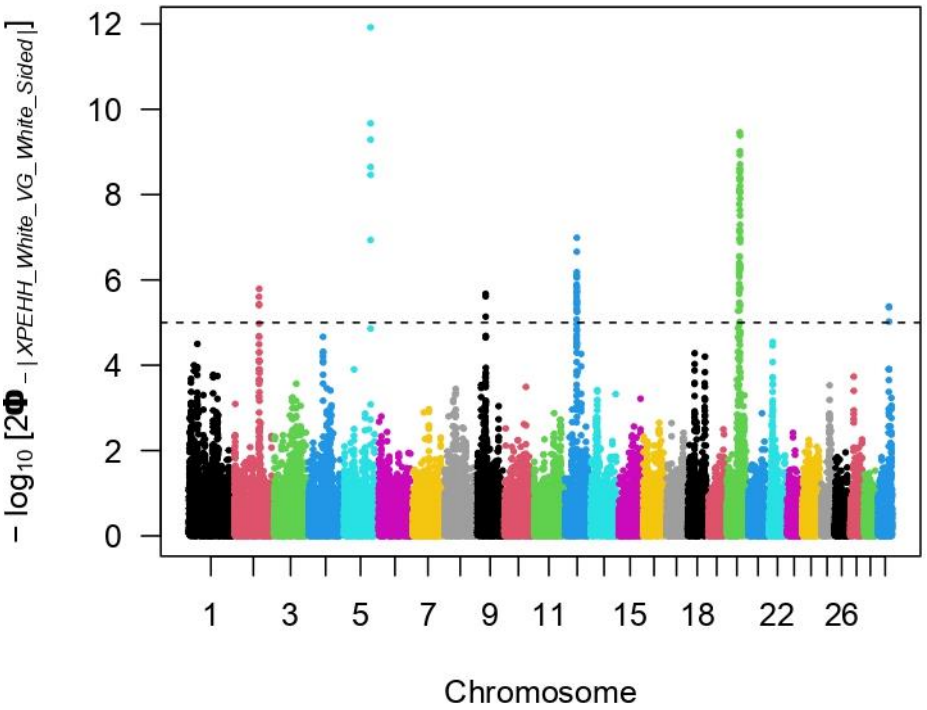

Figure 5.33: Selection signatures detected by an XP-EHH analysis between white and white-sided goats. The threshold for the detection of signatures was set to 5.

Between white and white-sided village goats, candidate signatures of selection were found on chromosome 5, and 29 in white goats, and on chromosome 2, 9, 12, and 20 in white-sided goats. Markers on chromosome 5 in white goats were linked to several genes including *GSG1; FAM234B; HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2;*

DUSP16; and *BORCS5*. Markers on chromosome 29 were linked to the following genes *PAG12; PAG-4*; and *PAG-11* which are involved in Protein digestion and absorption (chx04974). In white-side goats significant markers found on chromosome 2 were linked to the following genes, *MBD5* which is involved in Polycomb repressive complex (chx03083); *EPC2* which is involved in ATP-dependent chromatin remodeling (chx03082); *KIF5C* which is involved in Endocytosis (chx04144), Dopaminergic synapse (chx04728), Motor proteins (chx04814), Alzheimer disease (chx05010), Parkinson disease (chx05012), Amyotrophic lateral sclerosis (chx05014), Huntington disease (chx05016), Prion disease (chx05020), Pathways of neurodegeneration - multiple diseases (chx05022), Salmonella infection (chx05132), and Non-small cell lung cancer (chx05223); *LYPD6B*; and *LYPD6* which are involved in Neuroactive ligand-receptor interaction (chx04080). Markers on chromosome 9 were linked to four genes including *SEC63* which is involved in Protein export (chx03060), and Protein processing in endoplasmic reticulum (chx04141); *SCML4*; *SOBP*; and *PDSS2* which is involved in Terpenoid backbone biosynthesis (chx00900). On chromosome 12 the signal was linked to the *KLHL1* gene. Markers on chromosome 20 were linked to the *NPR3* and *CDH9* genes.

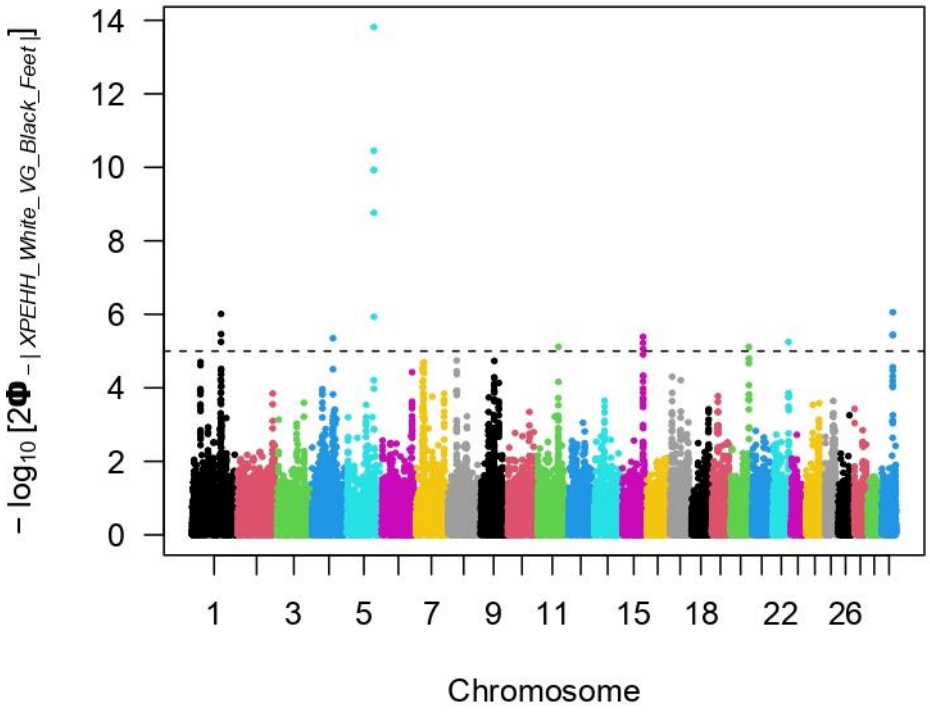

Figure 5.34: Selection signatures detected by an XP-EHH analysis between white and blacklegged goats. The threshold for the detection of signatures was set to 5.

Between white and black-legged village goats, candidate signatures of selection were found on chromosome 4, 5, 22 and 29 in white goats, and on chromosome 11, 15, and 20 on black-legged goats. In white goats, markers on chromosome 4 were linked to *ORC5*; and *RELN*. Markers on 5 were linked to several genes including *GSG1*; *FAM234B*; *HEBP1*; *GPRC5D*; *GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16*; and *BORCS5*. Markers on chromosome 22 were linked to the following genes *PLXND1; H1-8; RHO; IFT122; MBD4* which is involved in Base excision repair (chx03410); *EFCAB12; RPL32* which is involved in Ribosome (chx03010), Coronavirus disease - COVID-19 (chx05171); *CAND2*; and *RAF1* which is involved in numerous pathways including EGFR tyrosine kinase inhibitor resistance (chx01521), Endocrine resistance (chx01522), MAPK signalling pathway (chx04010), Melanogenesis (chx04916), Melanoma (chx05218), Bladder cancer (chx05219), Chronic myeloid leukemia (chx05220), Acute myeloid leukemia (chx05221), Non-small cell lung cancer (chx05223), Breast cancer (chx05224), Hepatocellular carcinoma (chx05225), Gastric cancer (chx05226), Central carbon metabolism in cancer (chx05230), Choline metabolism in cancer (chx05231), and PD-L1 expression and PD-1 checkpoint pathway in cancer (chx05235). Markers on chromosome 29 were linked to the *PAG12*; *PAG-4*; and *PAG-11* genes which are involved in Protein digestion and absorption (chx04974). In blacklegged goats, significant markers on chromosome 11 were linked to the *UBXN2A* which is involved in Protein processing in endoplasmic reticulum (chx04141); *ATAD2B*; and *KLHL29* genes. Markers on chromosome 15 were linked to the *ARHGAP42*; *PGR* which is involved in Oocyte meiosis (chx04114), Progesterone-mediated oocyte maturation (chx04914), Estrogen signalling pathway (chx04915), Chemical carcinogenesis - receptor activation (chx05207), and Breast cancer (chx05224); and *TRPC6* which is involved in cGMP-PKG signalling pathway (chx04022), and Axon guidance (chx04360). Markers on chromosome were linked to the *ADCY2* gene which play roles in numerous pathways including Thermogenesis (chx04714), Estrogen signalling pathway (chx04915), Thyroid hormone synthesis (chx04918), and Melanogenesis (chx04916).

Table 5.5: Genes under selection between village meat-type goats with different coat colours and coat colour patterns and their associated pathways

Breed		Chr	Genes	Pathways
White vs Red-TX	White	4	*MAGI2*	Rap1 signalling pathway; PI3K-Akt signalling pathway
		5	*HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5; MANSC1*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer; MAPK signalling pathway;
	Red-TX	2	*NABP1; MYO1B*	Motor proteins
		8	*FXN; TJP2; APBA1; GFRA2; DOK2; XPO7; NPM2; FGF17; BMP1; PHYHIP; POLR3D; PIWIL2; SLC39A14; PPP3CC; SORBS3; PDLIM2; CCAR2; BIN3; EGR3; PEBP4*	Porphyrin metabolism; Tight junction; Nucleocytoplasmic transport; MAPK signalling pathway; Ras signalling pathway; Rap1 signalling pathway; Calcium signalling pathway; PI3K-Akt signalling pathway; Regulation of actin cytoskeleton; Pathways in cancer; Chemical carcinogenesis - receptor activation; Melanoma; Breast cancer; Gastric cancer; RNA polymerase; Cytosolic DNA-sensing pathway; Ferroptosis; Alzheimer disease; Parkinson disease; MAPK signalling pathway; Calcium signalling pathway; cGMP-PKG signalling pathway; Oocyte meiosis; Cellular senescence; Wnt signalling pathway; Axon guidance; VEGF signalling pathway; Osteoclast differentiation; C-type lectin receptor signalling pathway; Natural killer cell mediated cytotoxicity; Th1 and Th2 cell differentiation; Th17 cell differentiation; T cell receptor signalling pathway; B cell receptor signalling pathway; Long-term potentiation; Glutamatergic synapse; Dopaminergic synapse; Oxytocin signalling pathway; Glucagon signalling pathway; Renin secretion; Alzheimer disease; Amyotrophic lateral sclerosis; Prion disease; Pathways of neurodegeneration - multiple diseases; Amphetamine addiction; Tuberculosis; Human cytomegalovirus infection; Human T-cell leukemia virus 1 infection; Kaposi sarcoma-associated herpesvirus infection; Human immunodeficiency virus 1 infection; PD-L1 expression and PD-1 checkpoint pathway in cancer; Lipid and atherosclerosis; C-type lectin receptor signalling pathway; Hepatitis B; Viral carcinogenesis
White vs Red-ZV	White	5	*FAM234B; HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer; MAPK signalling pathway;
	Red-ZV	2	*ACVR2A; MBD5; EPC2; KIF5C; LYPD6B; LYPD6*	Cytokine-cytokine receptor interaction; TGF-beta signalling pathway; Signalling pathways regulating pluripotency of stem cells; Fluid shear stress and atherosclerosis; Polycomb repressive complex; ATP-dependent chromatin remodeling; Endocytosis; Dopaminergic synapse; Motor proteins; Alzheimer disease; Parkinson disease; Amyotrophic lateral sclerosis; Huntington disease; Prion disease; Pathways of neurodegeneration - multiple diseases;

				Salmonella infection; Non-small cell lung cancer; Neuroactive ligand-receptor interaction; Neuroactive ligand-receptor interaction
		4	*HOXA1; HOXA2; HOXA3; HOXA4; HOXA5; HOXA6; HOXA7; HOXA9; HOXA10; HOXA11; HOXA13; EVX1*	Transcriptional misregulation in cancer;
		16	*CACNA1E; ZNF648*	
		20	*CTNND2; DAP*	Wnt signalling pathway
		25	*TMC7; TMC5; GDE1; CCP110; VPS35L; KNOP1*	
White vs Black	White	1	*SI*	
		4	*MAGI2*	Rap1 signalling pathway; PI3K-Akt signalling pathway
		5	*GSG1; FAM234B; HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5; MANSC1*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complication; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer; MAPK signalling pathway;
	Black	18	*FTO; IRX3; IRX5*	
White vs Grey	White	5	*HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5; MANSC1*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer; MAPK signalling pathway;
	Grey	2	*CUL3; FAM124B*	Ubiquitin mediated proteolysis; Hedgehog signalling pathway
		6	*UGT2A2*	
		12	*PCDH17*	
		28	*GNPAT; C1orf131; TRIM67; FAM89A; ARV1; TTC13; C1orf198*	Glycerophospholipid metabolism; Peroxisome
Black vs Grey	Black	20		
		24		

	Grey	15	*ARAP1; VPS26B; THYN1; ACAD8; GLB1L3*	Endocytosis; Valine, leucine and isoleucine degradation; Metabolic pathways;
		19	*PLD2; GLTPD2; TM4SF5; VMO1; ZMYND15; MED11; ARRB2; PELP1; ALOX15; ALOX12; RNASEK; C17orf49; BCL6B; SLC16A13; SLC16A11; ASGR2; ASGR1; DLG4; ACADVL; DVL2; PHF23; GABARAP; CTDNEP1; ELP5; CLDN7; SLC2A4; YBX2; GPS2; NEURL4; ACAP1*	Glycerophospholipid metabolism; Ether lipid metabolism; Metabolic pathways; Ras signalling pathway; cAMP signalling pathway; Sphingolipid signalling pathway; Phospholipase D signalling pathway; Endocytosis; Fc gamma R-mediated phagocytosis; Glutamatergic synapse; GnRH signalling pathway; Parathyroid hormone synthesis, secretion and action; Pathways in cancer; Chemical carcinogenesis - reactive oxygen species; Pancreatic cancer; Choline metabolism in cancer; MAPK signalling pathway; Chemokine signalling pathway; Endocytosis; Hedgehog signalling pathway; Dopaminergic synapse; Olfactory transduction; Relaxin signalling pathway; Parathyroid hormone synthesis, secretion and action; GnRH secretion; Morphine addiction; Chemical carcinogenesis - receptor activation; Arachidonic acid metabolism; Metabolic pathways; Antifolate resistance; Serotonergic synapse; Inflammatory mediator regulation of TRP channels; Thyroid hormone synthesis; Hippo signalling pathway; Glutamatergic synapse; Huntington disease; Pathways of neurodegeneration - multiple diseases; Cocaine addiction; Fatty acid degradation; Metabolic pathways; Fatty acid metabolism; Alcoholic liver disease; mTOR signalling pathway; Wnt signalling pathway; Notch signalling pathway; Hippo signalling pathway; Signalling pathways regulating pluripotency of stem cells; Melanogenesis; Cushing syndrome; Alzheimer disease; Pathways of neurodegeneration - multiple diseases; Human papillomavirus infection; Pathways in cancer; Basal cell carcinoma; Breast cancer; Hepatocellular carcinoma; Gastric cancer; FoxO signalling pathway; Autophagy – other; Mitophagy – animal; Autophagy – animal; NOD-like receptor signalling pathway; GABAergic synapse; Cell adhesion molecules; Tight junction; Leukocyte transendothelial migration; Hepatitis C; Human T-cell leukemia virus 1 infection; Endocytosis
White vs Speckled-TX	White	5	*FAM234B; HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer; MAPK signalling pathway;
		7		
	Speckled-TX	5	*TMEM117; TWF1; IRAK4; PUS7L; ADAMTS20; PRICKLE1; PPHLN1*	MAPK signalling pathway; NF-kappa B signalling pathway; Toll-like receptor signalling pathway; NOD-like receptor signalling pathway; Neurotrophin signalling pathway; Alcoholic liver disease; Salmonella infection; Pertussis; Yersinia infection; Leishmaniasis; Chagas disease; Toxoplasmosis; Tuberculosis; Hepatitis B; Measles; Influenza A; Herpes simplex virus 1 infection; Epstein-Barr virus infection; Human immunodeficiency virus 1 infection; Coronavirus disease - COVID-19; Lipid and atherosclerosis; Wnt signalling pathway;
		17	*CUX2; MYL2; CCDC63; PPP1CC; PABPC4L*	Cardiac muscle contraction; Adrenergic signalling in cardiomyocytes; Apelin signalling pathway; Focal adhesion; Adherens junction; Tight junction; Leukocyte transendothelial migration; Regulation of actin cytoskeleton; Motor proteins; Salmonella infection; Hypertrophic cardiomyopathy; Dilated cardiomyopathy; mRNA surveillance pathway; cGMP-PKG signalling pathway; cAMP signalling pathway; Oocyte meiosis; Cellular senescence; Adrenergic signalling in cardiomyocytes; Vascular smooth muscle contraction; Hippo signalling pathway; Focal adhesion; Platelet activation; Long-term potentiation; Dopaminergic synapse; Inflammatory mediator regulation of

				TRP channels; Regulation of actin cytoskeleton; Insulin signalling pathway; Oxytocin signalling pathway; Insulin resistance; Amphetamine addiction; Alcoholism; Herpes simplex virus 1 infection; Proteoglycans in cancer; Diabetic cardiomyopathy; mRNA surveillance pathway; RNA degradation
		28	*VSTM4; LRRC18; ARHGAP22*	
White vs Speckled-ZV	White	19	*CRK; MYO1C; INPP5K; PITPNA; SLC43A2; SCARF1; RILP; PRPF8; TLCD2; WDR81; SERPINF2; SERPINF1; SMYD4; RPA1; RTN4RL1; DPH1*	MAPK signalling pathway; ErbB signalling pathway; Rap1 signalling pathway; Chemokine signalling pathway; Focal adhesion; Fc gamma R-mediated phagocytosis; Neurotrophin signalling pathway; Regulation of actin cytoskeleton; Insulin signalling pathway; Growth hormone synthesis, secretion and action; Bacterial invasion of epithelial cells; Yersinia infection; Human cytomegalovirus infection; Human immunodeficiency virus 1 infection; Pathways in cancer; MicroRNAs in cancer; Renal cell carcinoma; Chronic myeloid leukemia; Motor proteins; Inositol phosphate metabolism; Metabolic pathways; Phagosome; Salmonella infection; Spliceosome; Complement and coagulation cascades; Wnt signalling pathway; RNA polymerase;
		28	*REEP3; JMJD1C*	Transcriptional misregulation in cancer
	Speckled-ZV	5	*TRHDE; ATXN7L3B*	
		14	*SNTG1; PPDPFL; SNAI2*	Hippo signalling pathway; Adherens junction
		24	*RNF152; PIGN; RELCH*	mTOR signalling pathway; Glycosylphosphatidylinositol (GPI)-anchor biosynthesis; Metabolic pathways
White vs Patchy	White	5	*HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer; MAPK signalling pathway
		6	*SULT1B1; SULT1E1; ODAM*	
		17	*EDNRA; PRMT9*	Calcium signalling pathway; cGMP-PKG signalling pathway; cAMP signalling pathway; Neuroactive ligand-receptor interaction; Vascular smooth muscle contraction; Renin secretion; Pathways in cancer
	Patchy	6	*MYOZ2; SYNPO2; SEC24D; METTL14; PRSS12; NDST3; HPGDS; SMARCAD1; ATOH1; GRID2; CCSER1; MMRN1; SNCA; GPRIN3; TIGD2; FAM13A; HERC3; NAP1L5; HERC5; HERC6; PPM1K; ABCG2; PKD2; SPP1; NCAPG; LCORL*	Protein processing in endoplasmic reticulum; Glycosaminoglycan biosynthesis - heparan sulfate / heparin; Metabolic pathways; Glutathione metabolism; Arachidonic acid metabolism; Metabolism of xenobiotics by cytochrome P450; Drug metabolism - cytochrome P450; Metabolic pathways; Chemical carcinogenesis - DNA adducts; Signalling pathways regulating pluripotency of stem cells; Neuroactive ligand-receptor interaction; Long-term depression; Alzheimer disease; Parkinson disease; Pathways of neurodegeneration - multiple diseases; Ubiquitin mediated proteolysis; Antifolate resistance; ABC transporters; Bile secretion; PI3K-Akt signalling pathway; Apelin signalling pathway; Focal adhesion; ECM-receptor interaction; Toll-like receptor signalling pathway; GnRH secretion; Human papillomavirus infection

White vs Belted-TX	White	1	*SLITRK3; SI*	Cell adhesion molecules;
		5	*HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5; MANSC1*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer, MAPK signalling pathway;
		6	*ATP10D; CORIN; NFXL1; CNGA1; NIPAL1; TXK*	cGMP-PKG signalling pathway; cAMP signalling pathway; Phototransduction; Leukocyte transendothelial migration
		17	*EDNRA; PRMT9*	Calcium signalling pathway; cGMP-PKG signalling pathway; cAMP signalling pathway; Neuroactive ligand-receptor interaction; Vascular smooth muscle contraction; Renin secretion; Pathways in cancer
	Belted-TX	2	*TNP1; IGFBP5*	
		9	*EPHA7; CNR1; SPACA1; AKIRIN2*	Axon guidance; Rap1 signalling pathway; Neuroactive ligand-receptor interaction; Thermogenesis; Retrograde endocannabinoid signalling
		18	*IRX6; MMP2; LPCAT2; SLC6A2*	Endocrine resistance; Leukocyte transendothelial migration; GnRH signalling pathway; Estrogen signalling pathway; Relaxin signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Pathways in cancer; Proteoglycans in cancer; Bladder cancer; Diabetic cardiomyopathy; Fluid shear stress and atherosclerosis; Glycerophospholipid metabolism; Ether lipid metabolism; Metabolic pathways; Synaptic vesicle cycle
		29	*E2F8; CSRP3; ZDHHC13; PTPN5; TMEM86A; SPTY2D1; MISFA*	MAPK signalling pathway
White vs Belted-ZV	White	5	*GSG1; FAM234B; HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5; MANSC1*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer; MAPK signalling pathway;
		24	*B4GALT6; TTR; DSG2; DSG3; DSG4; DSG1; DSC1*	Lactosylceramide biosynthesis; Thyroid hormone synthesis; Arrhythmogenic right ventricular cardiomyopathy; Staphylococcus aureus infection
	Belted-ZV	14	*ZFPM2*	MicroRNAs in cancer
		18	*PHLDB3; ETHE1; ZNF575; XRCC1; IRGQ; ZNF576; ZNF428; CADM4; PLAUR; IRGC;*	Sulfur metabolism; Metabolic pathways; Base excision repair; Insulin secretion; GnRH secretion; Salivary secretion; Protein digestion and absorption; Herpes simplex virus 1 infection;

			SMG9; KCNN4; LYPD5; ZNF404; ZNF45; ZNF234; ZNF226	
		20	*ADCY2*	Purine metabolism; Metabolic pathways; Endocrine resistance; Rap1 signalling pathway; Calcium signalling pathway; cGMP-PKG signalling pathway; cAMP signalling pathway; Chemokine signalling pathway; Phospholipase D signalling pathway; Oocyte meiosis; Longevity regulating pathway; Longevity regulating pathway - multiple species; Adrenergic signalling in cardiomyocytes; Vascular smooth muscle contraction; Apelin signalling pathway; Gap junction; Platelet activation; Circadian entrainment; Thermogenesis; Retrograde endocannabinoid signalling; Glutamatergic synapse; Cholinergic synapse; GABAergic synapse; Inflammatory mediator regulation of TRP channels; Insulin secretion; GnRH signalling pathway; Ovarian steroidogenesis; Progesterone-mediated oocyte maturation; Estrogen signalling pathway; Melanogenesis; Thyroid hormone synthesis; Oxytocin signalling pathway; Glucagon signalling pathway; Regulation of lipolysis in adipocytes; Aldosterone synthesis and secretion; Relaxin signalling pathway; Cortisol synthesis and secretion; Parathyroid hormone synthesis, secretion and action; Cushing syndrome; Growth hormone synthesis, secretion and action; Salivary secretion; Gastric acid secretion; Pancreatic secretion; Bile secretion; Morphine addiction; Human cytomegalovirus infection; Human T-cell leukemia virus 1 infection; Pathways in cancer; Chemical carcinogenesis - receptor activation; Dilated cardiomyopathy
White vs White-Sided	White	5	*GSG1; FAM234B; HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer; MAPK signalling pathway;
		29	*PAG12; PAG-4; PAG-11*	Protein digestion and absorption; Protein digestion and absorption
	White-Sided	2	*MBD5; EPC2; KIF5C; LYPD6B; LYPD6*	Polycomb repressive complex; ATP-dependent chromatin remodeling; Endocytosis; Dopaminergic synapse; Motor proteins; Alzheimer disease; Parkinson disease; Amyotrophic lateral sclerosis; Huntington disease; Prion disease; Pathways of neurodegeneration - multiple diseases; Salmonella infection; Non-small cell lung cancer; Neuroactive ligand-receptor interaction; Glycerophospholipid metabolism; Ether lipid metabolism; Arachidonic acid metabolism; Linoleic acid metabolism; alpha-Linolenic acid metabolism; Metabolic pathways; Ras signalling pathway; Vascular smooth muscle contraction; Pancreatic secretion; Fat digestion and absorption
		9	*SEC63; SCML4; SOBP; PDSS2*	Protein export; Protein processing in endoplasmic reticulum; Terpenoid backbone biosynthesis
		12	*KLHL1*	
		20	*NPR3; CDH9;*	
White vs Blacklegs	White	4	*ORC5; RELN*	Cell cycle; PI3K-Akt signalling pathway; Focal adhesion; ECM-receptor interaction; Spinocerebellar ataxia; Human papillomavirus infection

	5	*GSG1; FAM234B; HEBP1; GPRC5D; GPRC5A; DDX47; APOLD1; CDKN1B; GPR19; CREBL2; DUSP16; BORCS5*	Endocrine resistance; ErbB signalling pathway; HIF-1 signalling pathway; FoxO signalling pathway; Cell cycle; PI3K-Akt signalling pathway; AGE-RAGE signalling pathway in diabetic complications; Cushing syndrome; Measles; Human papillomavirus infection; Epstein-Barr virus infection; Pathways in cancer; Transcriptional misregulation in cancer; Viral carcinogenesis; MicroRNAs in cancer; Prostate cancer; Chronic myeloid leukemia; Small cell lung cancer; Gastric cancer; MAPK signalling pathway;
	22	*PLXND1; H1-8; RHO; IFT122; MBD4; EFCAB12; RPL32; CAND2; RAF1*	Base excision repair; Ribosome; Coronavirus disease - COVID-19; EGFR tyrosine kinase inhibitor resistance; Endocrine resistance; MAPK signalling pathway; ErbB signalling pathway; Ras signalling pathway; Rap1 signalling pathway; cGMP-PKG signalling pathway; cAMP signalling pathway; Chemokine signalling pathway; FoxO signalling pathway; Sphingolipid signalling pathway; Phospholipase D signalling pathway; Autophagy – animal; mTOR signalling pathway; PI3K-Akt signalling pathway; Apoptosis; Cellular senescence; Vascular smooth muscle contraction; Axon guidance; VEGF signalling pathway; Apelin signalling pathway; Focal adhesion; Gap junction; Signalling pathways regulating pluripotency of stem cells; Neutrophil extracellular trap formation; C-type lectin receptor signalling pathway; JAK-STAT signalling pathway; Natural killer cell mediated cytotoxicity; T cell receptor signalling pathway; B cell receptor signalling pathway; Fc epsilon RI signalling pathway; Fc gamma R-mediated phagocytosis; Long-term potentiation; Neurotrophin signalling pathway; Serotonergic synapse; Long-term depression; Regulation of actin cytoskeleton; Insulin signalling pathway; GnRH signalling pathway; Progesterone-mediated oocyte maturation; Estrogen signalling pathway; Melanogenesis; Prolactin signalling pathway; Thyroid hormone signalling pathway; Oxytocin signalling pathway; Relaxin signalling pathway; Parathyroid hormone synthesis, secretion and action; GnRH secretion; Growth hormone synthesis, secretion and action; Alzheimer disease; Pathways of neurodegeneration - multiple diseases; Alcoholism; Salmonella infection; Tuberculosis; Hepatitis C; Hepatitis B; Human cytomegalovirus infection; Influenza A; Human papillomavirus infection; Kaposi sarcoma-associated herpesvirus infection; Human immunodeficiency virus 1 infection; Pathways in cancer; Proteoglycans in cancer; MicroRNAs in cancer; Chemical carcinogenesis - receptor activation; Chemical carcinogenesis - reactive oxygen species; Colorectal cancer; Renal cell carcinoma; Pancreatic cancer; Endometrial cancer; Glioma; Prostate cancer; Melanoma; Bladder cancer; Chronic myeloid leukemia; Acute myeloid leukemia; Non-small cell lung cancer; Breast cancer; Hepatocellular carcinoma; Gastric cancer; Central carbon metabolism in cancer; Choline metabolism in cancer; PD-L1 expression and PD-1 checkpoint pathway in cancer
	29	*PAG12; PAG-4; PAG-11*	Protein digestion and absorption
Blacklegs	11	*UBXN2A; ATAD2B; KLHL29*	Protein processing in endoplasmic reticulum;
	15	*ARHGAP42; PGR; TRPC6*	Oocyte meiosis; Progesterone-mediated oocyte maturation; Estrogen signalling pathway; Chemical carcinogenesis - receptor activation; Breast cancer; cGMP-PKG signalling pathway; Axon guidance
	20	*ADCY2*	Purine metabolism; Metabolic pathways; Endocrine resistance; Rap1 signalling pathway; Calcium signalling pathway; cGMP-PKG signalling pathway; cAMP signalling pathway; Chemokine signalling pathway; Phospholipase D signalling pathway; Oocyte meiosis; Longevity regulating pathway; Longevity regulating pathway - multiple species; Adrenergic signalling in cardiomyocytes; Vascular smooth muscle contraction; Apelin signalling pathway; Gap junction; Platelet activation; Circadian entrainment; Thermogenesis; Retrograde endocannabinoid signalling;

	Glutamatergic synapse; Cholinergic synapse; GABAergic synapse; Inflammatory mediator regulation of TRP channels; Insulin secretion; GnRH signalling pathway; Ovarian steroidogenesis; Progesterone-mediated oocyte maturation; Estrogen signalling pathway; Melanogenesis; Thyroid hormone synthesis; Oxytocin signalling pathway; Glucagon signalling pathway; Regulation of lipolysis in adipocytes; Aldosterone synthesis and secretion; Relaxin signalling pathway; Cortisol synthesis and secretion; Parathyroid hormone synthesis, secretion and action; Cushing syndrome; Growth hormone synthesis, secretion and action; Salivary secretion; Gastric acid secretion; Pancreatic secretion; Bile secretion; Morphine addiction; Human cytomegalovirus infection; Human T-cell leukemia virus 1 infection; Pathways in cancer; Chemical carcinogenesis - receptor activation; Dilated cardiomyopathy.

5.4 Discussion

South African goats found in extensive communal production systems are a unique genetic reservoir represented by their diverse coat coloration which has gone mostly unstudied (Sevane et al., 2018). The analysis of signatures of selection has previously been used to report on vital candidate genes that influence numerous economically important traits for goat production such as growth and development, metabolism, meat and milk production, reproduction, mohair, and disease resistance (Sejian et al., 2021; Henkel et al., 2019; Gou et al., 2018; Akis et al., 2012; Becerril et al., 1996). This information has been useful in improving the accuracy of selection through the implementation of genomic selection which uses genotype and phenotype information to inform animal genetic improvement strategies (Torres-Hernández et al., 2022). Although there has been progress made in this regard, information about the candidate genes of several other production traits such as coat colour with the knowledge of selection pressure is still lacking, especially in the ecotype populations found in South African village goats (Henkel et al., 2019).

The absence of improved breeding schemes within communal production systems is resulting in the prevalence of indiscriminate crossbreeding within indigenous village goats which has put these populations at higher risk of extinction (Rajaganapathy et al., 2018; Ojo et al., 2015). Because of this, the implementation of improved selection schemes within communal production systems is one of the main strategies for the genetic improvement of these local breeds. To achieve this, some farmers have turned to selecting replacements based on appearance, considering coat colour as a selection criterion (Henkel et al., 2019; Akis et al., 2012; Adalsteinsson et al., 1994). However, the usefulness of coat colour in selection schemes is limited as information on coat colour candidate genes for local populations is lacking (Becerril et al., 1996). Based on the above, the objective of this study was to detect candidate genes for the coat colours and coat colour patterns of South African meat-type village goats. This information has the potential to provide insights into the mechanisms underlying goat coat colour variation, as well as shed light on its relationship with other economical traits important for goat productivity, and adaptation to diverse climatic and environmental conditions. Such information will be vital towards the consideration of coat colour in selection schemes and its incorporation into genetic improvement programs in marginal or small-scale systems.

In small communal production systems, coat colour has previously been demonstrated as an important production trait as farmer and consumer preferences for specific skin colours and

fiber characteristics, such as diameter, length, and colour have resulted in certain coat colours and coat colour patterns having higher economic value compared to others (Peng et al., 2019). For instance, white coats tend to have a higher economic value, due to consumer preferences based on their use in cultural and religious ceremonies, and because the white fleece can be dyed to create coats of any colour (Bhat et al., 2019). In Angora goats, it has been reported that directional selection for white fibers has resulted in the prevalence of the white coat (Dong et al., 2015). A relationship between coat colour and animal price has also been reported in local populations from other regions in Africa. For instance, in West Africa, black goats have the highest economic value (Chokoe et al., 2020). In contrast, in Ethiopia black goats have low economic value as it is the least-preferred coat colour (Nguluma et al., 2020). Such relationships between coat colour and animal economic value illustrate the importance of considering coat colour during selection even beyond fiber-producing animals where it has a direct impact on profitability.

Coat colour has also been linked to other goat production traits, such as reproduction, and adaptability. For instance, in Black Bengal goats, black males have a higher market value as they reproduce more than males with brown/reddish coats which have lower economic value (Berhanu et al., 2012). It is suspected that the black goats have improved reproductive parameters due to their coat which plays a role in their environmental adaptability by protecting them against photochemical damage from UV rays (Ferreira et al., 2021). Furthermore, the black coat of these goats has been linked to numerous other economically important traits, with black goats demonstrating increased performance for milk production, kidding interval and litter size compared to other colours (Baenyi Simon et al., 2020; Mia et al., 2020; Choudhury et al., 2013). In West African Dwarf goats, it was demonstrated that coat colour impacts prolificacy, fecundity, and litter size at birth and weaning, as well as in weights at weaning (Daramola et al., 2009). Although these studies have demonstrated the associations between coat colour and other production traits, such information for South African meat-type village goats is lacking.

Past studies have revealed that goats coat colour depends on the presence and distribution of eumelanin and pheomelanin, with Eumelanin found to be responsible for darker coats including black and brown pigments, while pheomelanin is responsible for lighter coats including red and yellow pigments (Peng et al., 2019). Furthermore, the distribution and expression of these two types of melanin is controlled by the Agouti Signalling Protein (*ASIP*) (Peng et al., 2019). Studies have shown that *ASIP* signals the production of pheomelanin and as a result has been

detected in high expression in white coated animals (Peng et al., 2019). This is because *ASIP* inhibits the production of eumelanin by out competing the α-melanocyte-stimulating hormone (α-MSH) to the binding site of the melanocortin 1 receptor (*MC1R*) (Nasti and Timares, 2015). In this sense, coat coloration in goats is controlled by a complex signal-gene regulatory network that is involved the differentiation and maturation of melanin producing cells which determine the amount, quality, and distribution of melanin in the body. Through this complex network, melanin production in melanocytes is triggered by the signalization induced by *MC1R* and culminates with the regulation of Tyrosinase (*TYR*), Tyrosine Protein Related 1 (*TRP1*), and *TRP2* (Robinson et al., 2000).

In recent years, studies have revealed that along with numerous candidate genes, different epigenetic mechanisms such as DNA methylation, histone modification, chromatin remodelling, and noncoding RNA influence melanogenesis both directly and indirectly by regulating key genes and signalling pathways that are involved in melanogenesis (Beerman and Rossi, 2015). For instance, recent epigenetic studies have shown that the effects of *ASIP* on the synthesis of pigmentation depends on the methylation grade on Cytosine Phospho-Guanine (CpG) retrotransposon intracisternal A particle (IAP) (Barbot et al., 2002). In addition, Laible et al. (2021), revealed that a mutation of the PEML gene is responsible for coat colour dilution in cattle. While the study by Van Buren et al. (2020), identified a variant of the *MLPH* gene that causes coat colour dilution in dogs.

However, because the composition of coat colour is a complex process which involves numerous candidate genes, some of which have not yet been sequenced (or had their variability levels quantified), the molecular base of the heritage for the trait is still poorly understood in goats. In addition, some techniques, such as quantitative real time-polymerase chain reaction have shown that there is no single locus explaining colour divergence, suggesting that this is a polygenic trait with prominent pleiotropic, epistatic, and non-epistatic effects influenced by the diet, age, and environment (Zhao, et al., 2015). Therefore, genetic analysis focused on understanding both genetic and epigenetic effects on coat colour variation can help to understand the genomic architecture of the trait and hence offer the necessary knowledge breeders need to design more efficient selection strategies for genetic improvement.

5.4.1 Within-population signatures of selection

The Boer, Savannah and Kalahari Red breeds have been selected and reared under intensive commercial systems of production. Hence, it was expected that the detection of selection

signatures within these populations would reveal genes related to economically important production traits. Although no pathways were found to be associated with the *CDH12* gene in goats (found within the Boer breed), the gene has previously been linked to the regulation of skin colour in humans (Charfeddine et al.2020). The genes detected within the other two commercial breeds (i.e., the Savanna and Kalahari Red) were revealed to be associated with pathways that play roles in biological functions including goat metabolism (i.e., *BAAT, GRIN3A*), and immune response (i.e., *IKZF2, PPP2R2B, TLERG1, ARAP3, DIAPH1, RELL2*). As such, these results revealed that such traits are the likely targets of the intensive selection the goats undergo within commercial production systems. Previous studies have reported that goats within commercial production systems have higher degrees of inbreeding which has limited the goats' environmental adaptation (Mdladla et al., 2016). Therefore, the detection of genes involved in skin colour, metabolism, and immune response within these commercial breeds suggests that there is a strong desire for goats with improved environmental adaptation within the commercial sector.

Unlike the commercial breeds, village goat populations (i.e., ecotypes) are constantly exposed to adverse climatic and environmental conditions (Mdladla et al., 2016). These adverse climatic and environmental effects impose natural selection pressures on the populations, resulting in the increase of the allele frequencies of genes involved in goat adaptation. Therefore, it was expected that studying the selection signatures present within these populations would reveal the targets of natural selection pressures, in turn revealing genetic loci underlying local goat adaptation. Such insights are crucial for guiding genetic improvement strategies and conservation programs for these populations. As expected, candidate genes under selection within these village goats with various coat colours and patterns were mostly involved in pathways associated with numerous traits including metabolism, reproduction, and immune response.

These results reflect the effects of natural selection pressures resulting from climatic and environmental factors that have influenced the goats' adaptations to local environments and communal production systems which are usually exposed to harsh environmental conditions and scarce feed. These findings offer possible explanations for reports that have revealed that ecotype populations have a greater capacity for adaptation and higher efficiency in consumption and utilization of fodder, which has allowed them to develop, produce and reproduce in harsh environments (Torres-Hernández et al., 2022; Akhtar et al., 2021). Furthermore, ecotypes have been revealed to be more resilient to higher temperatures, while

possessing smaller corporal size, which proportionally produce less metabolic heat (Nair et al., 2021).

Along with genes involved with economical traits such as metabolism, reproduction, and immune response, numerous candidate genes involved in goat coat colour were detected within the village meat-type goats (Table 5.3). For instance, of the significant markers detected within white village goats, genes potentially involved with white coat colour formation included the *DUSP16* gene found on chromosome 5 which is involved in the MAPK signalling pathway (chx04010). The MAPK signalling pathway has been reported to be involved in the goat coat pigmentation through the regulation of melanogenesis (D'Mello et al., 2016). Another coat colour candidate gene was the *CDKN1B* gene on chromosome 5, which is involved in several pathways including the PI3K-Akt signalling pathway (chx04151) which has been linked to melanogenesis and the formation of coat pigment in various livestock species including cattle (D'Mello et al., 2016). Furthermore, the *TYRP1* gene on chromosome 8 was found to be under selection within the white village goats. The gene is involved in several pathways including Tyrosine metabolism (chx00350), Metabolic pathways (chx01100), as well as Melanogenesis (chx04916). The gene has previously been linked to the coat colouration of various animals including pigs (Ren et al., 2011), dogs (Schmutz et al., 2002), rabbits (Jia et al., 2021), cattle (Berryere et al. 2003), cats (Lyons et al. 2005), and sheep (Gratten et al. 2007). Hence, it's detection in this study suggests that it may be implicated in goat coat colour as well.

Within red goats, the *KDR* gene on chromosome 6 was found to be under selection. The gene is involved in several pathways including the MAPK signalling pathway (chx04010), and the PI3K-Akt signalling pathway (chx04151), two pathways that have been associated with the regulation of coat colour variation in mammals. Furthermore, David et al., (2014) revealed that the *KDR* gene forms a complex with *KIT*, and *PDGFRA* which together are responsible for the red coat colour pattern in Angus cattle. The gene has been revealed to also play a role in pig coat colour (Xu et al., 2020). Under selection within black village goats, the *FGF10* gene was found on chromosome 20. The *FGF10* gene is likely a candidate gene for black coat colour as it is involved in several pathways that have been linked to coat colour formation including the MAPK signalling pathway (chx04010), the Rap1 signalling pathway (chx04015), as well as the PI3K-Akt signalling pathway (chx04151). The gene has previously been found under selection in black Xichuan black-bone chickens (Li et al., 2020). Candidate genes under selection within grey coat coloured village goats included the *COL6A3* gene on chromosomes

3, and the *HEY2* gene on chromosome 9. The *COL6A3* gene is involved in PI3K-Akt signalling pathway (chx04151), while the *HEY2* gene is involved in several pathways including the Notch signalling pathway (chx04330). The Notch signalling pathway has previously been linked to hair greying in mice (Schouwey and Beermann, 2008).

Candidate coat colour genes under selection within goats with the speckled coat pattern included *TYRP1* on chromosome 8 which was detected in speckled goats within the TZ cluster, as well as the *DUSP16* gene on chromosome 5, which was found in speckled goats within the ZV cluster. The *TYRP1* is involved in Tyrosine metabolism (chx00350), Metabolic pathways (chx01100), and Melanogenesis (chx04916). The *DUSP16* gene is involved in MAPK signalling pathway (chx04010). Candidate coat colour genes shared by speckled goats in both clusters were found on chromosome 19 and these were linked to several genes including the *CRK* gene which is involved in MAPK signalling pathway (chx04010), and the *SERPINF1* gene which is involved in Wnt signalling pathway (chx04310). Candidate coat colour genes under selection for the patchy coat pattern included the *COL6A3* gene on chromosome 3, and the *TYRP1* gene on chromosome 8. The *COL6A3* gene is involved in the PI3K-Akt signalling pathway (chx04151), while the *TYRP1* is involved in Tyrosine metabolism (chx00350), Metabolic pathways (chx01100), and Melanogenesis (chx04916).

Candidate coat colour genes under selection for the belted coat pattern included the *TYRP1* gene on chromosome 8, and the *MAP3K7* gene on chromosome 9, which were found within belted goats in the TX cluster. The *TYRP1* gene is involved in Tyrosine metabolism (chx00350), Metabolic pathways (chx01100), and Melanogenesis (chx04916), while the *MAP3K7* gene is involved in MAPK signalling pathway (chx04010), and Wnt signalling pathway (chx04310). Whereas candidate coat colour genes within belted goats in the ZV cluster included *CRK* on chromosome 19, which is involved in MAPK signalling pathway (chx04010), and Rap1 signalling pathway (chx04015), as well as the *SERPINF1* gene on the same chromosome which is involved in Wnt signalling pathway (chx04310). Candidate coat colour genes under selection for the white sided coat pattern included the *CREB5*, *RELN*, and the *WNT2* genes on chromosome 4, as well as the *CRK* and the *SERPINF1* genes on chromosome 19. The *CREB5* and the *RELN* genes are involved in several pathways including PI3K-Akt signalling pathway (chx04151), while the *WNT2* gene is involved in Wnt signalling pathway (chx04310), and Melanogenesis (chx04916). The *CRK* gene is involved in MAPK signalling pathway (chx04010), while the *SERPINF1* gene is involved in Wnt signalling

pathway (chx04310). Candidate coat colour genes under selection for the black-legs coat pattern included the *CRK* and the *SERPINF1* genes on chromosome 19 which are involved in the MAPK signalling pathway (chx04010), and the Wnt signalling pathway (chx04310), respectively. These results reflect the effects of natural selection pressures resulting from climatic and environmental factors that have influenced goat coat colour and coat colour pattern diversity due to adaptations to local environments.

The coat of an animal (skin and coat) is significantly linked to the environment, as it acts as the primary barrier of protection from external threats (Al-Dawood, 2017). The trait is highly variable and can differ in terms of colour, depth, and length, and has been shown to be a key factor of an animal's adaptation to different climates with animals with darker coats having higher thermotolerance due to their ability to absorb more direct solar radiation compared to animals with lighter coats which usually have lower thermotolerance (Ferreira et al., 2021; Joy et al., 2020a,b; Onasanya et al., 2018; Al-Dawood, 2017). Such findings suggest that goats with different coat colours possess different thermoregulatory abilities (Acharya et al., 1995). This is supported by studies that have investigated heat tolerance in livestock which have indicated that coat colour affects heat tolerance criterions such as rectal temperature, and respiratory rate (Joy et al., 2020a,b;). In addition, several studies have reported lower urinary frequency in dark-coated animals compared to light-coated animals which were found to have higher solar radiation reflection, lower heat absorption, and lower surface temperature (Joy et al., 2020a,b, de Souza et al., 2014).

Other researchers point out that black goats are better at adapting to cold, as the black pigment helps them warm up faster (Chokoe et al., 2020). Similarly, Ferreira et al. (2021) state that, in rainfall seasons, the black colour lets the goats catch more sun energy in cold mornings, supporting homeostasis. Furthermore, Nigerian farmers prefer black goats because white goats are at greater risk of being stolen and of being attacked by predators due to their coats (Torres-Hernández et al., 2022; Berhanu et al., 2012). However, because coat colour variation is influenced by both genetic and epigenetic factors (i.e., age or intensity of solar radiation including) (Rosero Alpala et al., 2021), further research is necessary to evaluate, in a particular way, the productive response of the animals with different coat colours in every environment. Such information has the potential to aid farmers determine coat colours that are ideal for a given climate or environment.

5.4.2 Cross population signatures of selection

Signatures found between the Savanna and the Kalahari Red breeds were linked to various genes linked to economical traits such as adaptation and reproduction. For instance, genes such as *GRIN3A*, and *SCH3*, under selection in Kalahari Red goats were found to have roles in metabolism and reproduction, respectively. While in the Savanna goats, genes such as *EXTL2*, and *ROCK2* were revealed to be involved in metabolism and environmental adaptation, respectively. Signatures found between the Savanna and the Boer breeds were linked to various genes linked to economical traits such as growth and development (i.e., *LTBP1* in Boer goats), coat colour (i.e., *RAF1* in Savanna goats), and adaptation (i.e., *PPARG* in Savanna goats). Signatures found between the Kalahari Red and Boer goats were involved in coat colour (i.e., *RAF1*, and *MITF*), and immune response (i.e., *UBA3*). The *MITF* gene has previously been shown to regulate white coat pigmentation in various animals including llamas (Anello et al., 2019), dogs (Schmutz et al., 2009), horses (Henkel et al., 2019), and sheep (Han et al., 2015).

Genes under selection between the Savanna breed and the white village goats were linked to metabolism. Genes such as *EXTL2, COL11A1, S1PR1, RAP1A* and *VAV3* were found in the Savanna goats, signalling artificial selection for adaptation to intensive production systems. Genes under selection between the Boer breed and the white village goats were similarly linked to metabolism. These genes included the *VAV3, SLC36A2*, and *BAAT*. Genes under selection between the Kalahari Red breed and the red village goats were revealed to be involved with multiple economical traits including growth and development (i.e., *CKS2*, and *SHC3*), metabolism (i.e., *ALDOB, GRIN3A, BAAT, CDK20, PSPH, PHKG1, VKDRC1L1*, and *PLPPR1*), and immune response (i.e., *ZNF189*). Selection signatures detected between commercial and village goats with similar coat colours (i.e., white village goats and Savanna goats; as well as red village goats and Kalahari Red goats), highlight key genetic differences between these populations suggesting that selection under intensive production systems has resulted in artificial selection pressures that have targeted traits related to improved goat productivity for the commercial breeds compared to the village goat populations from which these improved breeds likely descend.

Genes under selection between white and the red village goats were revealed to be involved with metabolism (*FXN*), and coat colour. Of the coat colour genes, *MAGI2, CDKN1B*, and *DUSP16* were found in white goats. While *FGF17, PPP3CC*, and *CTNND2* were found in red goats. These genes likely highlight the genetic difference between red and white village goats. Between white and black village goats, genes under selection involved with coat colour included the *MAGI2* and the *DUSP16* genes in white goats. While the *FTO, IRX3*, and the

IIRX5 genes, which were found in the black goats were associated with metabolism, and diseases, respectively. Genes under selection between white and the grey village goats were revealed to be involved with multiple economical traits including coat colour (i.e., *MAGI2, CDKN1B,* and *DUSP16*), immune response (i.e., *CUL3, FAM124B, FAM89A, UUUGT2A2* and *TRIM67*), and metabolism (i.e., *GNPAT*). Genes under selection between black and grey village goats were revealed to be involved with multiple economical traits including immune response (i.e., *ARAP1, RNASEK, ASGR1*, and *ACAP1*), metabolism (i.e., *ACAD8, PLD2, ALOX12, DLG4, ACADVL, GABARAP*, and *GPS2*), and coat colour (i.e., *DUL2*, and *ARRB2*).

Genes under selection between white and speckled village goats were revealed to be involved with multiple economical traits including metabolism (i.e., *PPP1CC, INPP5K, PIGN*, and *SNA12*), immune response (i.e., *RILP, PRPF8*, and *SERPINF2*), and coat colour (i.e., *ADAMTS20, CRK, IRAK4, SERPINF1* and *PRICKLE1*). The *ADAMTS20* gene has previously been associated with black/brown/grey French goats (Oget et al., 2019). Genes under selection between white and patchy village goats were revealed to be involved with multiple economical traits including metabolism (i.e., *ABCG2, SEC24D, NDST3,* and *HPGD5*), development (i.e., *SMARCAD1*), and coat colour (i.e., *EDNRA*, and *SPP1*). Genes under selection between white and belted village goats were revealed to be involved with multiple economical traits including metabolism (i.e., *LPCAT2, ETHE1*, and *KZNN4*), coat colour (i.e., *EDNRA, B4GALT6, PTPN5*, and *ADCY2*), immune response (i.e., *MMP2*, *DSG1*, and *ZNF404*), and adaptation (i.e., *CNR1*). Genes under selection between white and white-sided village goats were revealed to be involved with economical traits such as metabolism (i.e., *PAG11, PAG-4, PAG12, EPC2, SEC63, PDSS2*) and immune response (i.e., *KIF5C*). Genes under selection between white and blacklegged village goats were revealed to be involved with economical traits such as metabolism (i.e., *PAG11, PAG-4, PAG11,* and *UBXN2A*), reproduction (i.e., *PGR*), and coat colour (i.e., *ADCY2*, and *RAF1*). The detection of several genes involved in different goat economical traits such goat metabolism, growth and reproduction, immune response, and adaptability together with coat colour genes may signal a relationship between selection for coat colour and overall goat productivity.

Although the relationship between coat colour and goat productivity is still relatively unclear, there have been studies that have suggested that it influences the morpho-structural composition of the animals (Ayoola et al., 2018). For instance, coat colour in Black Bengal goats has been showed to affect morphometry and production (Choudhury et al., 2013), leading to the belief that coat colour is associated with reproductive and productive traits (Anzures-

Olvera et al., 2019). Furthermore, Mia et al. (2020) has revealed that coat colour can affect goat productivity by influencing animals' caloric load which impacts their metabolism. Daramola et al. (2009) found an association between coat colour and milk production, observing increased milk production with the decrease of coat colour intensity. This finding was further supported by the review by Gupta and Mondal (2019), which inferred that there is a drastic reduction in milk quantity and quality with the advance of thermal stress. Concerning meat production, some African villages select for darker coats believing that eating their meat has higher religious and health benefits (Dossa et al., 2015). Archana et al. (2018) found that Salem black goats have increased meat quality due to their higher resilience to thermal stress. In addition, several studies found that black coated animals performed better in regard to growth rate, weight gain, and survival rate, compared to lighter animals (Hossain et al., 2020; Choudhury et al., 2013; Daramola et al., 2009).

As demonstrated above, coat colour has been revealed to impact goat productivity by affecting production traits such as environmental adaptation, reproduction, and milk and meat production. However, the trait is often neglected in livestock improvement programs that aim to increase animal productivity. Incorporating coat colour as a vital part of the selection criteria during the selection of reproductive animals will prove advantageous, especially in low scale production systems (Abraham et al., 2018; Odubote, 1994). Instances of producers employing indirect selection criteria based on coat colour have already been reported (Berhanu et al., 2012). For instance, in some agro-pastoral production systems in Africa, bucks are selected based on their appearance, with a predilection to spotted or mottled animals over flat coloured ones (Misbah et al., 2015). In Ethiopian and South African communal production systems, goats are bred with a variety of colours for specific religious/cultural rituals (Mdladla et al., 2016). Consequently, the distribution of colour variations within the local populations is closely influenced by smallholder and consumer preferences and demands for specific coat colours, which relate to the animals' meat and skin quality, goat adaptability and overall productivity (Akhtar et al., 2021).

In this sense, developing indirect selection schemes that integrate coat colour as a selection criterion is relevant for these production systems (Nguluma et al., 2020). However, in goats, information about the relationships between coat colour and other goat productivity traits is still minimal. While some efforts have been made to reveal the effects of coat colour on local goats' longevity and phenotypic differentiation, more information needs to be generated and corroborated. The analysis of the coat colour will help avoid the loss of these goat local

populations, as it has happened in other species, such as Menz sheep, where the loss of the black colour due to crossbreeding led to decreased population productivity (Getachew et al., 2020). Therefore, including morphologic characteristics related to goat productivity and adaptation, such as coat colour, is a relevant element in genetic improvement programs, which will ensure the survival of local breeds as they are faced with high risk of erosion due to indiscriminate crossbreeding and replacement with more productive commercial breeds.

Furthermore, as coat colour is one of the most important means of identification between and within breeds (Berhanu et al., 2012), an indirect selection criterion based on animal appearance would allow greater ease for the recognition of desirable animals providing small producers with an affordable strategy to improve the production yield of their flocks (Maldonado-Jáquez et al., 2023; Ramos-Martínez et al., 2020; Torres-Hernández et al., 2020; Abraham et al., 2018; Tadesse et al., 2014). Therefore, the results from this study which revealed that coat colour is under selection within South African meat-type village goats and suggested a likely association between coat colour and other goat production traits such as goat metabolism, reproduction, and immune response will be useful in that they will aid the implementation of coat colour as a selection criterion in goat breeding programs in low-scale communal production systems. These finding agree with several previous studies which have demonstrated associations between coat colour and various goat production traits. However, the analysis of more animals with different coat colour phenotypes may further strengthen these findings and confirm the effect of the genes identified for coat colour variation.

5.5 References

ABRAHAM, H., GIZAW, S. & URGE, M. 2018. Identification of breeding objectives for Begait goat in western Tigray, North Ethiopia. *Trop Anim Health Prod,* 50**,** 1887-1892.

ACHARYA, R. M., GUPTA, U. D., SEHGAL, J. P. & SINGH, M. 1995. Coat characteristics of goats in relation to heat tolerance in the hot tropics. *Small Ruminant Research,* 18**,** 245-248.

ADALSTEINSSON, S., SPONENBERG, D. P., ALEXIEVA, S. & RUSSEL, A. J. 1994. Inheritance of goat coat colors. *J Hered,* 85**,** 267-72.

AKHTAR, A., HOQUE, M., BHUIYAN, A., AMIN, M. & HABIB, M. 2021. Phenotypic performance and response to selection for body weight gains of Black Bengal goat in a community breeding program. *African Journal of Agricultural Research,* 17**,** 1058-1067.

AKIS, I., ÖZTABAK, K., ESEN GURSEL, F. & ÜN, C. 2012. Goat agouti gene polymorphism and its association with coat color in indigenous turkish goat breeds. *Veterinarija ir Zootechnika,* 60**,** 3-6.

AL-DAWOOD, A. 2017. Towards heat stress management in small ruminants–a review. *Annals of Animal Science,* 17**,** 59-88.

ANELLO, M., DAVERIO, M. S., SILBESTRO, M. B., VIDAL-RIOJA, L. & DI ROCCO, F. 2019. Characterization and expression analysis of KIT and MITF-M genes in llamas and their relation to white coat color. *Animal genetics,* 50**,** 143-149.

ANZURES-OLVERA, F., VÉLIZ, F. G., DE SANTIAGO, A., GARCÍA, J. E., MELLADO, J., MACÍAS-CRUZ, U., AVENDAÑO-REYES, L. & MELLADO, M. 2019. The impact of hair coat color on physiological variables, reproductive performance and milk yield of Holstein cows in a hot environment. *Journal of Thermal Biology,* 81**,** 82-88.

ARCHANA, P. R., SEJIAN, V., RUBAN, W., BAGATH, M., KRISHNAN, G., ALEENA, J., MANJUNATHAREDDY, G. B., BEENA, V. & BHATTA, R. 2018. Comparative assessment of heat stress induced changes in carcass traits, plasma leptin profile and skeletal muscle myostatin and HSP70 gene expression patterns between indigenous Osmanabadi and Salem Black goat breeds. *Meat Sci,* 141**,** 66-80.

AYOOLA, S., OLAYINKA, A., MUSA, A., MATHEW, W., POPOOLA, M. & BUNJAH, U. 2018. Effect of coat colour on morphological characteristics of Red Sokoto, Sahel and West African Dwarf goats. *Journal of Animal Science and Veterinary Medicine,* 3**,** 47-57.

BAENYI SIMON, P., JUNGA, J., KEAMBOU TIAMBO, C., BIRINDWA, B., KATCHO, K., TAREKEGN, G., OCHIENG, J., TIAMBO, J., BIRINDWA, C., KARUME, A., TAREKEGN, K. & BAENYI, S. 2020. Production Systems, Genetic Diversity and Genes Associated with Prolificacy and Milk Production in Indigenous Goats of Sub-Saharan Africa: A Review. *Open Journal of Animal Sciences,* 10**,** 735-749.

BARBOT, W., DUPRESSOIR, A., LAZAR, V. & HEIDMANN, T. 2002. Epigenetic regulation of an IAP retrotransposon in the aging mouse: progressive demethylation and de-silencing of the element by its repetitive induction. *Nucleic Acids Res,* 30**,** 2365-73.

BECERRIL, C. M., WILCOX, C. J. & GUERRERO, V. M. 1996. Holstein white coat color and performance: phenotypic, genetic and environmental correlations. *Brazilian Journal of Genetics,* 19.

BEERMAN, I. & ROSSI, DERRICK J. 2015. Epigenetic Control of Stem Cell Potential during Homeostasis, Aging, and Disease. *Cell Stem Cell,* 16**,** 613-625.

BERHANU, T., THIENGTHAM, J., TUDSRI, S., ABEBE, G., TERA, A. & PRASANPANICH, S. 2012. Purposes of keeping goats, breed preferences and selection criteria in pastoral and agro-pastoral districts of South Omo Zone. *Livestock Research for Rural Development,* 24.

BERRYERE, T. G., SCHMUTZ, S. M., SCHIMPF, R. J., COWAN, C. M. & POTTER, J. 2003. TYRP1 is associated with dun coat colour in Dexter cattle or how now brown cow? *Anim Genet,* 34**,** 169-75.

BHAT, B., SINGH, A., IQBAL, Z., KAUSHIK, J. K., RAO, A., AHMAD, S. M., BHAT, H., AYAZ, A., SHEIKH, F. & KALRA, S. 2019. Comparative transcriptome analysis reveals the genetic basis of coat color variation in Pashmina goat. *Scientific Reports,* 9**,** 6361.

CHARFEDDINE, C., DALLALI, H., ABDESSALEM, G., GHEDIRA, K., HAMDI, Y., ELOUEJ, S., LANDOULSI, Z., DELAGUE, V., LAGARDE, A., LEVY, N., EL-AMRAOUI, A., BOUBAKER, M. S., ABDELHAK, S. & MOKNI, M. 2020. Identification of a CDH12 potential candidate genetic variant for an autosomal dominant form of transgrediens and progrediens palmoplantar keratoderma in a Tunisian family. *Journal of Human Genetics,* 65**,** 397-410.

CHOKOE, T., MATELELE, T., MAQHASHU, A., FHULUFHELO, R., TUMUDI, M., MPOFU, T., NEPHAWE, K. & MTILENI, B. 2020. Phenotypic Diversity of South African Indigenous Goat Population in Selected Rural Areas. *American Journal of Animal and Veterinary Sciences,* 15**,** 59-66.

CHOUDHURY, M., SARKER, S. C., ISLAM, F., ALI, A., BHUIYAN, A. K. F. H., IBRAHIM, M. N. M. & OKEYO, A. 2013. Morphometry and performance of Black Bengal goats at the rural community level in Bangladesh. *Bangladesh Journal of Animal Science,* 41.

D'MELLO, S. A., FINLAY, G. J., BAGULEY, B. C. & ASKARIAN-AMIRI, M. E. 2016. Signaling Pathways in Melanogenesis. *Int J Mol Sci,* 17.

DARAMOLA, J. O. & ADELOYE, A. A. 2009. Physiological adaptation to the humid tropics with special reference to the West African Dwarf (WAD) goat. *Trop Anim Health Prod,* 41**,** 1005-16.

DAVID, V. A., MENOTTI-RAYMOND, M., WALLACE, A. C., ROELKE, M., KEHLER, J., LEIGHTY, R., EIZIRIK, E., HANNAH, S. S., NELSON, G. & SCHÄFFER, A. A. 2014. Endogenous retrovirus insertion in the KIT oncogene determines white and white spotting in domestic cats. *G3: Genes, Genomes, Genetics,* 4**,** 1881-1891.

DE SOUZA, P. T., SALLES, M. G. F., DA COSTA, A. N. L., CARNEIRO, H. A. V., DE SOUZA, L. P., RONDINA, D. & DE ARAÚJO, A. A. 2014. Physiological and production response of dairy goats bred in a tropical climate. *International Journal of Biometeorology,* 58**,** 1559-1567.

DONG, Y., ZHANG, X., XIE, M., AREFNEZHAD, B., WANG, Z., WANG, W., FENG, S., HUANG, G., GUAN, R. & SHEN, W. 2015. Reference genome of wild goat (capra aegagrus) and sequencing of goat breeds provide insight into genic basis of goat domestication. *BMC genomics,* 16**,** 1-11.

DOSSA, L. H., SANGARÉ, M., BUERKERT, A. & SCHLECHT, E. 2015. Production objectives and breeding practices of urban goat and sheep keepers in West Africa: regional analysis and implications for the development of supportive breeding programs. *SpringerPlus,* 4**,** 281.

FERREIRA, J., SILVEIRA, R. M. F., DE SOUSA, J. E. R., DE VASCONCELOS, A. M., GUILHERMINO, M. M. & FAÇANHA, D. A. E. 2021. Evaluation of homeothermy, acid-base and electrolytic balance of black goats and ewes in an equatorial semi-arid environment. *Journal of Thermal Biology,* 100**,** 103027.

GANDINI, G., TURRI, F., RIZZI, R., CROTTA, M., MINOZZI, G. & PIZZI, F. 2017. Economic evaluation of genetic improvement in local breeds: The case of the Verzaschese goat. *Italian Journal of Animal Science,* 16**,** 1-12.

GAUTIER, M. & VITALIS, R. 2012. rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics,* 28**,** 1176-7.

GETACHEW, T., HAILE, A., TESSEMA, T., DEA, D., EDEA, Z. & RISCHKOWSKY, B. 2020. Participatory identification of breeding objective traits and selection criteria for indigenous goat of the pastoral communities in Ethiopia. *Trop Anim Health Prod,* 52**,** 2145-2155.

GRATTEN, J., BERALDI, D., LOWDER, B. V., MCRAE, A. F., VISSCHER, P. M., PEMBERTON, J. M. & SLATE, J. 2007. Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proc Biol Sci,* 274**,** 619-26.

GUO, J., TAO, H., LI, P., LI, L., ZHONG, T., WANG, L., MA, J., CHEN, X., SONG, T. & ZHANG, H. 2018. Whole-genome sequencing reveals selection signatures associated with important traits in six goat breeds. *Scientific Reports,* 8**,** 10405.

GUPTA, M. & MONDAL, T. 2019. Heat stress and thermoregulatory responses of goats: a review. *Biological Rhythm Research,* 52**,** 1-27.

HAN, J.-L., MIN, Y., GUO, T.-T., YUE, Y.-J., LIU, J.-B., NIU, C.-E., WANG, C.-F. & YANG, B.-H. 2015. Molecular characterization of two candidate genes associated with coat color in Tibetan sheep (Ovis arise). *Journal of integrative agriculture,* 14**,** 1390-1397.

HENKEL, J., SAIF, R., JAGANNATHAN, V., SCHMOCKER, C., ZEINDLER, F., BANGERTER, E., HERREN, U., POSANTZIS, D., BULUT, Z. & AMMANN, P. 2019. Selection signatures in goats reveal copy number variants underlying breed-defining coat color phenotypes. *PLoS genetics,* 15**,** e1008536.

HOSSAIN, S., AFROZ, F., HABIB, M., MIRAZ, M. F. & JALIL, M. 2020. Genetic and phenotypic performance of three coat color variants of Black Bengal Goat in a closed nucleus flock at BLRI.

JIA, X., DING, P., CHEN, S., ZHAO, S., WANG, J. & LAI, S. 2021. Analysis of MC1R, MITF, TYR, TYRP1, and MLPH Genes Polymorphism in Four Rabbit Breeds with Different Coat Colors. *Animals (Basel),* 11.

JOY, A., DUNSHEA, F. R., LEURY, B. J., CLARKE, I. J., DIGIACOMO, K. & CHAUHAN, S. S. 2020a. Resilience of Small Ruminants to Climate Change and Increased Environmental Temperature: A Review. *Animals (Basel),* 10.

JOY, A., SEJIAN, V., KRISHNAN, G., MADIAJAGAN, B., PRATHAP, P. & BHATTA, R. 2020b. Heat stress impact on blood biochemical response and plasma aldosterone level in three different indigenous goat breeds. *Journal of Animal Behaviour and Biometeorology,* 8**,** 266-275.

KANEHISA, M. & SATO, Y. 2020. KEGG Mapper for inferring cellular functions from protein sequences. *Protein Science,* 29**,** 28-35.

LAIBLE, G., COLE, S. A., BROPHY, B., WEI, J., LEATH, S., JIVANJI, S., LITTLEJOHN, M. D. & WELLS, D. N. 2021. Holstein Friesian dairy cattle edited for diluted coat color as a potential adaptation to climate change. *BMC Genomics,* 22**,** 856.

LEE, C. N., BAEK, K. S. & PARKHURST, A. 2016. The impact of hair coat color on longevity of Holstein cows in the tropics. *J Anim Sci Technol,* 58**,** 41.

LI, D., SUN, G., ZHANG, M., CAO, Y., ZHANG, C., FU, Y., LI, F., LI, G., JIANG, R., HAN, R., LI, Z., WANG, Y., TIAN, Y., LIU, X., LI, W. & KANG, X. 2020. Breeding history and candidate genes responsible for black skin of Xichuan black-bone chicken. *BMC Genomics,* 21**,** 511.

LYONS, L. A., FOE, I. T., RAH, H. C. & GRAHN, R. A. 2005. Chocolate coated cats: TYRP1 mutations for brown color in domestic cats. *Mamm Genome,* 16**,** 356-66.

MABROUK, O., SGHAIER, N., COSTA, R., GADDOUR, A., BENAMMAR ELGAAÏED, A. & DELGADO, J. V. 2010. The effect of non-genetic factors on the early body weights of Tunisian local goats. *Revista Brasileira De Zoologia - REV BRAS ZOOL,* 39.

MAHANJANA, A. & CRONJE, P. 2000. Factors affecting goat production in a communal farming system in the Eastern Cape region of South Africa. *South African Journal of Animal Science,* 30**,** 149-155.

MALDONADO-JÁQUEZ, J. A., TORRES-HERNÁNDEZ, G., BAUTISTA, Y., BAÉZ, P. & GRANADOS-RIVERA, L. D. 2023. Effect of coat and skin color variation on longevity, ethnological and functional indices in local goats. *Revista Científica de la Facultad de Ciencias Veterinarias,* 33**,** 1-7.

MDLADLA, K., DZOMBA, E. & MUCHADEYI, F. 2016. P5039 A landscape genomic approach to unravel the genomic mechanism of adaptation in indigenous goats of South Africa. *Journal of Animal Science,* 94**,** 134-135.

MIA, M. & MONDAL, A. 2020. A STUDY ON PRODUCTIVE AND REPRODUCTIVE PERFORMANCES OF BLACK BENGAL GOAT IN RELATION TO COAT COLORS.

MISBAH, F., BELAY, B. & HAILE, A. 2015. Participatory definition of trait preference and pastorals' indigenous knowledge on goat breeding strategy around Aysaita district, Ethiopia. *Livestock Research for Rural Development,* 27.
NAIR, M. R. R., SEJIAN, V., SILPA, M. V., FONSÊCA, V. F. C., DE MELO COSTA, C. C., DEVARAJ, C., KRISHNAN, G., BAGATH, M., NAMEER, P. O. & BHATTA, R. 2021. Goat as the ideal climate-resilient animal model in tropical environment: revisiting advantages over other livestock species. *Int J Biometeorol,* 65**,** 2229-2240.
NASTI, T. H. & TIMARES, L. 2015. MC1R, Eumelanin and Pheomelanin: Their Role in Determining the Susceptibility to Skin Cancer. *Photochemistry and Photobiology,* 91**,** 188-200.
NGULUMA, A. S., HYERA, E., NZIKU, Z., SHIRIMA, E. M., MASHINGO, M. S. H., LOBO, R. N. B., GETACHEW, T., RISCHKOWSKY, B. & HAILE, A. 2020. Characterization of the production system and breeding practices of indigenous goat keepers in Hai district, Northern Tanzania: implications for community-based breeding program. *Trop Anim Health Prod,* 52**,** 2955-2967.
ODUBOTE, K. 1994. Influence of qualitative traits on the performance of West African Dwarf Goats. *Nigeria Journal of Animal Production,* 21**,** 25-28.
OGET, C., SERVIN, B. & PALHIÈRE, I. 2019. Genetic diversity analysis of French goat populations reveals selective sweeps involved in their differentiation. *Animal Genetics,* 50**,** 54-63.
OJO, O., AKPA, G., ORUNMUYI, M. & ADEYINKA, I. 2015. Genetic Differentiation among Nigerian Indigenous Goat Populations. *Journal of Agricultural Science,* 7**,** 39-47.
ONASANYA, G., DE CAMPOS, J., ANWO, O., IKEOBI, C. & AMUSAN, S. 2018. *Impact of Coat Colour Genes and Coat types on zoometric Traits of African Dwarf Sheep bred under Traditional Extensive System*.
PENG, Y., WANG, Y., WANG, R., GENG, L., MA, R., ZHANG, C., LIU, Z., GONG, Y., LI, J. & LI, X. 2019. Exploring differentially expressed genes associated with coat color in goat skin using RNA-seq. *Canadian Journal of Animal Science,* 99**,** 357-366.
PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M. A., BENDER, D., MALLER, J., SKLAR, P., DE BAKKER, P. I. & DALY, M. J. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics,* 81**,** 559-575.
RAJAGANAPATHY, J., JAYASHANKAR, M., NAGARAJA, C., SHRIKRISHNA, I. & SATYANARAYANA, K. 2018. Genetic characterization of local goats of Karnataka by microsatellite marker analysis. *Indian Journal of Animal Research*.
RAMOS-MARTINEZ, J., SALINAS-GONZÁLEZ, H., MEDINA-ELIZONDO, M., FIGUEROA-VIRAMONTES, U. & MALDONADO-JÁQUEZ, J. 2021. Organization and grouping as a key axis to design attention schemes for goat farmers in northern Mexico: Case study. *Abanico veterinario,* 10**,** 1-19.
R CORE TEAM. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
REN, J., MAO, H., ZHANG, Z., XIAO, S., DING, N. & HUANG, L. 2011. A 6-bp deletion in the TYRP1 gene causes the brown colouration phenotype in Chinese indigenous pigs. *Heredity (Edinb),* 106**,** 862-8.
ROBINSON, S. W., DINULESCU, D. M. & CONE, R. D. 2000. GENETIC MODELS OF OBESITY AND ENERGY BALANCE IN THE MOUSE. *Annual Review of Genetics,* 34**,** 687-745.
ROSERO ALPALA, J. A., RANGEL GARCIA, W. D., ROJAS BARRETO, A. & ORLANDO BURGOS-PAZ, W. 2021. Contribution of genomic data in defining the breed composition of dual-purpose cattle. *Revista mexicana de ciencias pecuarias,* 12**,** 1008-1024.
RUMOSA GWAZE, F., CHIMONYO, M. & DZAMA, K. 2008. Communal goat production in Southern Africa: a review. *Tropical Animal Health and Production,* 41**,** 1157.
SCHMUTZ, S. M., BERRYERE, T. G. & DREGER, D. L. 2009. MITF and White Spotting in Dogs: A Population Study. *Journal of Heredity,* 100**,** S66-S74.
SCHMUTZ, S. M., BERRYERE, T. G. & GOLDFINCH, A. D. 2002. TYRP1 and MC1R genotypes and their effects on coat color in dogs. *Mamm Genome,* 13**,** 380-7.
SCHOUWEY, K. & BEERMANN, F. 2008. The Notch pathway: hair graying and pigment cell homeostasis. *Histol Histopathol,* 23**,** 609-19.
SEJIAN, V., SILPA, M. V., RESHMA NAIR, M. R., DEVARAJ, C., KRISHNAN, G., BAGATH, M., CHAUHAN, S. S., SUGANTHI, R. U., FONSECA, V. F. C., KÖNIG, S., GAUGHAN, J. B., DUNSHEA, F. R. & BHATTA, R. 2021. Heat Stress and Goat Welfare: Adaptation and Production Considerations. *Animals,* 11**,** 1021.
SEVANE, N., CORTÉS, O., GAMA, L. T., MARTÍNEZ, A., ZARAGOZA, P., AMILLS, M., BEDOTTI, D. O., BRUNO DE SOUSA, C., CAÑON, J., DUNNER, S., GINJA, C., LANARI, M. R., LANDI, V.,

SPONENBERG, P. & DELGADO, J. V. 2018. Dissection of ancestral genetic contributions to Creole goat populations. *Animal,* 12**,** 2017-2026.

TADESSE, D., URGE, M., ANIMUT, G. & MEKASHA, Y. 2014. Perceptions of households on purpose of keeping, trait preference, and production constraints for selected goat types in Ethiopia. *Tropical animal health and production,* 46**,** 363-370.

TORRES-HERNÁNDEZ, G., MALDONADO-JÁQUEZ, J. A., GRANADOS-RIVERA, L. D., WURZINGER, M. & CRUZ-TAMAYO, A. 2022. Creole goats in Latin America and the Caribbean: a priceless resource to ensure the well-being of rural communities. *International Journal of Agricultural Sustainability,* 20**,** 368-380.

TORRES-HERNÁNDEZ, G., VAQUERA HUERTA, H., MALDONADO-JÁQUEZ, J. A., SALINAS, H., PASTOR, F., GRANADOS-RIVERA, L. D., VAQUERA, H. & BAUTISTA, Y. 2020. Influence of milk production records on the estimation of typical and extended curve lactation parameters in local goats under grazing management. *The Indian journal of animal sciences,* 90**,** 643-646.

VAN BUREN, S. L., MINOR, K. M., GRAHN, R. A., MICKELSON, J. R., GRAHN, J. C., MALVICK, J., COLANGELO, J. R., MUELLER, E., KUEHNLEIN, P. & KEHL, A. 2020. A Third MLPH Variant Causing Coat Color Dilution in Dogs. *Genes (Basel),* 11.

XU, J., FU, Y., HU, Y., YIN, L., TANG, Z., YIN, D., ZHU, M., YU, M., LI, X., ZHOU, Y., ZHAO, S. & LIU, X. 2020. Whole genome variants across 57 pig breeds enable comprehensive identification of genetic signatures that underlie breed features. *Journal of Animal Science and Biotechnology,* 11**,** 115.

ZHAO, F., MCPARLAND, S., KEARNEY, F., DU, L. & BERRY, D. P. 2015. Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genetics Selection Evolution,* 47**,** 49.

Chapter 6: General Discussion and Conclusion

6.1 General discussion

The clustering of the commercial breeds revealed by the PCA and Admixture plots suggested close genetic relatedness between the breeds. In contrast to their commercial counterparts, the ecotypes' clustering suggested higher levels of genetic diversity in the village goats. Furthermore, genetic variation within the populations (91.44%) was higher than among populations (4.77%), and small average F-statistics detected in all breeds indicated small genetic differentiation between population likely due to interpopulation gene flow. While genetic differences between these populations were likely the result of village goat adaptation to differences in climate. These finding also suggested that the coat colour distributions within the ecotype populations is likely due to the goats' adaptation to the different climatic conditions in these regions, as well as breeder preferences.

Coat colour has important socio-cultural and economic value in veld goat communities, as a result, most farmers select their goats based on their coat colours and patterns (Rumosa Gwaze et al. 2008; Mahanjana & Cronje 2000). For instance, Mahanjana and Cronje (2000) found that white coloured goats are in higher demand in the Eastern Cape due to their perceived value in traditional/cultural ceremonies, resulting in selection preferences towards white goats. Therefore, the prevalence of specific coat colours at specific regions in the country is due to the goats' adaptation to the different climatic conditions in these regions, as well as the selective breeding practices of goat farmers which are informed by their socio-cultural and religious preferences.

Genome-wide Association studies carried out for all the coat colours (i.e., white, red, black, grey) and the coat colour patterns (i.e., white-body red head, speckled, patchy, belted, white-sided, blacklegged) under investigation in this study to reveal genes associated with the various coat colour phenotypes of South African meat-type goats. Results from the GWASs revealed several genes associated with the different coat colours including white (*CDK5*), red (*TLE6*), black (*CACNA2D1*), grey (*GSK3B*), and coat colour patterns including white-body red head (*CADPS2*, *SLC13A1*), speckled (*KIT*, *TYRP1*), patchy (*GNAI3*), belted (*TYRP1*), white-sided (*AHCY*), blacklegged (*AHCY*). The genes detected by the GWAS are involved in pathways that have been shown to influence coat pigmentation. For instance, *TYRP1* is involved in pathways including Tyrosine metabolism; Metabolic pathways; and Melanogenesis. *CACNA2D1* is involved in several pathways including MAPK signalling pathway. The *GSK3B* gene is

involved in several pathways including PI3K-Akt signalling pathway, Wnt signalling pathway, and Melanogenesis. The *GNAI3* gene is involved in several pathways including melanogenesis. The *KIT* gene is involved several pathways including MAPK signalling pathway, PI3K-Akt signalling pathway, and Melanogenesis. The *AHCY* gene is involved in Cysteine and methionine metabolism, and Metabolic pathways. *CDK5* is involved in several biological pathways and has been linked to lighter coat colour (Dong et al., 2017). *TLE6* is involved in the Wnt signalling pathway, and the Notch signalling pathway. These results suggest that coat colour is a polygenic trait which is influenced by multiple genes. However, due to the small sample sizes used in this study, the GWAS detection power was low and as such these results warrant further validation with larger sample sizes.

Copy number variation analysis revealed that copy number variation plays a role in goat coat colour variation similar to other livestock species such as cattle. The CNV analysis revealed more copy number losses then gains. Furthermore, 279 CNVRs overlapping 5 222 genes involved in several biological processes, molecular functions, cellular components, and pathways were found. Several coat colour genes overlapping these CNVRs included *GSK3B, Notch1, Notch2,* CDK5, *ADAMTS20, TYRO3, MAP2K1, ITCH, ASIP, AHCY, SLC45A3, EDNRA, ADCY2,* and *TYR.*, suggesting that copy number variation pays a role in coat colour in goats.5 genes were detected by both GWAS and CNV analysis (*TYR/TYRP1, GSK3B, CDK5, AHCY,* and *EDNRA*).

The Boer, Savannah and Kalahari Red breeds have been selected and reared under intensive commercial systems of production. Hence, it was expected that the detection of selection signatures within these populations would reveal genes related to economically important production traits. The *CDH12* gene found within the Boer breed, the gene has previously been linked to the regulation of skin colour in humans (Charfeddine et al.2020). The genes detected within the other two commercial breeds (i.e., the Savanna and Kalahari Red) were revealed to be associated with pathways that play roles in biological functions including goat metabolism (i.e., *BAAT, GRIN3A*), and immune response (i.e., *IKZF2, PPP2R2B, TLERG1, ARAP3, DIAPH1, RELL2*). As such, these results reflected that such traits are the likely targets of the intensive selection the goats undergo within commercial production systems. Therefore, the detection of genes involved in skin colour, metabolism, and immune response within these commercial breeds suggests that there is a strong desire for goats with improved environmental adaptation within the commercial sector. This would agree with reports that have suggested that the intense selection pressures present within commercial production systems have

previously resulted in higher degrees of inbreeding within the commercial goat populations which has limited the goats' environmental adaptation.

Candidate genes under selection in the village goats were found to be involved in pathways associated with melanin production, these include the MAPK signalling pathway (e.g., *DUSP16*, *KDR*, *FGF10*, *MAP3K7*), the PI3K-Akt pathway (*CDKN1B*, *COL6A3*, *CRK*, *CREB5*), the Wnt signalling pathway (*SERPINF1*, *WNT2*), the Notch signalling pathway (*HEY2*), and Melanogenesis (*TYRP1*) suggesting that coat colour is under selection in communal goat production systems. For instance, The MAPK signalling pathway has been reported to be involved in the goat coat pigmentation through the regulation of melanogenesis. The *CDKN1B* gene on chromosome 5, is involved in several pathways including the PI3K-Akt signalling pathway which has been linked to melanogenesis and the formation of coat pigment in various livestock species including cattle (D'Mello et al., 2016). Furthermore, the *TYRP1* gene on chromosome 8 found within the white village goats is involved in several pathways including Tyrosine metabolism, Metabolic pathways, as well as Melanogenesis. The gene has previously been linked to the coat colouration of various animals including pigs (Ren et al., 2011), dogs (Schmutz et al., 2002), rabbits (Jia et al., 2021), cattle (Berryere et al. 2003), cats (Lyons et al. 2005), and sheep (Gratten et al. 2007). In addition, the gene was detected in both the GWAS and CNVs analysis. Hence, it's detection in this study suggests that it may be implicated in goat coat colour as well.

The *KDR* gene found within red goats is involved in several pathways including the MAPK signalling pathway, and the PI3K-Akt signalling pathway, two pathways that have been associated with the regulation of coat colour variation in mammals. Furthermore, it has been revealed that the *KDR* gene forms a complex with *KIT*, and *PDGFRA* which together are responsible for the red coat colour pattern in Angus cattle (David et al., 2014). The gene has been revealed to also play a role in pig coat colour (Xu et al., 2020). The *FGF10* gene within black village goats is involved in several pathways that have been linked to coat colour formation including the MAPK signalling pathway, the Rap1 signalling pathway, as well as the PI3K-Akt signalling pathway. The gene has previously been found under selection in black Xichuan black-bone chickens (Li et al., 2020). Candidate genes under selection within grey coat coloured village goats included the *COL6A3* gene on chromosomes 3, and the *HEY2* gene on chromosome 9. The *COL6A3* gene is involved in PI3K-Akt signalling pathway, while the *HEY2* gene is involved in several pathways including the Notch signalling pathway. The Notch

signalling pathway has previously been linked to hair greying in mice (Schouwey and Beermann, 2008).

Furthermore, genes associated with various traits important for goat productivity and adaptation were detected. For instance, within commercial breeds, genes involved in metabolism (*BAAT, GRIN3A*), and immune response (*IKZF2, PPP2R2B, TLERG1, ARAP3*) were detected suggesting that artificial selection pressures within intensive production systems target goat productivity. While, within, the ecotype population several genes involved in metabolism (*FXN, FTO, IRX3, EDNRA*), reproduction (*PGR*), growth and development (*SMARCAD1*), immune response (*ACAD8, DLG4, GPS2*), and environmental adaptation (*CNR1*) were detected, suggesting that natural selection pressures within extensive communal production systems target goat adaptability to adverse environmental conditions.

Unlike the commercial breeds, village goat populations are constantly exposed to adverse climatic and environmental conditions. These adverse climatic and environmental effects impose natural selection pressures on the populations, resulting in the increase of the frequency of advantageous alleles linked to genes involved in goat adaptation. Therefore, it was expected that studying the selection signatures present within these populations would reveal the targets of natural selection pressures, in turn revealing genetic loci underlying local adaptation through natural selection. Such insights are crucial for guiding genetic improvement strategies and conservation programs for these populations. As expected, candidate genes under selection within these village goats with various coat colours and patterns were mostly involved in pathways associated with numerous traits including metabolism, reproduction, and immune response. These results reflect the effects of natural selection pressures resulting from climatic and environmental factors that have influenced the goats' adaptations to local environments and communal production systems which are usually exposed to harsh environmental conditions and scarce feed. These findings offer possible explanations for reports that have revealed that ecotype populations have a greater capacity for adaptation and higher efficiency in consumption and utilization of fodder, which has allowed them to develop, produce and reproduce in harsh environments. Furthermore, ecotypes have been revealed to be more resilient to higher temperatures, while possessing smaller corporal size, which proportionally produce less metabolic heat.

6.2 Conclusion

Results from this study revealed that goat coat colour is a polygenic trait which is influenced by multiple genes as well as copy number variation. In addition, the results showed that coat colour is under selection within South African meat-type village goats, furthermore, the detection of numerous genes involved in goat metabolism, reproduction, and immune response suggested a likely association between coat colour and other goat production traits. These finding agree with several previous studies which have demonstrated associations between coat colour and various goat production traits including environmental adaptation, reproduction, and milk and meat production, as well as studies that have shown that coat colour is influenced by multiple genes and CNVs within genomic regions overlapping these genes. Findings from this study will be useful in that they will aid the implementation of coat colour as a selection criterion in goat breeding programs in low-scale communal production systems.

6.3 Study limitations and considerations for the future

Although this study was able to detect several candidate genes that have been previously linked to coat colour, the identification and characterisation of the gene-gene interactions responsible for colour variations is still a major challenge and requires deeper insights for the effective implementation of genomic selection in livestock animals. A major challenge for the current study was in regard to the small sample sizes of the various coat colours and coat colour patterns investigated. Although 329 South African meat-type goats with various coat colours and patterns including plain coats such as white, red, black, grey, and coat colour patterns such as speckled, patchy, belted, white-sided, and black-legs were included in the study several other coat colour and coat colour patterns were considered for the ecotypes, however high variation of coat colour patterns (i.e., some animals represented complex coat colour patterns which were difficult to classify into a single phenotype) with even smaller sample sizes prevented their inclusion in this study. Thus, the samples collected were limited and they did not originate from a single homogeneous population. As a result, the final dataset had a small sample size comprised of different populations and breeds with different degrees of relatedness. As such GWAS detection power and precision was likely low. GWAS has been shown to produce more reliable/repeatable results with larger the sample sizes, as they reduce the likelihood of false positives. In this study, the GWAS accounted for possible false positive results by using GAPIT which has been shown to reduce the likelihood of false positives by accounting for multiple levels of relatedness through controlling for both population structure and unequal familial relationships using kinship matrix and by using the Bonferroni correction for multiple hypothesis testing to determine the threshold value for significant markers.

However, GWAS power has been proven to increase with larger sample sizes and higher quality of phenotypes. Low sample sizes and the inclusion of samples that do not originate from a homogeneous population increase the risk of differences in allele frequencies due to population stratification, cryptic relatedness, and confounding by nongenetic factors which may affected test statistics across the genome. In this regard, the small sample sizes in this study likely resulted in an underpowered GWAS that detected SNPs above the significance level (dotted line) but underestimated significant GWAS peaks. Therefore, although SNPs associated to genes underlying coat colour were found above the significance level (dotted line), the detection power and precision of the analysis was limited by the study's small sample size, as such the study's GWAS results need validation with larger population sizes to be able to draw overall conclusions on South African goat populations. Such validation studies may be favoured by collecting larger samples, running independent within-breed association analysis that incorporate principal components to consider ancestry information.

Furthermore, GWAS has been shown to be especially efficient when genetic determinism of traits that involve only a few genes but less adapted to highly polygenic traits such as coat colour which result from the small effects of several variants. In addition, the SNP array has the disadvantage of including only a small number of known SNPs, and marker distribution is biased (SNP ascertainment bias). Hence, only a few markers can be detected. To account for these limitations, GWAS based on genome-wide sequencing (WGS) is recommend for validation studies as WGS studies have been shown to overcome challenges such as uncovering the underlying biological mechanisms of several markers with small effects on a trait (i.e., extremely polygenic traits) by discovering rare variants of large effect for which causal mechanisms are generally easier to determine.

It is also noted that the SNP probes used to detect CNVs are neither dense enough nor uniformly distributed to achieve an unbiased high resolution CNV map. In this regard, future studies may consider using alternative techniques for CNV detection, such as next generation sequencing of individuals with various coat colours and coat colour patterns to achieve better coverage of goat CNVs. Finally, a portion of the goat genome is not annotated, and this could have influenced the results of the gene ontology analysis and the use of different quality control parameter, different CNV detection algorithms, and differences in sample sizes may contributed to differences in the CNVs detected between this study and previous studies in goats. Therefore, results from this study warrants further validations in the future.

6.4 References

BERRYERE, T. G., SCHMUTZ, S. M., SCHIMPF, R. J., COWAN, C. M. & POTTER, J. 2003. TYRP1 is associated with dun coat colour in Dexter cattle or how now brown cow? *Anim Genet,* 34**,** 169-75.

CHARFEDDINE, C., DALLALI, H., ABDESSALEM, G., GHEDIRA, K., HAMDI, Y., ELOUEJ, S., LANDOULSI, Z., DELAGUE, V., LAGARDE, A., LEVY, N., EL-AMRAOUI, A., BOUBAKER, M. S., ABDELHAK, S. & MOKNI, M. 2020. Identification of a CDH12 potential candidate genetic variant for an autosomal dominant form of transgrediens and progrediens palmoplantar keratoderma in a Tunisian family. *Journal of Human Genetics,* 65**,** 397-410.

D'MELLO, S. A., FINLAY, G. J., BAGULEY, B. C. & ASKARIAN-AMIRI, M. E. 2016. Signaling Pathways in Melanogenesis. *Int J Mol Sci,* 17.

DAVID, V. A., MENOTTI-RAYMOND, M., WALLACE, A. C., ROELKE, M., KEHLER, J., LEIGHTY, R., EIZIRIK, E., HANNAH, S. S., NELSON, G. & SCHÄFFER, A. A. 2014. Endogenous retrovirus insertion in the KIT oncogene determines white and white spotting in domestic cats. *G3: Genes, Genomes, Genetics,* 4**,** 1881-1891.

DONG, C., YANG, S., FAN, R., JI, K., ZHANG, J., LIU, X., HU, S., XIE, J., LIU, Y., GAO, W., WANG, H., YAO, J., SMITH, G. W. & HERRID, M. 2017. Functional Role of Cyclin-Dependent Kinase 5 in the Regulation of Melanogenesis and Epidermal Structure. *Scientific Reports,* 7**,** 13783.

GRATTEN, J., BERALDI, D., LOWDER, B. V., MCRAE, A. F., VISSCHER, P. M., PEMBERTON, J. M. & SLATE, J. 2007. Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proc Biol Sci,* 274**,** 619-26.

JIA, X., DING, P., CHEN, S., ZHAO, S., WANG, J. & LAI, S. 2021. Analysis of MC1R, MITF, TYR, TYRP1, and MLPH Genes Polymorphism in Four Rabbit Breeds with Different Coat Colors. *Animals (Basel),* 11.

LI, D., SUN, G., ZHANG, M., CAO, Y., ZHANG, C., FU, Y., LI, F., LI, G., JIANG, R., HAN, R., LI, Z., WANG, Y., TIAN, Y., LIU, X., LI, W. & KANG, X. 2020. Breeding history and candidate genes responsible for black skin of Xichuan black-bone chicken. *BMC Genomics,* 21**,** 511.

LYONS, L. A., FOE, I. T., RAH, H. C. & GRAHN, R. A. 2005. Chocolate coated cats: TYRP1 mutations for brown color in domestic cats. *Mamm Genome,* 16**,** 356-66.

MAHANJANA, A. & CRONJE, P. 2000. Factors affecting goat production in a communal farming system in the Eastern Cape region of South Africa. *South African Journal of Animal Science,* 30**,** 149-155.

REN, J., MAO, H., ZHANG, Z., XIAO, S., DING, N. & HUANG, L. 2011. A 6-bp deletion in the TYRP1 gene causes the brown colouration phenotype in Chinese indigenous pigs. *Heredity (Edinb),* 106**,** 862-8.

RUMOSA GWAZE, F., CHIMONYO, M. & DZAMA, K. 2008. Communal goat production in Southern Africa: a review. *Tropical Animal Health and Production,* 41**,** 1157.

SCHMUTZ, S. M., BERRYERE, T. G. & GOLDFINCH, A. D. 2002. TYRP1 and MC1R genotypes and their effects on coat color in dogs. *Mamm Genome,* 13**,** 380-7.

SCHOUWEY, K. & BEERMANN, F. 2008. The Notch pathway: hair graying and pigment cell homeostasis. *Histol Histopathol,* 23**,** 609-19.

XU, J., FU, Y., HU, Y., YIN, L., TANG, Z., YIN, D., ZHU, M., YU, M., LI, X., ZHOU, Y., ZHAO, S. & LIU, X. 2020. Whole genome variants across 57 pig breeds enable comprehensive identification of genetic signatures that underlie breed features. *Journal of Animal Science and Biotechnology,* 11**,** 115.