The Identification of Genomic Regions Associated with Tick Count, Growth Traits, and Skin Thickness in F_2 Angus x Nguni Cattle Using Genome-Wide Association Analysis

by
Nelisiwe Mkize

Dissertation presented for the degree of
Doctor of Philosophy (Animal Sciences) in the Faculty of
AgriSciences at Stellenbosch University

Supervisor: Prof. Kennedy Dzama
Co-supervisors: Prof. Ntanganedzeni Mapholi; Dr Bekezela Dube;
Prof. Azwihangwisi Maiwashe

December 2023

The financial assistance of the National Research Foundation (NRF) and the Agricultural Research Council (ARC) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to the NRF or ARC.

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2023

Dedication

This was an extremely hard journey for me without the support of my family. I dedicate this work to my late parents, siblings, and grandmother. Additionally, I also dedicate this work to myself, for not giving up when I had plenty of reasons to use as an excuse to quit.

Acknowledgement

Firstly, I would like to thank God for granting me the opportunity to be part of this research work and for carrying me through from start to finish.

I would also like to convey my sincere appreciation to:

All my supervisors (Prof N.O. Mapholi, Dr B. Dube, Prof K. Dzama and Prof A. Maiwashe), for taking me in, to be part of the marker detection research project, for proving data analysis scientific insights, for believing in me, for their patience, and for the never-ending teachings and words of encouragement.

The Agricultural Research Council, the Red Meat Research and Development Trust of South Africa, and the National Research Foundation for providing infrastructure and financial support.

The Animal Genetics team for helping with data collection, namely: Mr Lucas Pale, Matome Madilindi, Khanyisani Nxumalo, Ndivho Nemukondeni, Galatsang Hutang, Thendo Tshilate, and Michael Rampedi. Collecting data would have not been successful without your devoted contributions, most of you abandoned your own research work to help us, and I do not take that lightly. May God richly bless you for your generosity.

Dr O Matika for assisting with data analysis and scientific advice related to the research topic.

My dearest son Luthando Khumalo for his perseverance. Ever since I embarked on the journey of studying, our time spend as family has been compromised. I thank you nkosanayami for being understanding and respectful.

My dearest friend Galaletsang Hutang for all the support, laughter…. the list is endless. You make the world a better place.

All my friends, Sanelisiwe Zondi, Zamuxolo Kingstone Mkhize, Thembeka Banditi, Tapiwa Diamond Chadya, Njabulo Felix Msomi, Mbuyiselo Ndlovu, Philani Mnikathi, Thandanani Zondi, and the late Lungisani Busani and Bavelile Hlongwa. Ngiyabonga kakhulu for the love and support you gave me during this journey, kwande.

The Khumalo family for looking after my son.

Lastly, Nontokozo Sibahle Kunene for providing accommodation, food, and emotional support when our residence was burnt. Ngiyabonga kakhulu maMtimadze for coming through for me during a tough time in my life. That situation triggered my previous adverse experiences, but the tremendous support I received from you afforded me an opportunity to relocate myself and quickly refocus on my studies. God bless you.

Research outputs from this study

Publications in peer reviewed scientific journals:

Mkize N., Maiwashe A., Dzama K., Dube B. and Mapholi N. 2021. Suitability of GWAS as a Tool to Discover SNPs Associated with Tick Resistance in Cattle: A Review. *Pathogens*. 2021; 10(12):1604. https://doi.org/10.3390/pathogens10121604

Conference Contributions:

N. Mkize, A. Maiwashe, B. Dube, K. Dzama, and N.O. Mapholi. 2023. Genome-wide association study for tick count, growth traits, and skin thickness in F2 Angus x Nguni cattle. 8thAll Africa Conference on Agriculture, September 26-29, Botswana.

N. Mkize, A. Maiwashe, B. Dube, K. Dzama and N.O. Mapholi. 2023. Post-GWAS functional annotation for tick count, growth traits, and skin thickness in F2 Angus x Nguni crossbred cattle. The 39th International Society for Animal Genetics Conference, July 2-7, Cape Town, South Africa.

N. Mkize, B. Dube, A. Maiwashe, N.O. Mapholi, F. Muchadeyi, and K. Dzama. 2023. Post-GWAS functional annotation for tick count in F2 Angus x Nguni crossbred cattle. 30th Plant and Animal Genome XXX Conference, January 13-18, San Diego, USA.

N. Mkize, B. Dube, A. Maiwashe, N.O. Mapholi, F. Muchadeyi, and K. Dzama. 2021. Genetic parameter estimates for tick resistance in F2 Angus x Nguni cattle artificially infested with Amblyomma hebraeum. The 52nd SASAS Congress, virtual proceeding, 10–12 August 2021.

N. Mkize, G. B. Hutang, B. Dube, K. Dzama, A. Maiwashe and N.O. Mapholi. 2019. Preliminary results: Heritability estimate for tick count and identification of genomic regions associated with tick tolerance in F2 Angus x Nguni population. 37th International Society for Animal Genetics, July 7-12, Lleida, Spain.

N. Mkize, B. Dube, A. Maiwashe, , N.O. Mapholi, F. Muchadeyi, and K. Dzama. 2018. Identification of Genomic Regions Associated with Tick Resistance in F2 Angus x Nguni Cattle Artificially Infested with Amblyomma hebraeum: Preliminary Results. 27th Plant and Animal Genome XXVII Conference, January 12-16, San Diego, USA.

Summary

The aim of the study was to identify genetic determinants associated with tick count, growth traits, and skin thickness in F2 Angus x Nguni crossbred cattle. Two hundred and sixty-six animals were used in the study. Genetic parameters were estimated by fitting a sire model using pedigree records and by fitting an animal model using a kinship matrix in the ASReml software. Hair samples from the animals were genotyped using Illumina BovineSNP150 assay and there were 115 143 SNPs after quality control. A single SNP approach was used for the association tests, using the mixed model function of the mmscore, while adjusting for population stratification and relatedness in GenABEL. A nominal threshold of Pnominal = 5 x 10-5 was applied to the data. Post-GWAS functional annotation, pathway, and cluster analyses were performed using DAVID and ShinyGo bioinformatics resources with the *Bos taurus* genome used as a background. Fisher's exact test with false discovery rate (FDR) adjustment was used at p (FDR) < 0.05.

The heritability estimates from the sire model for tick count, birth, weaning and yearling weights, and skin thickness were 0.21±0.22, 0.36±0.26, 0.32±0.14, 0.32±0.19, and 0.40±0.15, respectively. The respective heritability estimates obtained from an animal model fitting kinship matrix were 0.12±0.15, 0.26±0.16, 0.70±0.18, 0.38±0.18 and 0.58±0.21. These heritability estimates showed the presence of genetic variation for these traits. Thus, improvement of these traits can be expected from selection. Single nucleotide polymorphisms (SNP) for tick count were observed on chromosomes 2 and 4, while those associated with birth weight were on chromosomes 3 and 12. Chromosomes 2 and 18 contained SNPs for weaning weight and chromosome 8 had a SNP for yearling weight. The SNPs for skin thickness were observed on chromosomes 2, 7, 8, 9, and 12. For tick count, functional annotation uncovered enriched gene ontology (GO) terms and pathways related to cellular processes, regulation of biological processes, and response to stimulus. These were relevant for immune system response necessary for understanding host-tick interaction. Candidate genes such as *ZNF746, GIMAP8,* and *RARRES2* could be potential biomarkers for tick control in cattle. The enriched GO terms and pathways for growth traits were related to ubiquinol-cytochrome-c reductase activity, plasma membrane raft, DNA ligase, ATP dependent among others. These categories were relatable to cell differentiation, skeletal muscle development, and metabolism regulation. The uncovered potential candidate genes for growth traits were *TRPM8, SPP2, UQCRFS1, MB, TMEFF1, CAVIN4, and MSANTD3*. For skin thickness, the enriched GO terms and pathways were related to immune response, cell differentiation, and transmembrane receptor tyrosine-protein kinase among others. The categories were enriched

through *FER* and *NTRK2* genes, which could be regarded as potential candidate genes for skin thickness in this population. The study uncovered some novel and previously identified genes as well as biological mechanisms related to tick count, growth traits, and skin thickness in the F_2 Angus x Nguni population. These genes could be used to facilitate genetic improvement of tick count, growth traits, and skin thickness in this population.

Preface

This thesis is an original work by Nelisiwe Mkize. This study forms part of a marker detection research project, which was established by the Animal Production Campus of the Agricultural Research Council. The research project was based on F_2 Angus x Nguni cattle population and the Agricultural Research Council Ethics Committee granted the ethical clearance **(Ref: APIEC 17/04).**

This thesis is structured as described below:

Chapter 1: General Introduction

This is an introductory chapter, which gives background information of the research undertaken, the problem statement, motivation, research question and aims and objectives of the study.

Chapter 2: Suitability of GWAS as a Tool to Discover SNPs Associated with Tick Resistance in Cattle: A Review

This chapter reviewed available literature to explore whether GWAS was a suitable tool to study tick resistance in cattle. To be specific, the study looked at computer software available to conduct GWAS, available genotyping platforms, testing for association, Post-GWAS analysis.

Chapter 3: Genetic parameter estimates for tick resistance in F_2 Angus x Nguni cattle artificially infested with Amblyomma hebraeum ticks

The focus of this chapter was on the estimation of genetic parameters for tick count, growth traits and skin thickness using a sire model through ASReml software.

Chapter 4: Genome-wide association study for tick count, growth traits and skin thickness in F_2 Angus x Nguni cattle

Genome wide association analysis was used to search for SNPs associated with tick count, growth traits, and skin thickness in F_2 Angus x Nguni cattle

Chapter 5: Post-GWAS analysis to find genes associated with tick count, growth traits and skin thickness

This chapter presents the Post-GWAS analysis aimed to understand the biological relevance of the identified SNPs with the traits of interest.

Chapter 6: General conclusion and recommendations

This chapter presents the general summary of the findings, scientific contribution, recommendations, future work and limitations of the study.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

SNP	Single Nucleotide Polymorphism
QTL	Quantitative Trait Loci
ARC-AP	Agricultural Research Council-Animal Production
GWAS	Genome-Wide Association Studies
GS	Genetic Selection
F_1	First Filial Generation
F_2	Second Filial Generation
TTBDs	Ticks and Tick-Borne Diseases
LD	Linkage Disequilibrium
GEMMA	Genome-Wide Efficient Mixed Model Association
MLM	Mixed Linear Model
BSLM	Bayesian Sparse Linear Mixed Model
IBD	Identity by Descent
SVS	SNP and Variation Suite
NCBI	National Center for Biotechnology Information
EMBL-EBI	European Molecular Biology Laboratory-European Bioinformatics Institute
BTA	Bos *taurus* Chromosome
UCSC	University of California Santa Cruz
QTLdb	Quantitative Trait Loci Database
VEGA	Vertebra Genome Annotation
Refseq	Reference Sequence
FDR	False Discover Rate
GLM	General Linear Model
MAF	Minor Allele Frequency
HWE	Hardy Weinberg Equilibrium
QC	Quality Control
GxE	Genotype-Environment Interaction
GIGI	Genotype Imputation Given Inheritance
mm	Milimeters
Kg	Kilograms
BWT	Birth Weight
WWT	Weaning Weight
YWT	Yearling Weight
TCOUNT	Tick Count

Y_INF	Year of Infestation
CC	Coat Colour
Chr	Chromosome
STHICK	Skin Thickness
h^2	Narrow Sense Heritability
MAS	Marker Assisted Selection
DNA	Deoxyribonucleic Acid
PC	Principal Component
MSD	Multidimensional Scaling
q-q	Quantile-Quantile
DAVID	The Database for Annotation, Visualization and Integrated Discovery

1. Chapter 1: General Introduction

1.1 Introduction

Cattle production contributes significantly to the economies of many countries and most people depend on it for food security and income [1–4]. Cattle food products have a nutritive role in the human diet; they contribute beneficial sources of proteins, vitamins, fibre, fatty acids, minerals, and energy [5–7]. These are valuable nutritive elements, which are important for human growth and development [8]. The rapid expansion of the world population contributes to the great demand for food and the majority of the world's population resides in hunger-stricken communities [9]. The United Nations have established 17 global Sustainable Development Goals. Sustainable Development Goal 2 aims to create a world that is free of hunger by the year 2030 [10]. The actual focus of the goal is to end hunger, achieve food security and improved nutrition, and promote sustainable agriculture [10]. The repercussions of continuous climate change are a threat to the United Nations' Sustainable Development Goal 2. Climate change creates environmental conditions, which threaten cattle production systems. Ultimately contributing to the instabilities in animal welfare and food insecurity.

About 70% of global cattle production occurs in subtropical regions, which have a high prevalence of ticks and tick-borne diseases [11]. Ticks and tick-borne diseases (TTBDs) affect about 80% of the world's cattle population [12]. As a result, ticks are of veterinary and public health importance globally and are associated with the transmission of various diseases (e.g. Bovine babesiosis, anaplasmosis, ehrlichiosis, Lyme disease, Rocky Mountain Spotted Fever, and others) in animals and humans [13,14].

The widespread presence of TTBDs has dramatic effects on cattle production traits, morbidity, and mortality [12,15,16]. Losses due to ticks have long been of concern to livestock producers, governments, and researchers [17]. Currently, global losses range from US$20-30 billion per annum [18] and this is significantly higher than previous losses of US$13.9-18.7 billion per annum that were reported by de Castro in 1997 [13]. This shows that presence of TTBDs in livestock production is an ongoing problem, which requires serious attention. In Africa, around 1.1 million cattle die annually because of TTBDs, resulting in economic losses of $160 million [19]. In India, the economic losses due to TTBDs in animal production are estimated to be US$498.7 million per annum [20]. In Brazil, previous reports suggest that the country loses about two billion dollars per annum because of ticks [21]. These losses are incurred despite

the widespread use of tick control measures, which include the application of acaricides and vaccines.

In Africa, the tick species that have been reported to be troublesome in cattle production, include *Amblyomma spp and Rhipicephalus spp* [22]. Studies investigating host susceptibility/resistance to tick infestations in South African cattle have reported different prevalence rates on different breeds such as Nguni, Bonsmara, Angus and others [16, 23–26]. These studies have highlighted that South Africa is one of the countries facing a tick-endemic crisis in cattle production. In the regions where ticks are endemic, different tick control measures such as the use of chemical compounds, biological systems and grazing management have been adopted to mitigate the burden of tick infestations in cattle production [17].

The success of different tick control methods employed to mitigate the burden of TTBDs in cattle production has been hindered by various shortcomings. Currently, the main tick control method in use is based on chemical compounds known as acaricides [27,28]. The application of acaricides is associated with high costs, the emergence of acaricide-resistant tick strains, the presence of acaricide residues in animal products and environmental contamination [29,30]. The emergence of acaricide-resistant strains of ticks underlines the need to continuously develop new chemicals. Given that acaricide-resistant strains are evolving faster than the development of new acaricides [31], this approach may not be a stand-alone solution. Apart from the acaricides usage, vaccines are currently being exploited as alternative anti-tick compounds. However, it has been reported that ticks mutate the targeted epitopes into unfamiliar forms, and this nullifies the effect of these drugs [32,33]. Moreover, the natural specificity of vaccines disadvantages the ability of the drugs to effectively confer protection against various tick species [34]. Therefore, there is a need for the development of alternative tick control strategies.

Host resistance to bovine tick species is the natural ability of an animal to limit the number of ticks that survive to maturity; this occurs because of the immunological response of the animal [35]. Differences in tick burdens between and within cattle breeds have been observed [16,26]. It is known that *Bos indicus* cattle and their crosses are more resistant to TTBDs compared to *Bos taurus* cattle [36,37]. This is explained by the genetic architecture of the breeds, which is currently being exploited globally, to gain insight into the biological mechanisms influencing the resistance and susceptibility of cattle against ticks [26,30,38–43]. The identification of genetic variants underlying resistance to ticks in cattle has the potential to provide

opportunities to improve tick resistance through genomic selection [44,45]. The exploratory measures to search for genetic determinants associated with tick resistance are currently being made possible by the availability and advancement of high-throughput single nucleotide polymorphisms (SNPs) genotyping arrays for cattle [46]. The SNP genotyping arrays generate information, which is studied using computational and statistical methods, to attain insight into the genetic variation among cattle, translating to the performance of economic traits [47,48]. When studying genetic variations among economically important traits, different genomic approaches are used. The approach that is gaining momentum in research is the genome-wide association studies (GWASs). The simple principle for this approach is to test the SNPs influencing the trait of interest by associating the trait's phenotype with genomic data of the species tested. In cattle tick resistance research, the GWAS approach has been used in different regions (such as Brazil, Australia, and South Africa) on different breeds to investigate genomic regions associated with resistance to ticks. Previous GWAS research identified genomic regions associated with resistance to ticks in cattle [26,30,38,43]. Despite the findings reported in these studies, information on molecular markers associated with tick resistance in cattle is still limited. Therefore, this elicits the need to conduct more studies to uncover biological mechanisms underlying tick resistance in cattle. More studies will provide knowledge, which will facilitate the development of alternative tick-control strategies through selective breeding for tick resistance. Breeding for host resistance to ticks may reduce the costs of vaccine and acaricides administration as well as reduce cattle mortalities. This would also improve animal welfare. The environment and consumers stand to benefit through reduced contamination of animal products and the environment with chemical residues.

Body weight is considered an important economic trait that describes growth in beef cattle production [49,50]. At the farm level, growth traits (body weights) are used as an indicator of meat production and as a criterion to select animals for further breeding [51–54]. One of the breeding goals in beef breeding programs is to improve body weight (e.g. birth weight, weaning weight and yearling weight). The improvement of body weight is beneficial for controlling calving ease (with birth weight), preventing stock fatalities, increasing productivity, and ensuring food security. Selling et al. [55] noted that the improvement of growth traits can be enhanced if Deoxyribonucleic acid (DNA) polymorphisms influencing the expression of the traits are uncovered and used for selective breeding. In this context, the improvement of growth traits could act as a proxy to improve tick resistance in cattle, since positive correlations have been established between these traits [44]. Furthermore, skin thickness is another economic trait that has been noted to play a role in tick infestation in cattle [56]. The thickness of the skin has also been suspected to have an association with growth traits [57]. However,

the information related to the relationship between growth traits and skin thickness is still limited.

Commercial beef production in South Africa is practiced mainly using exotic breeds such as Angus, Simmental, Santa Gertrudis, and Wagyu [58]. Angus cattle are one of the preferred breeds for meat production, because of their ease of calving, marbling, and growth traits [59]. However, Muchenje et al. [16] reported that the Angus breed tends to be susceptible to TTBDs that affect meat production and reduce profit margins. Previous studies have reported that there are cattle breeds with natural resilience to tick infestations [23,26,60,61]. South African indigenous breeds that have been reported to be highly resistant to tick infestations include the Afrikaner and Nguni [23,61–64]. The Nguni breed has captured international interest concerning its resilience to TTBDs, high reproductive performance, good gait, minimal maintenance requirements, and good foraging ability [16,65]. The demand for new tick control measures in cattle production has led to the implementation of studies that are focused on the crossbreeding of naturally tick-resistant breeds with breeds known for producing high-quality meat but are susceptible to tick infestations. In Southern Africa, crossbreeding to improve tick resistance has not been well explored. This warrants the need to implement crossbreeding research aimed at improving tick resistance using the Nguni and Angus breeds.

1.2 Problem statement

Ticks and tick-borne diseases (TTBDs) threaten animal welfare, food security, food and environmental safety, and profit margins of pasture-based cattle production systems. TTBDs are prevalent in cattle production systems globally and are expected to increase due to climate change. Farm management systems make use of acaricides to control TTBDs, however, these drugs are costly, and their residues pollute the environment and food products of cattle origin. A contaminated environment raises public health concerns and underscores the need for remedial actions to be taken swiftly. Furthermore, the prolonged use of acaricides in production systems results in the emergence of tick strains that are resistant to these drugs. Therefore, there is a need for alternative tick control strategies that will not only be cost-effective, but sustainable and environmentally friendly, complementing conventional tick control strategies. The use of genomic information in breeding to improve host natural resistance to ticks has been recommended as an environmentally safe alternative strategy to control cattle ticks.

In South Africa, beef production systems are mostly practiced under extensive environments

characterized by elevated levels of infestation by ticks. Commercial beef production is prevalently practiced using exotic breeds. Not only are exotic breeds known to be highly susceptible to tick infestations but also they are less adaptable to harsh environmental conditions such as drought. During unfavourable environmental conditions, commercial cattle production is heavily affected.

The variation in resistance to ticks that have been observed in South African cattle breeds has not been well explored as a potential approach to control ticks. Thus, the Nguni cattle's superior resistance to ticks has not been utilized in improving the level of resistance in South African beef production systems.

There is evidence of additive variation that has been linked to tick count, skin thickness and growth traits. The genetic improvement of these traits requires knowledge of the genetic architecture of the determinants which influence the trait. Genome-wide association analysis presents a better approach to elucidate the genetic architecture of biological mechanisms influencing tick count, growth traits and skin thickness.

1.3 Motivation of the study

The establishment of South African research studies aimed to uncover the genetic determinants associated with tick count, growth traits and skin thickness will provide information, which will inform cattle breeding programs. Selection for tick count, growth traits and skin thickness, may help control tick infestations, reduce production losses and the intensive use of acaricides, as well as increase food and environmental safety from acaricides contaminations.

1.4 Aims and objectives

The aim of this study was to identify genetic determinants associated with tick count, growth traits, and skin thickness in F_2 Angus x Nguni crossbred cattle.

The specific objectives of the study were:

(i) To estimate genetic parameters for tick count, growth traits and skin thickness in F_2 Angus x Nguni cattle.

(ii) To identify SNPs associated with tick count, growth traits and skin thickness in F_2

Angus x Nguni cattle using Genome Wide Association analysis.

(iii) To conduct post-genome wide association analyses to identify genes associated with biological processes relevant to tick count, growth traits and skin thickness.

1.5 Research question

The study aimed to address the following research questions:

(iv) Does genetic variation exist for tick count, growth traits and skin thickness in F_2 Angus x Nguni cattle?

(v) Are there any genetic determinants associated with the expression of tick count, growth traits, and skin thickness in F_2 Angus x Nguni cattle?

(vi) What are the biological functions of the genetic determinants underlying tick count, growth traits and skin thickness?

1.6 References

[1] Upton M. The Role of Livestock in Economic Development and Poverty Reduction n.d.:66.

[2] Abu Hatab A, Cavinato MER, Lagerkvist CJ. Urbanization, livestock systems and food security in developing countries: A systematic review of the literature. Food Secur 2019;11:279–99. https://doi.org/10.1007/s12571-019-00906-1.

[3] Thornton PK. Livestock production: recent trends, future prospects. Philos Trans R Soc B Biol Sci 2010;365:2853–67. https://doi.org/10.1098/rstb.2010.0134.

[4] Ogunkoya FT. Socio-economic factors that affect livestock numbers: a case study of smallholder cattle and sheep farmers in the Free State province of South Africa n.d.:119.

[5] Schönfeldt HC, Pretorius B, Hall N. The impact of animal source food products on human nutrition and health. South Afr J Anim Sci 2013;43:394–412. https://doi.org/10.4314/sajas.v43i3.11.

[6] Beef 101: Nutrition Facts and Health Effects. Healthline 2019. https://www.healthline.com/nutrition/foods/beef (accessed November 24, 2022).

[7] Buchanan JW, Reecy JM, Garrick DJ, Duan Q, Beitz DC, Koltes JE, et al. Deriving Gene Networks from SNP Associated with Triacylglycerol and Phospholipid Fatty

Acid Fractions from Ribeyes of Angus Cattle. Front Genet 2016;7:116. https://doi.org/10.3389/fgene.2016.00116.

[8] Fite MB, Tura AK, Yadeta TA, Oljira L, Roba KT. Consumption of animal source food and associated factors among pregnant women in eastern Ethiopia: A community-based study. PLoS ONE 2022;17:e0270250. https://doi.org/10.1371/journal.pone.0270250.

[9] Delgado CL, Rosegrant MW, Steinfeld H, Ehui SK, Courbois C. Livestock to 2020: The next food revolution. International Food Policy Research Institute; 1999.

[10] Martin. Goal 2: Zero Hunger. U N Sustain Dev n.d. https://www.un.org/sustainabledevelopment/hunger/ (accessed November 24, 2022).

[11] Porto Neto LR, Jonsson NN, D'Occhio MJ, Barendse W. Molecular genetic approaches for identifying the basis of variation in resistance to tick infestation in cattle. Vet Parasitol 2011;180:165–72. https://doi.org/10.1016/j.vetpar.2011.05.048.

[12] World Livestock 2011. Livestock in food security |Policy Support and Governance| Food and Agriculture Organization of the United Nations n.d. https://www.fao.org/policy-support/tools-and-publications/resources-details/en/c/1262785/ (accessed August 1, 2022).

[13] de Castro JJ. Sustainable tick and tickborne disease control in livestock improvement in developing countries. Vet Parasitol 1997;71:77–97. https://doi.org/10.1016/s0304-4017(97)00033-2.

[14] Jongejan F, Uilenberg G. The global importance of ticks. Parasitology 2004;129 Suppl:S3-14. https://doi.org/10.1017/s0031182004005967.

[15] Rajput ZI, Hu S, Chen W, Arijo AG, Xiao C. Importance of ticks and their chemical and immunological control in livestock. J Zhejiang Univ Sci B 2006;7:912–21. https://doi.org/10.1631/jzus.2006.B0912.

[16] Muchenje V, Dzama K, Chimonyo M, Raats JG, Strydom PE. Tick susceptibility and its effects on growth performance and carcass characteristics of Nguni, Bonsmara and Angus steers raised on natural pasture. Anim Int J Anim Biosci 2008;2:298–304. https://doi.org/10.1017/S1751731107001036.

[17] Mapholi NO, Marufu MC, Maiwashe A, Banga CB, Muchenje V, MacNeil MD, et al. Towards a genomics approach to tick (Acari: Ixodidae) control in cattle: a review. Ticks Tick-Borne Dis 2014;5:475–83. https://doi.org/10.1016/j.ttbdis.2014.04.006.

[18] Lew-Tabor AE, Rodriguez Valle M. A review of reverse vaccinology approaches for the development of vaccines against ticks and tick borne diseases. Ticks Tick-Borne Dis 2016;7:573–85. https://doi.org/10.1016/j.ttbdis.2015.12.012.

[19] Olwoch JM, Reyers B, Engelbrecht FA, Erasmus BFN. Climate change and the tick-borne disease, Theileriosis (East Coast fever) in sub-Saharan Africa. J Arid Environ 2008;72:108–20. https://doi.org/10.1016/j.jaridenv.2007.04.003.

[20] Singh U, Deb R, Alyethodi RR, Alex R, Kumar S, Chakraborty S, et al. Molecular markers and their applications in cattle genetic research: A review. Biomark Genomic Med 2014;6:49–58. https://doi.org/10.1016/j.bgm.2014.03.001.

[21] Grisi L, Leite RC, Martins JR de S, Barros ATM de, Andreotti R, Cançado PHD, et al. Reassessment of the potential economic impact of cattle parasites in Brazil. Rev Bras Parasitol Veterinária 2014;23:150–6. https://doi.org/10.1590/S1984-29612014042.

[22] Días TS. Some data concerning the ticks (Acarina-Ixodoidea) presently known in Mozambique. Undefined 1991.

[23] Marufu MC. Prevalence of ticks and tick-borne diseases in cattle on communal Rangelands in the highland areas of the Eastern Cape province, South Africa. Thesis. University of Fort Hare, 2008.

[24] Nyangiwe N, Goni S, Hervé-Claude LP, Ruddat I, Horak IG. Ticks on pastures and on two breeds of cattle in the Eastern Cape province, South Africa. Onderstepoort J Vet Res 2011;78:320. https://doi.org/10.4102/ojvr.v78i1.320.

[25] Spickett AM, Heyne IH, Williams R. Survey of the livestock ticks of the North West province, South Africa. Onderstepoort J Vet Res 2011;78:1–12.

[26] Mapholi NO, Maiwashe A, Matika O, Riggio V, Bishop SC, MacNeil MD, et al. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks Tick-Borne Dis 2016;7:487–97. https://doi.org/10.1016/j.ttbdis.2016.02.005.

[27] Shyma KP, Gupta JP, Singh V. Breeding strategies for tick resistance in tropical cattle: a sustainable approach for tick control. vol. 39. 2013. https://doi.org/10.1007/s12639-013-0294-5.

[28] Nejash AA. Review of Economically Important Cattle Tick and Its Control in Ethiopia. Vector Biol J 2016;1. https://doi.org/10.4172/2473-4810.1000104.

[29] Castro-Janer E, Martins JR, Mendes MC, Namindome A, Klafke GM, Schumaker TTS. Diagnoses of fipronil resistance in Brazilian cattle ticks (Rhipicephalus (Boophilus) microplus) using in vitro larval bioassays. Vet Parasitol 2010;173:300–6. https://doi.org/10.1016/j.vetpar.2010.06.036.

[30] Machado MA, S Azevedo AL, Teodoro RL, Pires MA, CD Peixoto MG, de Freitas C, et al. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (Bos taurus × Bos indicus). BMC Genomics 2010;11:280. https://doi.org/10.1186/1471-2164-11-280.

[31] Kemp DH, McKenna RV, Thullner R (1999) Strategies for tick control in a world of acaricide resistance. In: Proceedings of the IV Seminario Internacional de Parasitologia Animal, CONASAGA, Puerto Vallarta, Mexico, pp 1–10 - Google Searchn.d.https://www.google.co.za/search?tbm=bks&hl=en&q=Kemp+DH%2C+McKenna+RV%2C+Thullner+R+%281999%29+Strategies+for+tick+control+in+a+world+of+acaricide+resistance.+In%3A+Proceedings+of+the+IV+Seminario+Internacional+de+Parasitologia+Animal%2C+CONASAGA%2C+Puerto+Vallarta%2C+Mexico%2C+pp+1%E2%80%9310 (accessed August 3, 2022).

[32] Bowman AS, Nuttall PA. Ticks: Biology, Disease and Control. Cambridge University Press; 2008.

[33] Medicine I of, Practice B on PH and PH, Science C on LD and OT-BDTS of the. Critical Needs and Gaps in Understanding Prevention, Amelioration, and Resolution of Lyme and Other Tick-Borne Diseases: The Short-Term and Long-Term Outcomes: Workshop Report. National Academies Press; 2011.

[34] de la Fuente J, Contreras M. Tick vaccines: current status and future directions. Expert Rev Vaccines 2015;14:1367–76. https://doi.org/10.1586/14760584.2015.1076339.

[35] Utech KBW, Seifert GW, Wharton RH. Breeding Australian Illawarra shorthorn cattle for resistance to Boophilus microplus. 1. Factors affecting resistance 1978.

[36] Piper EK, Jackson LA, Bielefeldt-Ohmann H, Gondro C, Lew-Tabor AE, Jonsson NN. Tick-susceptible Bos taurus cattle display an increased cellular response at the site of larval Rhipicephalus (Boophilus) microplus attachment, compared with tick-

resistant Bos indicus cattle. Int J Parasitol 2010;40:431–41. https://doi.org/10.1016/j.ijpara.2009.09.009.

[37] Tabor AE, Ali A, Rehman G, Rocha Garcia G, Zangirolamo AF, Malardo T, et al. Cattle Tick Rhipicephalus microplus-Host Interface: A Review of Resistant and Susceptible Host Responses. Front Cell Infect Microbiol 2017;7.

[38] Otto PI, Guimarães SEF, Verardo LL, Azevedo ALS, Vandenplas J, Soares ACC, et al. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J Dairy Sci 2018;101:11020–32. https://doi.org/10.3168/jds.2017-14223.

[39] Gasparin G, Miyata M, Coutinho LL, Martinez ML, Teodoro RL, Furlong J, et al. Mapping of quantitative trait loci controlling tick [Riphicephalus (Boophilus) microplus] resistance on bovine chromosomes 5, 7 and 14. Anim Genet 2007;38:453–9. https://doi.org/10.1111/j.1365-2052.2007.01634.x.

[40] Machado, M.A, Azevedo, A.L., Gasparin, K. Fine mapping of BTA 10 and BTA 11 for tick resistance in Gir 9 Holstein cross. In: 9th World Congress of Genetics Applied to Livestock Production, Leipzig, Germany, 2010b, p. 607.

[41] Martinez ML, Machado MA, Nascimento CS, Silva MVGB, Teodoro RL, Furlong J, et al. Association of BoLA-DRB3.2 alleles with tick (Boophilus microplus) resistance in cattle. Genet Mol Res GMR 2006;5:513–24.

[42] Marima JK, Nel CL, Marufu MC, Jonsson NN, Dube B, Dzama K. A genetic and immunological comparison of tick-resistance in beef cattle following artificial infestation with Rhipicephalus ticks 2020. https://doi.org/10.1007/s10493-020-00480-8.

[43] Sollero BP, Junqueira VS, Gomes CCG, Caetano AR, Cardoso FF. Tag SNP selection for prediction of tick resistance in Brazilian Braford and Hereford cattle breeds using Bayesian methods. Genet Sel Evol 2017;49:49. https://doi.org/10.1186/s12711-017-0325-2.

[44] Frisch JE. Towards a permanent solution for controlling cattle ticks. Int J Parasitol 1999;29:57–71; discussion 73-75. https://doi.org/10.1016/s0020-7519(98)00177-5.

[45] Porto Neto LR, Bunch RJ, Harrison BE, Prayaga KC, Barendse W. Haplotypes that include the integrin alpha 11 gene are associated with tick burden in cattle. BMC Genet 2010;11:55. https://doi.org/10.1186/1471-2156-11-55.

[46] Dekkers JCM. Application of Genomics Tools to Animal Breeding. Curr Genomics 2012;13:207–12. https://doi.org/10.2174/138920212800543057.

[47] Koopaee HK, Koshkoiyeh AE. SNPs genotyping technologies and their applications in farm animals breeding programs: review. Braz Arch Biol Technol 2014;57:87–95. https://doi.org/10.1590/S1516-89132014000100013.

[48] Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, et al. Development and Characterization of a High Density SNP Genotyping Assay for Cattle. PLOS ONE 2009;4:e5350. https://doi.org/10.1371/journal.pone.0005350.

[49] Martínez R, Bejarano D, Gómez Y, Dasoneville R, Jiménez A, Even G, et al. Genome-wide association study for birth, weaning and yearling weight in Colombian Brahman cattle. Genet Mol Biol 2017;40:453–9. https://doi.org/10.1590/1678-4685-GMB-2016-0017.

[50] Du L, Duan X, An B, Chang T, Liang M, Xu L, et al. Genome-Wide Association Study Based on Random Regression Model Reveals Candidate Genes Associated with Longitudinal Data in Chinese Simmental Beef Cattle. Anim Open Access J MDPI 2021;11:2524. https://doi.org/10.3390/ani11092524.

[51] Gomes Freire Guidolin D, Buzanskas M, Ramos S, Venturini G, Munari D, Oliveira J, et al. Genotype - environment interaction for post-weaning traits in Nellore beef cattle. Anim Prod Sci 2012;52:975–80. https://doi.org/10.1071/AN11037.

[52] Barwick SA, Henzell AL. Development successes and issues for the future in deriving and applying selection indexes for beef breeding. Aust J Exp Agric 2005.

[53] Cook BR, Tess MW, Kress DD. Effects of selection strategies using heifer pelvic area and sire birth weight expected progeny difference on dystocia in first-calf heifers. J Anim Sci 1993;71:602–7. https://doi.org/10.2527/1993.713602x.

[54] Boligon AA, Silva J a. V, Sesana RC, Sesana JC, Junqueira JB, Albuquerque LG. Estimation of genetic parameters for body weights, scrotal circumference, and testicular volume measured at different ages in Nellore cattle. J Anim Sci 2010;88:1215–9. https://doi.org/10.2527/jas.2008-1719.

[55] Snelling WM, Allan MF, Keele JW, Kuehn LA, McDaneld T, Smith TPL, et al. Genome-wide association study of growth in crossbred beef cattle. J Anim Sci 2010;88:837–48. https://doi.org/10.2527/jas.2009-2257.

[56] Fourie PJ, Foster LA, Neser FWC. Score and hide-thickness, together with tick burden and body condition score, in four cattle breeds in the South-eastern Free State province of South Africa n.d.;11:8.

[57] Maiorano AM, Oliveira MCS, Cyrillo JNSG, Albuquerque LG, Curi RA, Silva JAI. Genetic study of skin thickness and its association with postweaning growth in Nellore cattle: estimation of the genetic parameters. Genet Mol Res GMR 2016;15. https://doi.org/10.4238/gmr.15017124.

[58] South African Beef Breeds n.d. https://southafrica.co.za/south-african-beef-breeds.html (accessed August 3, 2022).

[59] Andersen HJ, Oksbjerg N, Young JF, Therkildsen M. Feeding and meat quality - a future approach. Meat Sci 2005;70:543–54. https://doi.org/10.1016/j.meatsci.2004.07.015.

[60] Yessinou RE, Adoligbe C, Akpo Y, Adinci J, Youssao Abdou Karim I, Farougou S. Sensitivity of Different Cattle Breeds to the Infestation of Cattle Ticks Amblyomma variegatum, Rhipicephalus microplus, and Hyalomma spp. on the Natural Pastures of Opkara Farm, Benin. J Parasitol Res 2018;2018:2570940. https://doi.org/10.1155/2018/2570940.

[61] Muchenje V, Dzama K, Chimonyo M, Raats JG, Strydom PE. Tick susceptibility and its effects on growth performance and carcass characteristics of Nguni, Bonsmara and Angus steers raised on natural pasture. Animal 2008;2:298.

[62] Marufu MC, Qokweni L, Chimonyo M, Dzama K. Relationships between tick counts and coat characteristics in Nguni and Bonsmara cattle reared on semiarid rangelands in South Africa. Ticks Tick-Borne Dis 2011;2:172–7. https://doi.org/10.1016/j.ttbdis.2011.07.001.

[63] Breeds - Africander. Cattle Site n.d. https://www.thecattlesite.com/breeds/beef/25/africander/ (accessed August 3, 2022).

[64] Spickett AM, De Klerk D, Enslin CB, Scholtz MM. Resistance of Nguni, Bonsmara and Hereford cattle to ticks in a Bushveld region of South Africa. Onderstepoort J Vet Res 1989;56:245–50.

[65] Strydom PE, Naude RT, Smith MF, Kotze A, Scholtz MM, Wyk JBV. Relationships between production and product traits in subpopulations of Bonsmara and Nguni cattle. South Afr J Anim Sci 2001;31:181-194–194. https://doi.org/10.4314/sajas.v31i3.3801.

2 Chapter 2: Literature Review

Suitability of GWAS as a Tool to Discover SNPs Associated with Tick Resistance in Cattle: A Review

Nelisiwe Mkize [1, 2, *], Azwihangwisi Maiwashe [1], Kennedy Dzama [2], Bekezela Dube [1] and Ntanganedzeni Mapholi [3]

[1] Agricultural Research Council-Animal Production Campus, Private Bag X2, Irene 0062 South Africa, norman@arc.agric.za (A.M.); dubeb@arc.agri.za (B.D.)

[2] Department of Animal Sciences, University of Stellenbosch, Private Bag X1, Matieland, Stellenbosch 7602, South Africa, kdzama@sun.ac.za

[3] Department of Agriculture and Animal Health, University of South Africa, Private Bag X6, Florida 1710, South Africa, maphon@unisa.ac.za

* Correspondence: nemkize@outlook.com

Abstract: Understanding the biological mechanisms underlying tick resistance in cattle holds the potential to facilitate genetic improvement through selective breeding. Genome wide association studies (GWAS) are popular in research on unravelling genetic determinants underlying complex traits such as tick resistance. To date, numerous studies have been published on single nucleotide polymorphisms (SNPs) associated with tick resistance in cattle. The discovery of SNPs related to tick resistance has led to the mapping of associated candidate genes. Despite the success of these studies, information on genetic determinants associated with tick resistance in cattle is still limited. This warrants the need for more studies to be conducted. In Africa, the cost of genotyping is still relatively expensive; thus, conducting GWAS is a challenge, as the minimum number of animals recommended cannot be genotyped. The required minimum population size and genotyping cost challenges may be overcome through the establishment of collaborations. Thus, the current review discusses GWAS as a tool to uncover SNPs associated with tick resistance, by focusing on the study design, association analysis, factors influencing the success of GWAS, and the progress on cattle tick resistance studies.

Keywords: tick control; genotyping technology; quality control; association test

2.1 Introduction

Traditionally, animal improvement programs were limited to phenotypic information only, which may be ineffective for traits that are costly to measure, such as tick resistance. In cattle production, the presence of bovine ticks is considered as one of the main sources of diseases,

which are detrimental to animal health [1]. Ticks and tick-borne diseases (TTBDs) have substantial effects on animal health and welfare wellbeing, as well as a serious economic impact in both developed and developing countries [2]. To mitigate the bovine tick burden, wide ranges of tick control strategies have been adopted; however, they are ineffective in completely eradicating ticks. Currently, farmers use acaricides, which were efficient when they were introduced; however, they later developed limitations which are detrimental to the animal production economy and to consumers. Various researchers have reported that the prolonged utilization of acaricides on food-producing animals creates the development of acaricide resistance by ticks, subsequently reducing the efficacy of chemicals [3-4]. Additionally, acaricide residues have been traced in milk and meat products [5]. The presence of these residues in food products and in the environment poses a health threat to human beings. This underlines the need for alternative tick control measures, which are chemically-free and environmentally friendly [6]. A potential alternative approach to control ticks would be the use of genomic information, which entails the exploitation of genetic variation in host resistance to tick infestation. The success of this approach depends on the discovery of genetic determinants associated with low tick load in cattle.

The development of high throughput genotyping technologies has provided an opportunity to identify novel genetic variants, such as single nucleotide polymorphisms (SNPs), associated with economic traits in cattle. SNPs are genetic markers of choice because they are heritable and abundantly distributed across the genome. Genome-wide association studies (GWAS) are increasingly becoming the common experimental approach to investigate SNP markers associated with various economic traits in animal production. This approach operates by associating the phenotype with the genotype data to investigate the causal genetic variants for traits of interest using statistical models. The use of SNP markers in breeding for tick resistance (low tick load in cattle) has the potential to assist breeders in making informed decisions to improve host resistance to ticks in cattle [7–9]. Using GWAS, several studies have been conducted to investigate genetic variants for tick resistance in cattle, in different breeds and regions.

Studies that have been conducted to date have presented evidence of the association of various genomic regions with low tick load in cattle and recommended the validation of the discovered regions. Some of the challenges associated with GWAS include different phenotyping methods and genotyping strategies. These challenges could be overcome by the standardization of phenotyping procedures for tick count [15]. Genotype imputation has been identified and recommended as a cost-effective approach to account for the missingness of

genotyped data and facilitate the improvement of GWAS power. Additionally, the establishment of collaborations holds the potential to solve issues associated with small discovery populations and running costs for GWAS for tick resistance.

Despite the number of tick resistance GWAS studies that have been conducted to date, the availability of data is still a challenge, which is a global challenge hindering the success of improving tick resistance traits through genomic selection. Several researchers noted that in depth knowledge of genome variation for tick resistance in cattle is required [7,16]. The generation of genomic information for host resistance to ticks is currently gaining more attention because there is a need for the knowledge on genetic determinants influencing this trait. This paper reviews GWAS as a genetic tool to identify genetic variants associated with resistance to ticks in cattle.

2.2 GWAS Overview

The advent of high-density SNP genotyping platforms has provided opportunities to detect quantitative trait loci (QTLs) and uncover the genetic architecture of quantitative traits. This development has stimulated interest among researchers to explore genetic variabilities associated with various diseases affecting animals using GWAS. Genome-wide association analysis relies on recombination to rearrange the genome. Its underlying principle is to seek correlation between phenotype and genotype based on a non-random association of alleles at two or more loci [17,18].

This GWAS approach has successfully uncovered genetic determinants associated with disease susceptibility and resistance in humans, animals, and plants [19–21] However, using this approach to uncover genetic determinants associated with traits which follow polygenic patterns of inheritance, such as tick resistance, is not straight forward, since such traits are controlled by multiple genes. Despite this challenge, the approach is used to search for marker variants indirectly associated with certain diseases or traits of interest by assuming that a marker is in linkage disequilibrium (LD) with the underlying causal variant [22]. Linkage disequilibrium refers to the non-random association of alleles at different loci in each population [23]. Currently, the GWAS approach is gaining popularity in mapping QTLs associated with traits of economic importance or complex traits. This is because GWAS can detect variants that can be in LD with the causal variant, and this information could be used to narrow genomic regions that harbour causal variants [24,25], providing genetic determinant information that could be useful for the genetic selection of economic traits, such as tick

resistance in cattle. The continued success of GWAS depends on careful population selection and collaborative analytical approaches. Work by [26] reviewed the guidelines for successful GWAS analysis intensively and presented a useful workflow, which might be of value when conducting GWAS.

Bovine GWAS have successfully discovered genetic determinants associated with distinctive disease resistance or susceptibility, such as tuberculosis [27], resistance to ticks [8,12], mastitis [28]; and foot and mouth disease [29]. Moreover, this approach has been used in the successful mapping of genetic variants associated with meat quality [30,31] and milk production [32,33]. Such studies assist in providing information on the genetic architecture of QTL, generate biological knowledge about the expression of economic traits and facilitate the improvement of genomic selection.

2.3 Computer Software for GWAS and Genomic Public Databases

The most used computer programs for GWAS are presented in Table 2.1. They perform the same activities, and their access is free. The effective use of these programs requires the user to have operating skills.

Table 2.1: Common publicly available computer programs for GWAS.

Software	Focus	Website	Reference
PLINK	Stratification, LD, and structured association mapping	http://pngu.mgh.harvard.edu/purcell/plink accessed on 20 May 2019	[34]
R (GenABEL)	Stratification, LD, and structured association mapping	https://cran.r-project.org/src/contrib/Archive/GenABEL accessed on 20 May 2019	[35]
SVS	Stratification, LD, haplotype blocs and structured association mapping	http://www.goldenhelix.com accessed on 1 June 2019	[36]
GenAMap	Stratification, LD, and structured association mapping	http://cogito-b.ml.cmu.edu/genamap accessed on 1 June 2019	[37]

GEMMA	Stratification, Fits LMM and BSLM models, IBD analysis, estimation of chip heritability, and association mapping.	http://www.xzlab.org/software.html accessed on 1 June 2019	[38]
Blupf90	Data conditioning, estimate variances using several methods, and use SNP information for improved accuracy of breeding values + for genome-wide association studies (GWAS)	http://nce.ads.uga.edu/wiki/doku.php?id=documentation accessed on 17 November 2021	[39]

GEMMA – genome-wide efficient mixed model association; LLM – linear mixed model; BSLM-Bayesian sparse linear mixed model; SNPs – single nucleotide polymorphisms, LD – linkage disequilibrium; IBD – identical by descent, SVS – SNP and variation suite

The information generated from genomic studies is housed on different web databases for public access. Such databases include NCBI, EMBL-EBI, Ensembl, Animal QTLdb and NAGRP. Some of these databases are not specific to any organism, while some are specific to livestock genomics (Table 2.2). To date, several QTLs and associations related to tick resistance have been identified in different cattle breeds using different research approaches. The information of the QTLs and associations is recorded in a database known as Animal QTLdb, which is currently updated whenever there is new information [40]. Based on the available information, it is noted that most of the QTLs and associations have been uncovered on chromosome 10, followed by BTA23, BTA14, BTA11, BTA2 and others. This is depicted in Figure 2.1, sourced from the Animals QTLdb database. Information on African indigenous breeds are lacking on these databases. The limited information contained on these databases concerning tick resistance in cattle breeds shows that more research still needs to be conducted, especially in African cattle breeds.

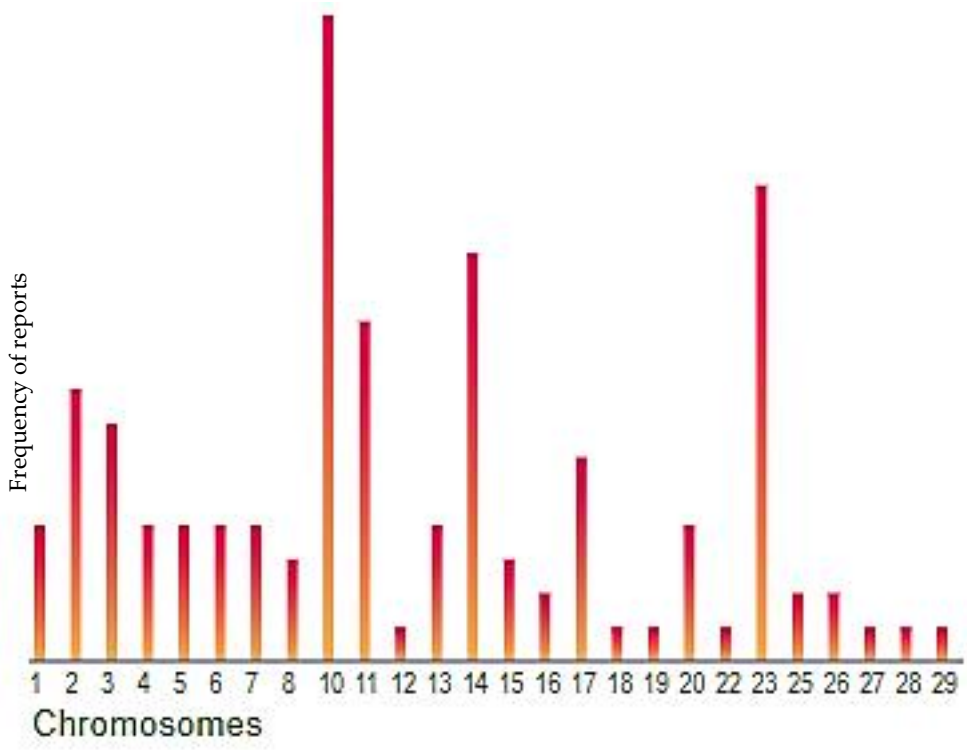

Figure 2.1: Distribution of tick resistance related QTL/associations in bovine, based on count of report data *(sourced from: https://www.animalgenome.org/cgi-bin/QTLdb/BT/index (accessed on 4 September 2020).*

Table 2.2: Some web databases that house genomic information associated with economic traits.

Genomic database	Description	URL
NCBI (Genbank)	Repository for biomedical and genomic information	https://www.ncbi.nlm.nih.gov/ accessed on 06 December 2019
Ensembel	Genome browser	https://www.ensembl.org/index.html accessed on 06 December 2019
Animal QTLdb	Animal QTL database	https://www.animalgenome.org/cgi-bin/QTLdb/index accessed on 04 September 2020
NAGRP	Genomic information browser	https://www.animalgenome.org/ accessed on 04 June 2020
EMBL-EBI	Genomic information database	https://www.ebi.ac.uk/ accessed on 15 June 2020
DDBJ	Genomic information browser	https://www.ddbj.nig.ac.jp/index-e.html accessed on 15 June 2020
UCSC	Genome browser	https://genome.ucsc.edu/ accessed on 15 June 2020
Refseq	Reference sequence database	https://www.ncbi.nlm.nih.gov/refseq/ accessed on 15 June 2020
VEGA	Genome browser	http://vega.archive.ensembl.org/index.html accessed on 15 June 2020

Animal QTLdb – Animal quantitative trait loci database;
NAGRP – National animal genome research program;
EMBI-EBI – European molecular biology laboratory-European bioinformatics institute;
DDBJ – DNA data bank of Japan;
UCSC – University of California Santa Cruz;
Refseq – Reference sequence;
VEGA – vertebra genome annotation

2.4 Available Genotyping Platforms and Coverage

The abundance of SNPs in the genome and its ability to be amenable to high throughput automated analysis make SNP genotyping the most preferred approach to studying genetic variation in animals, humans, and plants [41–43]. Moreover, SNPs are heritable and allow single base resolution, making the identification of causal markers easy [44]. Initially, SNP genotyping arrays were developed for human studies [45], and then, the technology was adopted in animal and plant research. In cattle research, three well-known commercial companies that produce commercial Bovine SNP arrays are Illumina, Neo-GeneSeek® and Affymetrix. These companies have developed SNP genotyping platforms with different densities, which are used for GWAS, identification of selection sweeps, and investigating genome-wide genetic diversity and relationships in cattle. The summary of SNP genotyping platforms available for Bovine is provided by [46]. The SNP genotyping platforms include different densities of SNPs, ranging from as low as Golden Gate Bovine 3K (2900 SNPs) to Bovine HD (777,962 SNPs) [47]. The density of an array plays a key role in the success of a GWAS study. It was recommended by [48] that denser SNP arrays should be used for crossbreed GWAS studies, due to its large hypothetical effective population size. GWAS experiments performed using data from denser SNP arrays are provided with enough marker density to dissect the genetic architecture of the trait of interest. In instances where low-density SNP arrays have been used, genotype imputation is advisable. This phenomenon will be briefly discussed later in this article.

The development of SNP arrays is advancing rapidly, making genomic selection feasible [47]. They noted that the development of genotyping platforms has not been regulated under a standardized system. This results in difficulties in making comparisons and merging data genotyped by different commercial companies. Therefore, there is need for standardization during SNP array development to minimize downstream research challenges [47]. Despite this success, studies on local African breeds are still facing the challenge of the SNP arrays being originally created using exotic breeds. This underscores the need for the development of SNP arrays that will incorporate ancestry data from African breeds. The development of such SNP arrays could potentially extend insights on the architecture of traits being investigated using African cattle populations. Although the decreasing cost of genotyping potentially makes GWAS possible and affordable, in Africa the cost of genotyping is still relatively expensive. This is a challenge for GWAS because it is difficult to genotype the minimum number of animals recommended [48]. The minimum number of animals required is determined by conducting a statistical power test, to ensure the low rate of discovering false

positive results in GWAS [49]. The issue of attaining a better number of samples for GWASs can be solved through the establishment of collaborations.

2.5 Testing for an Association

The main goal for GWAS is to test a null hypothesis, stating that there is no association between the genotype and the expression of the trait of interest. This is facilitated by choosing the appropriate association test approach, influenced by covariates, population structure, study population and pedigree structure [50]. A single locus statistical test and multiple locus tests are the two approaches that are currently being used, depending on the focus of the study. A single locus statistical test compares the genotype and the phenotype by focusing on one SNP at a time [51]. This test uses a regression model, assuming that the trait being studied is normally distributed and the variance is the same within a population. Testing one SNP at a time results in multiple tests, which may produce false positive and false negative results [48,52,53]. Therefore, it is necessary to correct for multiple testing to prevent spurious associations, making it impossible to conduct follow up studies. This can be performed using the false discovery rate (FDR) and Bonferroni correction. The adjustment through the FDR approach corrects for the expected rate of false discoveries, and gives the investigator an insight into the proportion of true associations in the study [54,55] However, this adjustment is considered less stringent compared to the Bonferroni correction [55,56].

After the adjustment, a SNP is considered statistically significant if its p value is less than or equals the adjusted genome-wide cut off [55]. A single SNP association approach worsens the missing heritability problem, which is the gap between the heritability measured using pedigree information and that measured through GWAS [57,58]. Biological mechanisms, such as epistasis, epigenetics, and others, are attributors of missing heritability. However, in a human-based study, none of these mechanisms accounted for the missing heritability [59]. The study included human microbiome information to understand the heritability of a given trait in humans. They reported that microbiome is associated with many important traits and encodes for extra genes which interact with human genes. The interaction can be a source of variation and phenotypic plasticity [59]. Thus, the inclusion of microbiome in GWAS for cattle can be used to solve the issue of missing heritability. Alternatively, the multilocus association approach can be adopted because it examines nonlinear relationships genome-wide. Multilocus models do not require Bonferroni correction; this is beneficial because it reduces the high chances of losing many loci associated with the targeted trait through failing to meet the stringent requirement for the significant test, as it happens in single locus models [60].

Multilocus models are more advantageous compared to single locus models, because of their ability to allow the estimation of three variance components, a high power of QTL detection, and the SNP effect is random [61]. Table 2.3 shows some examples of single and multilocus models that are currently being used in GWAS across distinct species.

Table 2.3: Models that can be used for GWAS analysis.

Model Type	Model	Reference
Single locus	General linear model (GLM)	[62]
	Mixed linear model (MLM)	[63]
	Logistic mixed model (LMM)	[64]
	Compressed mixed linear model (CMLM)	[63]
Multi-locus	Multilocus random SNP effect mixed linear models (mrMLM)	[65,66]
	Fast multilocus random SNP effect efficient mixed model association (FASTmrEMMA)	[67]

During the association analysis, population stratification is regarded as one of the confounding effects that can inflate the variance of the usual statistics [65]. The inflation of test statistics may potentially attribute high false positive discoveries; hence, it should be accounted for. There are two ways to account for population stratification in GWAS: the variance is adjusted using genetic control or principal component analysis. It was pointed out by [65] that genomic control adjusts the variance by calculating the statistics on data from null loci. Figure 2.2 shows the quartile–quantile plots before correction (A) and after (B). Figure 2.2A shows a departure from the diagonal observed, which indicates a high inflation rate. Figure 2.2B shows the improvement after correction using genomic control, where the inflation rate has decreased.

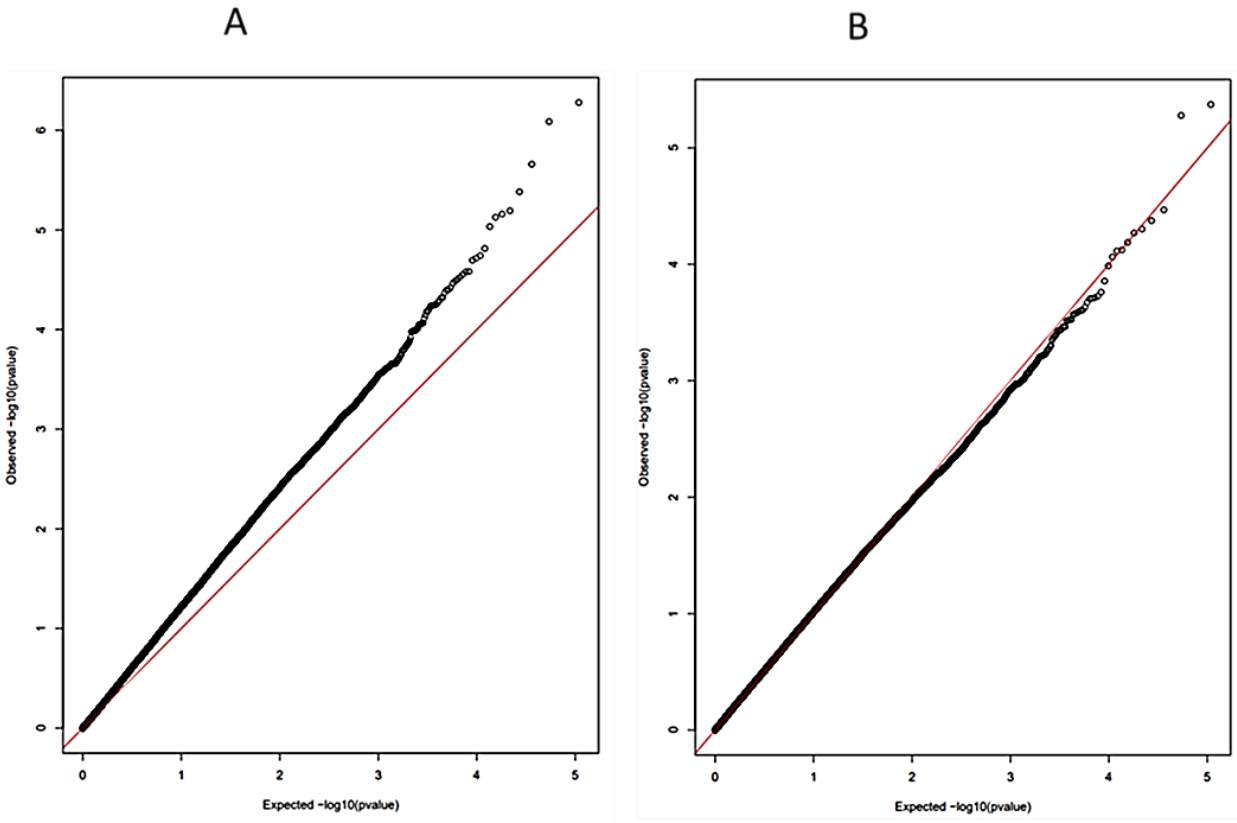

Figure 2.2: A-Quartile–quantile plots before correction. B-Quartile–quantile plots after correction.

At the end of the association analysis, an independent association test replication is recommended on the SNPs to validate the findings [66]. This is, however, still not applicable for most studies, because of the associated cost implications, time, and other factors associated with the study design. When a replication study is conducted, GWAS is performed on statistically significant SNPs. This process makes use of a small sample size, attributing to a low GWAS power and making it difficult to confirm the initial findings [67]. Lack of replications does not necessarily mean findings are not valid, but the study should be professionally designed, and all confounding factors accounted for, to ensure the validity of the findings.

The preferable mode to summarize and present GWAS findings visually is the use of a Manhattan plot (Figure 3). Figure 2.3 presents y-axis $-\log_{10}$(P-value) versus x-axis (the chromosome position for each SNP tested), where each circle signifies a SNP. The SNPs are stacked together to form a signal that is influenced by the level of LD amongst the SNPs with relation to the causal marker. Several studies have used Manhattan plots to present their GWAS findings in cattle [8,14]. In South Africa, [8] conducted a GWAS study on the indigenous Nguni and produced the plots presented in Figure 2.3. The two dotted horizontal lines depict the suggestive (red) and the actual genome-wide cut off line [grey], which represents the level of significance. If a SNP passes the grey line, it is considered significant. On this study, a

significant SNP associated with tick resistance in Nguni cattle was observed in BTA10. There are many free tools that can be used to graphically plot a GWAS Manhattan diagram; examples include SNPEVG [68], R package (qqman) [69], Stata [70], and Manhattan [71]. When a GWAS association has been detected, tools such as Manhattan harvester [72] and Locus Zoom [73] provide opportunities to study the detected region in depth. Both tools focus on the physical position of the chromosome of interest; however, LocusZoom is more informative because it allows the visualization of LD levels, recombination rates and genes [73]. The LocusZoom tool has mostly been used in human research. Figure 2.4 depicts an example of a LocusZoom plot, sourced from a human GWAS study [74], showing the architecture of a region of interest in chromosome 19 (significant SNPs, associated recombination rate, and genes). The use of tools such as LocusZoom in bovine research will provide an in depth insight into the landscape of genetic contribution associated with the expression of economic traits, such as tick resistance, growth traits, milk production and others.

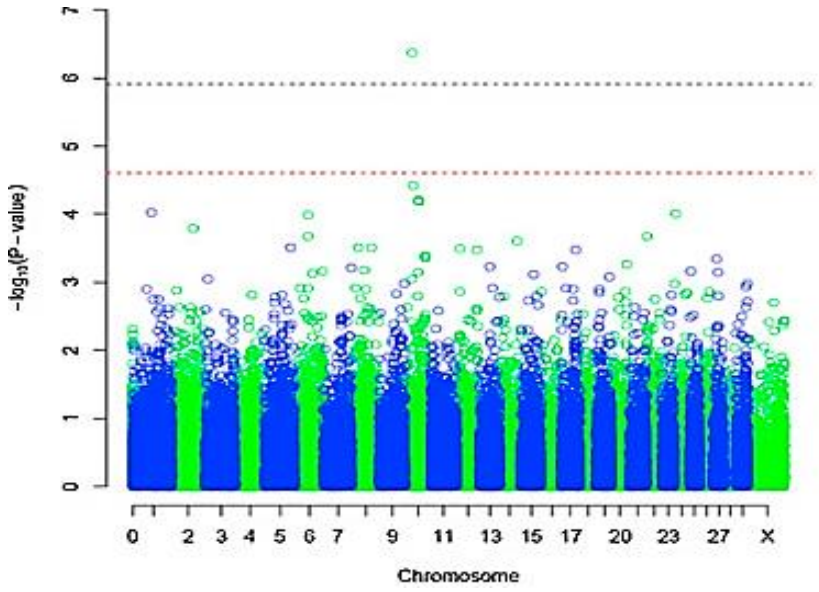

Figure 2.3: Manhattan plot showing findings for a single marker GWAS where the association of low tick load (total A. hebraeum ticks) and genotype was assessed in Nguni breed, using a genome-wide p value < 0.05 as a cut-off. The redline indicates suggestive threshold and the grey indicates the genome-wide cut off (taken from [8]).

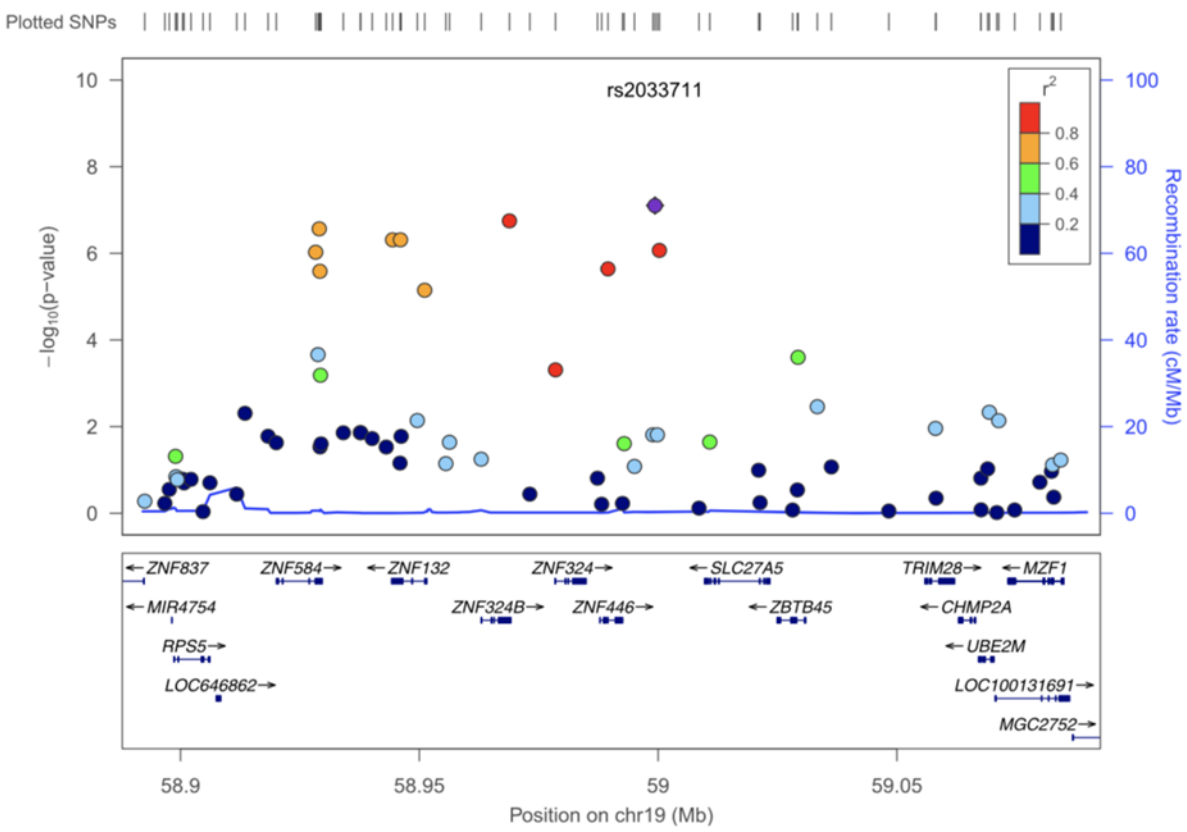

Figure 2.4: LocusZoom plot showing in depth findings for most significant SNPs in Chromosome 19, from a human GWAS that was focused to study genetic variation underlying renal uric acid excretion in Hispanic Children [74].

2.6 Post GWAS Analysis

The discovered GWAS hits are used to physically map candidate genes underlying the trait being studied. The National Centre of Biotechnology Information (https://www.ncbi.nlm.nih.gov/ accessed on 6 December 2019) and ENSEMBL(http://www.ensembl.org/index.html accessed on 6 December 2019) databases are used for gene annotation to identify candidate genes associated with identified SNP markers [75]. When the genes have been identified, their biological functionality or relevance can be verified using functional annotation databases such as DAVID and KEGG [58,76,77]. Furthermore, it is also possible to create gene networks using open source software such as Cytoscape [78]. The gene network provides a better biological understanding of the interaction of genes underlying the trait of interest. A recent GWAS by [14] identified SNPs and candidate genes (*TREM1, TREM2, CD83, MYO5A, TREML1, PRSS16*) associated with tick resistance, and then used the information to create a gene network. Tick resistance is a complex trait that is influenced by various determinants. Therefore, gene networks are particularly important to

give insight on the interconnection of genes responsible for the expression of the tick resistance trait. Despite the success that has been made, the inferring of true causal genes and biological mechanisms from GWAS results in tick resistance studies is still a challenge. This is due to the difficulties associated with the interpretation of GWAS findings and the limitation of available data.

2.7 Factors Influencing the Success of GWAS

2.7.1 GWAS Experimental Design

The success of GWAS requires proper study experimental design, in addition to factors such as population of interest, sample size and standardized data collection. In addition, concise pipelines to execute the actual analysis should be taken into consideration. Phenotyping data should be properly collected to reduce high rates of outliers, which may potentially create noisy data. The data is often tested for normality, to assess the distribution and the presence of outliers so that necessary steps can be taken to address the violation of normal distribution assumptions and the removal of outliers. The tick resistance trait is known to be not normally distributed and literature shows that some studies addressed this issue by transforming the data to confer normality. It has been noted that, although removal of outliers is important, it affects the size of the population tested. Selection of population for association analysis generates a structure that leads to specific genetic variation and an effect on the end use of association analysis [49]. In cattle, within family pure breeds and crossbreeds have been used for dairy and beef GWAS experiments to identify genomic regions associated with phenotypic variation in economically important traits. Crossbred lines, specifically the F_2 design, have been identified as an appropriate design because they exhibit a high level of LD compared to pure lines. High LD potentially increases the power of GWAS [79]. Studies have been conducted to study the LD patterns within the crossbred F_2 population [48,80,81].

In relation to cattle tick resistance studies, GWAS studies have been conducted using the crossbred F2 design in different regions [12,14]. However, it can be noted that some previous studies were conducted using data that was genotyped using microsatellite markers. The limitation of microsatellite markers involves the lack of an adequate number of informative markers, while on the other hand, SNPs are very abundant in the genome [82]. A study by [14] is currently the only known study that used the F_2 crossbred design to discover tick resistance genetic determinants using SNP genotyped data. Although all these studies have been conducted, and information generated, the development of crossbred F_2 populations remains a challenge. More time is required to build the required population, and this is

associated with inflated costs of SNP genotyping a large sample size, especially in developing countries. However, there is need for this population design to be used for the genomic improvement of the tick resistance trait, regardless of the associated shortcomings.

2.7.2 Phenotyping

Phenotyping for tick resistance in cattle is often assessed through natural or artificial infestation. The assessment of tick count under the approach of interest is one of the baseline indicators for tick resistance differences. Tick count is the commonly used mode of phenotyping in studies that assess tick resistance differences in a population. Counting ticks for GWAS phenotyping where a large population size is the requirement becomes a bottleneck. A study by [83] noted that counting ticks on every animal in a population is labor intensive, time consuming, and requires trained technicians and expensive infrastructure to constrain the animal. This stresses the animals through handling and the use of a trained technician does not rule out the possibilities of bias tendencies when counting. More challenges are faced under natural infestation because of multiple tick species and the need to categorize them accordingly when counting. All these shortcomings play a part in the reduced success rate of recording a tick count for big study populations. A study by [84] on Nguni cattle estimated the correlation between whole body tick count and tick count in different body parts. They also assessed correlations between tick species and different seasons, and observed that some tick species were prevalent in certain seasons, indicating that seasons are influential on tick distribution. The approach for this study is highly informative because it gave insight into the distribution of ticks in various parts of the animals. The challenge is that the process of counting ticks is labor intensive, and this will be exceedingly difficult for studies with large population size.

Artificial tick infestation has been used as an alternative approach; however, this approach does not represent the true reflection of natural infestation in regions where various tick species inhabit [83]. To solve this problem, [85] introduced a scoring method that overcame some disadvantages associated with tick count. However, [86] reported that the scoring method provides a low heritability, as compared to tick count. Low heritability constitutes a reduced power to detect association, since heritability is a good indicator that explains the genetic variance contribution towards the expressed phenotype [49]. Therefore, since the scoring method is associated with low heritability estimates, tick count remains a better method of phenotyping for tick resistance GWASs. However, there is need for high throughput phenotyping methods that will eliminate the issues of biasness when counting ticks and reduce associated costs and stress imposed to animals.

2.7.3 Population Size

Population size is amongst the limiting factors affecting the statistical power of a GWAS. A small population size results in reduced statistical power to detect a causal variant. This poses a major challenge in understanding the biological mechanism underlying economic traits such as tick resistance. A population comprising 1000 individuals is regarded as a better population size, constituting about 80% GWAS statistical power [87]. In certain research set ups, it is difficult to attain a minimum of 1000 samples to conduct a GWAS, because of numerous factors at play associated with infrastructure and costs. These factors are not limiting researchers' drive to find information that could be used to improve animal welfare and facilitate sustainable food security in the current trying times. According to [88], a population of individuals ranging from 100 to 500 can be suitably used to perform GWAS. However, such a population needs to be carefully selected based on the status of the genetic variation of the trait being studied, and other factors influencing phenotype such as environment. Moreover, the selection criterion must take population stratification into consideration to avoid false positive results. The selected GWAS population should result in low stratification[89] and large genetic variation to potentially detect true variants that can be used in breeding schemes to improve the trait of interest.

GWAS statistical power is defined as the probability of rejecting a null hypothesis under an assumed alternative hypothesis [87]. It depends on a clear study design and the number of samples used for the analysis. One thousand samples, equating to 80% power, are considered the normal number in human and animal studies [87]. The same applies to livestock studies. However, the issue of sample numbers remains a problem in developing countries. Hence, many GWAS livestock studies are conducted on samples below one thousand. Factors playing a role in this regard include the lack of recorded phenotypes for certain traits and the availability of resources to cover genotyping costs. Although such costs are gradually decreasing, researchers in Africa are still battling to afford genotyping on cattle numbers near the normal one thousand mark. The situation is even more difficult when high-density SNP genotyping chips are considered. GWAS analysis requires a large sample size to achieve sufficient statistical power [90,91]. Attaining a large population size for tick resistance studies is a challenge in developing countries, because of lack of funds, resources, well trained data collectors and infrastructure. To tackle this problem, the establishment of collaborations could be beneficial, since it will promote the sharing of information, insights, and funds. This will improve the development of GWAS and its quality.

2.7.4 Data Quality Control for GWAS

Prior to actual genome-wide association analysis, the genotyped data is subjected to quality control (QC) to decrease the chances of discovering false positive and false negative associations [92]. False positive association is defined as the occurrence of identifying a SNP association that is not profoundly influencing the trait of interest in the study. On the other hand, false negative association is defined as an incidence where a SNP that is influencing the trait of interest is not associated with the trait in the study. The genotyped data is subjected to stringent filters that are performed on a sample and SNP level. The quality control filters include missing call rate, minor allele frequency (MAF) and Hardy Weinberg equilibrium (HWE). The samples with a missing call rate higher than 1–5% could be a result of poor DNA quality potentially influencing genotypic errors [56]. The removal of SNPs with high missing genotypes may increase SNPs with accurate genotype calls for downstream analysis. However, in a study with a small population size, it is not good practice to lose samples because this negatively affects the power of the study. This can be overcome through genotype imputation to replace the missing SNP markers, explained later.

The other aspect of QC is to assess the format of genotyped data, where sometimes there is a mixing of the AB and ACTG formats, which needs to be corrected to form one uniform format. Additionally, data is assessed for MAF, which removes SNPs not complying with a given threshold for a particular study. Exceptionally low allele frequencies are less informative [93] and can result in the discovery of fake associations [94]. GWAS capitalizes on the LD that exists between the markers; thus, it is especially important to assess the deviation of SNPs from the HWE. The deviation of SNPs from HWE is set using Chi square test [95], where the significance level for rejecting is based on P values ranging from 10^{-5}–10^{-7} [50,56]. The SNPs that are not in compliance with a stipulated criterion are removed from the dataset that is used in the downstream analysis. Overall, statistical software such as PLINK, and R environment are used for quality control. However, PLINK is the most preferred because it is free and flexible to accommodate large scale data management [56]. It is noted there is no universal criterion to perform QC and there is no perfect QC pipeline that can capture all the problematic SNPs in a population being studied. Therefore, it is important to view the clustering intensity plots for SNPs, to ensure that there are no obvious clustering problems.

2.7.5 The Extent of LD Measures r2 in GWAS

GWAS relies on the extent of LD between markers across the genome, where its measure (r^2) is bounded between 0 and 1, with 1 considered as the perfect association [48]. Linkage

disequilibrium in a population can be affected by population structure, genetic drift, selection, recombination rate, migration, and mutations [96]. The development of LD maps and haplotype block structures at the population level are useful parameters for guiding GWAS. In association studies, the presence of LD creates two outcomes, namely, direct, and indirect association [97]. Direct association occurs when the SNP influencing a biological system is directly genotyped in the study and found to be statistically associated with the trait. The indirect association outcome occurs when the SNP is indirectly linked with the trait. The phenomenon is termed taq SNP and is graphically explained well in a review by [97]. The feasibility of GWAS strongly depends on the extent of LD, as the latter determines the required SNP markers and mapping resolution [98]. Therefore, it is important to study the extent of LD in a population of interest before the association analysis is performed.

2.7.6 The Effect of Genotype−Environment Interaction

According to Falconer et el. [99], genotype by environment (GxE) interaction is defined as the different responses of genotypes under different environmental conditions. GxE affects the ranking of animal performance under different environmental conditions. Therefore, it is important to closely monitor economical traits that are influenced by environmental factors [100]. In cattle production, tick resistance is among the traits that are highly influenced by environmental factors. Climate change influences the distribution and density of tick populations and because it influences the life cycle of ticks, it increases the chances of tick−host interactions.

Currently, genetic evaluations for tick tolerance in countries such as Australia and Brazil are performed routinely. Despite the success of these evaluations, Moat et al. [101] noted that GxE is not taken into consideration in these evaluations. They also pointed out that failure to consider GxE interaction in genetic evaluations can potentially affect genetic gain, as the selection of candidate comparisons is environmentally dependent. Thus, the cost implications of not accounting for GxE can be high [102]. This is because animals observed as top performers in one environment will not necessarily perform the same in a different environment, a phenomenon that is associated with the loss of genetic progress. Few studies have provided evidence that resistance to ticks in cattle can be influenced by various environmental effects [103]. A recent study investigated the existence of GxE using different models in Hereford and Braford cattle [101]. Their findings showed that the estimates of repeatability varied along the environmental gradient (range 0.18–0.45), indicating that resistance to ticks is environmentally influenced. Additionally, the posterior means of the

genetic correlations across the environmental tick infestation surface plot demonstrated a large plateau above 0.80. This indicates that there will be re-ranking of performance for a trait of interest between environments, which necessitates the separation of breeding programs for each environment [104,105]. GxE interaction contributes to the genetic architecture of complex traits and it affects the chances of discovering a true association between phenotype and genotype in GWASs. According to Cooley et al. [106], failure to adjust for environmental effects results in the reduced chances of predicting an association. This has been proven through assimilation studies, intensively in human epidemiological studies [107–109] and plant studies [110,111].

2.7.7 Batch Effect

For most study designs, samples are not genotyped at once, instead, they are handled in batches. One of the reasons behind this can be the use of substantial number of samples, which makes it impossible to genotype the samples at the same time. In addition, some studies collect samples at different time intervals, prompting the genotyping to be also conducted at different time intervals. For cattle tick studies, sometimes data is collected from different environments that are geographically spaced, leading to data being treated in batches. Additionally, for cattle ticks studies conducted on hybrids, the data is partitioned according to the development of the hybrids. It is known that the development of hybrid populations can never be carried out at one go. Instead, it is conducted in batches, for example, the development of an F_2 population. The partitioning of samples gives rise to the batch effect, which results in apparent associations confounded by the batches. It is therefore necessary to assess the dataset for a potential batch effect, since it has potential to yield spurious association if this is not accounted for. Before the actual GWAS association test, the batches' genotyped data should be handled independently to assess the presence of confounding effects [112].

2.7.8 Genotype Imputation as a Cost-Effective Approach to Improve the Power of GWASs

The main purpose of imputation is to infer missing genotypes of the SNPs that are not directly genotyped in the study, using *in silico* haplotype information from reference samples with genotypes from denser genotyping arrays [113,114]. Genotype imputation holds the potential to improve the statistical power to detect association by reducing the number of missing genotypes, thereby increasing the overall number of genotypes available for association

analysis [113]. Genotype imputation has the potential to boost GWAS statistical power by up to 10% over testing only genotyped SNPs [115]. The performance of association tests on typed SNPs may not lead to a significant association, especially when the sample size is small [113]. Their findings were different when the association test was conducted on imputed genotyped data, where there was also a more detailed view of the association region.

Various tools are available for imputation, and some are presented in Table 2.4. For example, BEAGLE has been built to handle genotype intensity data so that genotypes can be called using LD information between the SNPs, offering an improvement in genotyping error rates. Imputation is a cost-effective tool to generate genomic variants at denser platforms [46,116]. Imputation allows high-density genotypes to be imputed reliably from low-density SNP arrays potentially solving the affordability issue in developing countries, as more animals can be genotyped at low cost. However, the accuracy of imputation and the factors that affect it should be taken into consideration because they determine the reliability of the tool.

Table 2.4: Some available software packages for genotype imputation.

Software	Usage	Website
BEAGLE	Prephases haplotypes infers missing genotypes, and identifies IBD in related samples	https://Faculty.washington.edu/browning/beagle/old.beagle.html accessed on 9 July 2020
GIGI	Imputes missing genotypes on a pedigree	https://faculty.washington.edu/wijsman/progdists/gigi/software/GIGI/GIGI.html accessed on 9 July 2020
IMPUTE2	Prephases haplotypes, infers missing genotypes	https://mathgen.stats.ox.ac.uk/impute/impute_v2.html accessed on 9 July 2020
MaCH/ minimac3	Prephases haplotypes, infers missing genotypes	https://github.com/statgen/Minimac4 accessed on 9 July 2020

IBD – identical by descent; GIGI –Genotype imputation given inheritance

In humans, whole genome sequencing and imputation based GWAS strategies were used to refine the association signals and recover novel association signals for complex traits [117,118]. Sequencing and imputation GWAS is powerful and cost effective, and can also be applied on non-European populations [119]. In cattle, the whole genome sequencing and imputation GWAS strategies have been applied to study the genetic architecture of quantitative traits in beef cattle[50]. In relation to tick resistance, most GWAS were carried out

using 50K genotyping platforms (Table 2.5). Thus, some studies used imputation to ensure better statistical analysis effectiveness [13,14,120]. Imputation boosted the number of SNPs that were tested for association and subsequently improved the power. Imputation makes it applicable for cattle tick resistance-focused studies to be conducted using samples that were initially genotyped at low-density platforms, then imputed to high density. This has been proposed as a cost-effective approach than can solve the current problems associated with generating genomic data in cattle production, especially in developing countries. Currently, in different regions, there are ongoing studies that are investigating the feasibility of the tool for cattle research. The outcomes of these studies will give insights that could be used to properly apply this tool in GWAS aimed to study biological mechanisms underlying tick resistance.

2.8 Progress on Tick Resistance GWAS in Cattle

Quantitative trait loci studies using microsatellites and SNPs have been inconsistent, with an exceptionally low percentage relating phenotypic variation to tick infestation [121]. Most studies were conducted in subtropical regions such as Brazil, Australia, and Mexico (Table 2.5). Brazilian studies have successfully mapped genomic regions associated with resistance to ticks on F_2 Gyr x Holstein and on Hereford and Braford [13]. Similar studies were conducted in Australia by Turner et al. [122] . In some instances, different studies identified QTLs on similar chromosomes, regardless of the differences in the breeds and tick species used. This underlines the need for the validation of the role of these chromosomes in cattle tick resistance. Validation could be pursued through GWAS meta-analysis, which can be achieved through collaborations.

To date, only one study in Africa used SNP genotyping and GWAS as an approach to investigate genetic variants associated with tick resistance in cattle [8]. This study was conducted on South African Nguni cattle and identified several genomic regions harbouring QTLs associated with tick count traits. Despite the studies conducted previously, information on genetic determinants associated with cattle resistance to ticks is still limited. Further investigations focusing on unravelling genomic determinants associated with tick resistance will identify and provide understanding on biological mechanisms associated with TTBDs in cattle production. The information from the investigations will present a wonderful opportunity to improve breeding programs to produce animals that are more resistant to tick infestation, while enhancing productivity [7,12].

Table 2.5: Previous GWAS studies on genomic regions associated with tick resistance in different regions of the world.

Region	Breed	Sample size	Mode of infestation	Genotyping platform	Findings	Reference
Brazil	F_2 *B. taurus* × *B. indicus*	382	Artificial	Microsatellite	Identified significant genomic regions on chromosomes 5, 7 and 14	[11]
Brazil	F_2 Gyr × Holstein	376	Artificial	Microsatellite markers	Identified dry season specific QTL on BTA 2 and 10, rainy season specific QTL on BTA 5, 11 and 27 and BTA 23 for both seasons	[12]
Australia	Brown-Swiss, Holstein-Friesian, mixed taurine	189	Natural	MegAllele genotyping bovine10K SNP	Identified genes associated with tick burden, namely *TNFSF8 [CD30], and SIRPA*	[122]
South Africa	Nguni	586	Natural	Illumina BovineSNP50 BeadChip	Identified significant genomic regions on chromosomes 1, 3, 6, 7, 8, 10, 11, 12, 14, 15, 17, 19 and 26	[8]
Brazil	Braford and Hereford	3455	Natural	Illumina BovineSNP50 BeadChip	Identified 48 tag SNPs associated with tick resistance	[13]
Brazil	F2 Gir × Holstein	46	Artificial	Illumina BovineSNP50 BeadChip	Identified genes associated with immune system function, namely, *TREM1*, *TREM2*, *CD83, MYO5A, TREML1, and PRSS16*]	[14]

F_2 –Second filial generation; QTL – quantitative trait loci; BTA –Bos taurus; SNPs – single nucleotide polymorphisms

2.9 Breeding Cattle for Tick Resistance

Breeding for genetic resistance is a potentially promising strategy to control ticks [123]. Sufficient genetic variation is one of the key factors determining the success of breeding schemes in livestock production. Bovine quantitative genetics studies have demonstrated low to high heritability for resistance to ticks depending on the breed [124,125]. Such findings hold the potential for tick resistance to be included as a goal in breeding schemes. *Bos taurus* breeds are known to be highly productive; however, they are also known to be highly susceptible to ticks and this makes their use in tropical production systems unsustainable, especially by resource-poor farmers. On the other hand, *Bos indicus* breeds are known for being resilient toward ticks as compared to *B. Taurus*, but they yield lower production.

These differences motivate the development of crossbreeding programs between *Bos indicus* and *Bos Taurus* cattle. These programs aim to produce crossbred animals, to facilitate the improvement of production whilst controlling TTBDs using genomic selection. The success of genomic selection for tick resistance depends on the availability of proper genetic evaluation programs for the trait. The existence of genetic evaluation programs for tick resistance has the potential to generate information and facilitate the improvement of tick resistance. In countries such as Australia and Brazil, there are ongoing genetic evaluation programs for tick resistance. These develop crossbred animals that are productive under environmental conditions that have a high prevalence of ticks [126]. In Brazil, there is a genetic evaluation program known as Delta G Connection, which involves the Hereford and Braford cattle [101]. The Australians have developed the Australian Friesian Sahiwal, which produces acceptable levels of milk in an environment with ticks. The lack of information of such programs from other regions may suggest less enthusiasm from other regions, especially Africa and Asia [126]. There are concerns about the selection potential for resistance to ticks and the trade-offs with other traits of economic significance. In this regard, studies have demonstrated that there is exceptionally low genetic correlation between tick count and various productive, adaptive, and pubertal traits [104].

The existence of genetic components of variation in host resistance to ticks in cattle is currently being studied to discover tick resistance molecular markers that can be used in marker-assisted selection (MAS). This information from locally adapted breeds can be used through crossbreeding to upgrade local breeds using highly productive exotic breeds. The challenges hindering the application of MAS for tick resistance are costs, resource populations, requirements of technical skills and the validation of discovered molecular markers for each

population. All these challenges translate to the lack of molecular markers associated with tick resistance in cattle. However, despite the limitations, the development of studies that will investigate and generate genomic information on tick resistance and production traits holds the potential to increase the accuracy of selection. Therefore, the use of molecular genetics techniques, together with conventional breeding tools, is important in balancing the process of selection for tick resistance.

Host resistance to ticks is potentially an alternative tick control strategy that could solve the current TTBDs problems affecting the beef and dairy industry. The development of this alternative strategy requires the generation of knowledge that will broaden the understanding of biological mechanisms underlying tick resistance in cattle. The first step to control for a certain trait of interest requires the studying of genetic determinants influencing the expression of the trait. The advancement of technology has made it possible to use a genome-wide association approach to gain understanding on mechanisms underlying tick resistance in cattle and to generate knowledge. However, information is still limited, and this is one of the hurdles preventing the facilitation of breeding for tick resistance through genomic selection. The information gap warrants the need for more GWASs to be conducted, to provide an understanding on biological mechanisms underlying tick resistance.

The costs associated with TTBDs are a major constrain for beef and dairy production. Therefore, the use of genomic selection as a tool to breed for tick resistant cattle will reduce the costs associated with the intensive use of vaccines and acaricides. This will further ensure animal welfare, facilitate increased production, and subsequently increase profit margins. The implementation of genomic selection for tick resistance in developing countries is currently hindered by the inflated costs associated with the generation of a large cattle resource population, phenotyping, and genotyping. The success of genomic selection for tick resistance is possible using cost-effective genotype imputation methods to increase the power of GWASs and the accuracy of GEBVs estimates. There are countries that have included tick resistance trait in their breeding goals. However, from those countries, there are no studies that have been put in place to investigate the economic aspects of including the tick resistance trait in a genomic selection-breeding goal. Therefore, there is a need for studies that will investigate the prospect of cost-effective genomic selection for tick resistance**.**

2.10 Limitations for GWAS to Uncover Tick Resistance Causal Variants in Cattle

The discovery of common causal variants associated with tick resistance is limited by various attributes, which include the nature of the trait, methodological challenges, and the lack of financial resources. Tick resistance is a polygenic trait, influenced by more than one gene. This means the discovery of major genes using GWAS is impossible. Factors such as epistasis, epigenetics, microbiome, and environment attribute to missing heritability [127,128]. This results in GWAS not being able to capture all the genetic determinants underlying the expression of tick resistance trait [14]. The lack of biological understanding on how these factors influence missing heritability in cattle is part of the obstacles hindering the discovery of true loci associated with tick resistance. The failure of GWAS to capture common variants is not a limitation to GWAS only. To date, there is no perfect genomic technology available to capture all the genetic information underlying the expression of complex traits [129]. Despite this drawback, there is a pressing need to use the available technologies to discover information that could be used to improve the trait through selection. Methodological challenges such as the GWAS study design of choice, poor structure of the data, phenotyping uncertainties, genotyping errors, improper data analysis, and the poor interpretation of results plays a huge role in limiting the success of GWAS for the trait of interest. The use of a poor data structure and failure to ensure proper data analysis increase the discovery of type I and type II errors. Therefore, ensuring proper data analysis and a clear interpretation of results is crucial for the generation of information that will inform alternative controlling strategies for ticks. The fluctuating exchange rate makes it difficult for most researchers to conduct GWAS studies focused on studying the genetic architecture of tick resistance, since such studies require a large sample size that is expensive to develop, sustain and genotype.

2.11 Conclusions

When conducting GWAS, it is necessary to account for factors that affect the rate of discovering an association and control the rate of discovering spurious associations. GWASs have been phenomenally successful in the discovery of SNPs and candidate genes associated with tick resistance in various cattle breeds of different origins. Despite the success, more information is needed given that most GWASs on tick resistance in cattle are underpowered. This underlines the need for continuous data collection to enable larger and more powerful studies. For the studies that have been conducted using low-density markers, genotype imputation is the most appropriate cost-effective approach for GWAS for tick resistance in cattle. The availability of modest research populations, tools and funds are the current limitations of GWASs in developing countries. GWASs in African countries are

performed using SNP genotyping arrays developed using exotic breeds. Thus, the ascertainment bias effect leads to a low discovery rate of variants that influence the expression of the phenotype. This underscores the need for the inclusion of information from local breeds of different regions in the development of SNP genotyping arrays for cattle. This will facilitate a better understanding of variation in the breeds that are naturally adapted to the African production environments.

Author Contributions: A.M., N.M. (Ntanganedzeni Mapholi), and K.D. Conceptualized, designed, and sourced funding for the research. N.M. (Nelisiwe Mkize) and B.D coordinated the research. N.M. (Nelisiwe Mkize) collect the data for the manuscript and initially drafted the manuscript. A.M., K.D., B.D., N.M. (Ntanganedzeni Mapholi) critically revised the manuscript. N.M (Nelisiwe Mkize) and B.D. prepared the final draft of the manuscript. All authors have read and agreed to the published version of the manuscript

Funding: Agricultural research council, Stellenbosch University, University of South Africa, National Research Foundation, and Red Meat Research & Development South Africa, financially supported this work.

Acknowledgement: The authors are thankful to the Agricultural Research Council-AP, Stellenbosch University, University of South Africa, National Research Foundation, and Red Meat Research & Development South Africa for providing resources.

Conflicts of Interest: The authors declare that there is no conflict of interest.

2.12 References

[1] Cardoso FF, Gomes CCG, Sollero BP, Oliveira MM, Roso VM, Piccoli ML, Higa RH, Yokoo MJ, Caetano AR, Aguilar I. Genomic prediction for tick resistance in Braford and Hereford cattle. J Anim Sci 2015; 93:2693–705. https://doi.org/10.2527/jas.2014-8832.

[2] Muchenje V, Dzama K, Chimonyo M, Raats JG, Strydom PE. Tick susceptibility and its effects on growth performance and carcass characteristics of Nguni, Bonsmara and Angus steers raised on natural pasture. Animal 2008; 2:298–304. https://doi.org/10.1017/S1751731107001036.

[3] Ibelli AMG, Ribeiro ARB, Giglioti R [UNESP, Regitano LCA, Alencar MM, Chagas ACS, Paco AL [UNESP, Oliveira HN de [UNESP, Duarte JMS, Oliveira MCS. Resistance of cattle of various genetic groups to the tick Rhipicephalus microplus and the relationship with coat traits. Vet Parasitol 2012:425. https://doi.org/10.1016/j.vetpar.2011.11.019.

[4] Guerrero FD, Miller RJ, Pérez de León AA. Cattle tick vaccines: many candidate antigens, but will a commercially viable product emerge. Int J Parasitol 2012; 42:421–7. https://doi.org/10.1016/j.ijpara.2012.04.003.

[5] Castro-Janer E, Martins JR, Mendes MC, Namindome A, Klafke GM, Schumaker TTS. Diagnoses of fipronil resistance in Brazilian cattle ticks (Rhipicephalus (Boophilus) microplus) using in vitro larval bioassays. Vet Parasitol 2010; 173:300–6. https://doi.org/10.1016/j.vetpar.2010.06.036.

[6] Rajput ZI, Hu S, Chen W, Arijo AG, Xiao C. Importance of ticks and their chemical and immunological control in livestock. J Zhejiang Univ Sci B 2006; 7:912–21. https://doi.org/10.1631/jzus.2006.B0912.

[7] Mapholi NO, Marufu MC, Maiwashe A, Banga CB, Muchenje V, MacNeil MD, Chimonyo M, Dzama K. Towards a genomics approach to tick (Acari: Ixodidae) control in cattle: a review. Ticks Tick-Borne Dis 2014; 5:475–83. https://doi.org/10.1016/j.ttbdis.2014.04.006.

[8] Mapholi NO, Maiwashe A, Matika O, Riggio V, Bishop SC, MacNeil MD, Banga C, Taylor JF, Dzama K. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks Tick-Borne Dis 2016; 7:487–97. https://doi.org/10.1016/j.ttbdis.2016.02.005.

[9] Nyangiwe N, Yawa M, Muchenje V. Driving forces for changes in geographic range of cattle ticks (Acari: Ixodidae) in Africa: A review. South Afr J Anim Sci 2018; 48:829–41. https://doi.org/10.4314/sajas.v48i5.4.

[10] Acosta-Rodríguez R, Alonso R, Balladares S, Flores-Aguilar H, García-Vázquez Z, Gorodezky C. Analysis of BoLA class II microsatellites in cattle infested with Boophilus microplus ticks: Class II is probably associated with susceptibility. Vet Parasitol 2005; 127:313–21. https://doi.org/10.1016/j.vetpar.2004.10.007.

[11] Gasparin G, Miyata M, Coutinho LL, Martinez ML, Teodoro RL, Furlong J, Machado MA, Silva MVGB, Sonstegard TS, Regitano LCA. Mapping of quantitative trait loci controlling tick [Riphicephalus (Boophilus) microplus] resistance on bovine chromosomes 5, 7 and 14. Anim Genet 2007; 38:453–9. https://doi.org/10.1111/j.1365-2052.2007.01634.x.

[12] Machado MA, S Azevedo AL, Teodoro RL, Pires MA, CD Peixoto MG, de Freitas C, Prata MCA, Furlong J, da Silva MVG, Guimarães SE, et al. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (Bos taurus × Bos indicus). BMC Genomics 2010; 11:280. https://doi.org/10.1186/1471-2164-11-280.

[13] Sollero BP, Junqueira VS, Gomes CCG, Caetano AR, Cardoso FF. Tag SNP selection for prediction of tick resistance in Brazilian Braford and Hereford cattle breeds using

Bayesian methods. Genet Sel Evol 2017; 49:49. https://doi.org/10.1186/s12711-017-0325-2.

[14] Otto PI, Guimarães SEF, Verardo LL, Azevedo ALS, Vandenplas J, Soares ACC, Sevillano CA, Veroneze R, de Fatima A Pires M, de Freitas C, et al. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J Dairy Sci 2018; 101:11020–32. https://doi.org/10.3168/jds.2017-14223.

[15] Marufu M, Qokweni L, Chimonyo M, Dzama K. Relationships between tick counts and coat characteristics in Nguni and Bonsmara cattle reared on semiarid rangelands inSouth Africa. Ticks Tick-Borne Dis 2011; 2:172–7. https://doi.org/10.1016/j.ttbdis.2011.07.001.

[16] Ibeagha-Awemu EM, Kgwatalala P, Ibeagha AE, Zhao X. A critical analysis of disease-associated DNA polymorphisms in the genes of cattle, goat, sheep, and pig. Mamm Genome 2008; 19:226–45. https://doi.org/10.1007/s00335-008-9101-5.

[17] Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES. Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 2009; 21:2194–202. https://doi.org/10.1105/tpc.109.068437.

[18] Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, Yang J. 10 Years of GWAS Discovery: Biology, Function, and Translation. Am J Hum Genet 2017; 101:5–22. https://doi.org/10.1016/j.ajhg.2017.06.005.

[19] Sherva R, Farrer LA. Power, and Pitfalls of the Genome-Wide Association Study Approach to Identify Genes for Alzheimer's Disease. Curr Psychiatry Rep 2011;13:138–46. https://doi.org/10.1007/s11920-011-0184-4.

[20] Psifidi A, Banos G, Matika O, Desta TT, Bettridge J, Hume DA, Dessie T, Christley R, Wigley P, Hanotte O, et al. Genome-wide association studies of immune, disease and production traits in indigenous chicken ecotypes. Genet Sel Evol 2016; 48:74. https://doi.org/10.1186/s12711-016-0252-7.

[21] Patron J, Serra-Cayuela A, Han B, Li C, Wishart DS. Assessing the performance of genome-wide association studies for predicting disease risk. PloS One 2019; 14:e0220215. https://doi.org/10.1371/journal.pone.0220215.

[22] Wang Z, Chatterjee N. Increasing mapping precision of genomewide association studies: To genotype and impute, sequence, or both? Genome Biol 2017; 18. https://doi.org/10.1186/s13059-017-1255-6.

[23] Paes GP, Viana JMS, Silva FF e, Mundim GB. Linkage disequilibrium, SNP frequency change due to selection, and association mapping in popcorn chromosome regions

containing QTLs for quality traits. Genet Mol Biol 2016; 39:97–110. https://doi.org/10.1590/1678-4685-GMB-2015-0126.

[24] Amos CI. Successful design and conduct of genome-wide association studies. Hum Mol Genet 2007; 16 Spec No. 2:R220-225. https://doi.org/10.1093/hmg/ddm161.

[25] Greely HT. The uneasy ethical and legal underpinnings of large-scale genomic biobanks. Annu Rev Genomics Hum Genet 2007; 8:343–64. https://doi.org/10.1146/annurev.genom.7.080505.115721.

[26] Reed E, Nunez S, Kulp D, Qian J, Reilly MP, Foulkes AS. A guide to genome-wide association analysis and post-analytic interrogation. Stat Med 2015; 34:3769–92. https://doi.org/10.1002/sim.6605.

[27] Bermingham ML, Bishop SC, Woolliams JA, Pong-Wong R, Allen AR, McBride SH, Ryder JJ, Wright DM, Skuce RA, McDowell SW, et al. Genome-wide association study identifies novel loci associated with resistance to bovine tuberculosis. Heredity 2014; 112:543–51. https://doi.org/10.1038/hdy.2013.137.

[28] Sodeland M, Kent MP, Olsen HG, Opsal MA, Svendsen M, Sehested E, Hayes BJ, Lien S. Quantitative trait loci for clinical mastitis on chromosomes 2, 6, 14 and 20 in Norwegian Red cattle. Anim Genet 2011; 42:457–65. https://doi.org/10.1111/j.1365-2052.2010.02165.x.

[29] Lee B-Y, Lee K-N, Lee T, Park J-H, Kim S-M, Lee H-S, Chung D-S, Shim H-S, Lee H-K, Kim H. Bovine Genome-wide Association Study for Genetic Elements to Resist the Infection of Foot-and-mouth Disease in the Field. Asian-Australas J Anim Sci 2015; 28:166–70. https://doi.org/10.5713/ajas.14.0383.

[30] Xia J, Fan H, Chang T, Xu L, Zhang W, Song Y, Zhu B, Zhang L, Gao X, Chen Y, et al. Searching for new loci and candidate genes for economically important traits through gene-based association analysis of Simmental cattle. Sci Rep 2017; 7:42048. https://doi.org/10.1038/srep42048.

[31] Santiago GG, Siqueira F, Cardoso FF, Regitano LCA, Ventura R, Sollero BP, Souza MD, Mokry FB, Ferreira ABR, Torres R a. A. Genomewide association study for production and meat quality traits in Canchim beef cattle. J Anim Sci 2017; 95:3381–90. https://doi.org/10.2527/jas.2017.1570.

[32] Yue SJ, Zhao YQ, Gu XR, Yin B, Jiang YL, Wang ZH, Shi KR. A genome-wide association study suggests new candidate genes for milk production traits in Chinese Holstein cattle. Anim Genet 2017; 48:677–81. https://doi.org/10.1111/age.12593.

[33] Chen Z, Yao Y, Ma P, Wang Q, Pan Y. Haplotype-based genome-wide association study identifies loci and candidate genes for milk yield in Holsteins. PLOS ONE 2018; 13:e0192695. https://doi.org/10.1371/journal.pone.0192695.

[34] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet 2007; 81:559–75.

[35] Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics 2007; 23:1294–6. https://doi.org/10.1093/bioinformatics/btm108.

[36] Golden Helix | Citing our Software n.d. https://www.goldenhelix.com/products/citing.html (accessed October 13, 2021).

[37] Curtis RE, Kinnaird P, Xing EP. GenAMap: Visualization Strategies for Structured Association Mapping n.d.

[38] Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet 2012; 44:821–4. https://doi.org/10.1038/ng.2310.

[39] Misztal I, Lourenco D, Aguilar I, Legarra A, Vitezica Z. Manual for BLUPF90 family of programs n.d.:142.

[40] Hu Z-L, Park CA, Reecy JM. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. Nucleic Acids Res 2019; 47:D701–10. https://doi.org/10.1093/nar/gky1084.

[41] Vignal A, Milan D, SanCristobal M, Eggen A. A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol 2002; 34:275. https://doi.org/10.1186/1297-9686-34-3-275.

[42] Brumfield R, Beerli P, Nickerson D, Edwards S. Brumfield RT, Beerli P, Nickerson DA, Edwards SV.. The utility of single nucleotide polymorphisms in inferences of population history. Trends Ecol Evol 18: 249-256. Trends Ecol Evol 2003; 18:249–56. https://doi.org/10.1016/S0169-5347 (03)00018-1.

[43] Fan B, Du Z-Q, Gorbach DM, Rothschild MF. Development and Application of High-density SNP Arrays in Genomic Studies of Domestic Animals. Asian-Australas J Anim Sci 2010; 23:833–47. https://doi.org/10.5713/ajas.2010.r.03.

[44] Scheben A, Yuan Y, Edwards D. Advances in genomics for adapting crops to climate change. Curr Plant Biol 2016; 6:2–10. https://doi.org/10.1016/j.cpb.2016.09.001.

[45] Ganal MW, Polley A, Graner E-M, Plieske J, Wieseke R, Luerssen H, Durstewitz G. Large SNP arrays for genotyping in crop plants. J Biosci 2012; 37:821–8. https://doi.org/10.1007/s12038-012-9225-3.

[46] Lashmar SF, Muchadeyi FC, Visser C. Genotype imputation as a cost-saving genomic strategy for South African Sanga cattle: A review. South Afr J Anim Sci 2019; 49:262–80. https://doi.org/10.4314/sajas.v49i2.7.

[47] Nicolazzi EL, Biffani S, Biscarini F, Orozco Ter Wengel P, Caprera A, Nazzicari N, Stella A. Software solutions for the livestock genomics SNP array revolution. Anim Genet 2015; 46:343–53. https://doi.org/10.1111/age.12295.

[48] Schmid M, Bennewitz J. Invited review: Genome-wide association analysis for quantitative traits in livestock – a selective review of statistical models and experimental designs. Arch Anim Breed 2017; 60:335–46. https://doi.org/10.5194/aab-60-335-2017.

[49] Alqudah AM, Haile JK, Alomari DZ, Pozniak CJ, Kobiljski B, Börner A. Genome-wide and SNP network analyses reveal genetic control of spikelet sterility and yield-related traits in wheat. Sci Rep 2020; 10:2098. https://doi.org/10.1038/s41598-020-59004-4.

[50] Wang Y, Zhang F, Mukiibi R, Chen L, Vinsky M, Plastow G, Basarab J, Stothard P, Li C. Genetic architecture of quantitative traits in beef cattle revealed by genome wide association studies of imputed whole genome sequence variants: II: carcass merit traits. BMC Genomics 2020;21:38. https://doi.org/10.1186/s12864-019-6273-1.

[51] Yang N, Lu Y, Yang X, Huang J, Zhou Y, Ali F, Wen W, Liu J, Li J, Yan J. Genome Wide Association Studies Using a New Nonparametric Model Reveal the Genetic Architecture of 17 Agronomic Traits in an Enlarged Maize Association Panel. PLOS Genet 2014; 10:e1004573. https://doi.org/10.1371/journal.pgen.1004573.

[52] Fernando RL, Nettleton D, Southey BR, Dekkers JCM, Rothschild MF, Soller M. Controlling the Proportion of False Positives in Multiple Dependent Tests. Genetics 2004; 166:611–9. https://doi.org/10.1534/genetics.166.1.611.

[53] Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet Epidemiol 2008; 32:361–9. https://doi.org/10.1002/gepi.20310.

[54] Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med 1990; 9:811–8. https://doi.org/10.1002/sim.4780090710.

[55] De R, Bush WS, Moore JH. Bioinformatics challenges in genome-wide association studies (GWAS). Methods Mol Biol Clifton NJ 2014; 1168:63–81. https://doi.org/10.1007/978-1-4939-0847-9_5.

[56] Zeng P, Zhao Y, Qian C, Zhang L, Zhang R, Gou J, Liu J, Liu L, Chen F. Statistical analysis for genome-wide association study. J Biomed Res 2015; 29:285–97. https://doi.org/10.7555/JBR.29.20140007.

[57] Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, et al. Finding the missing heritability of complex diseases. Nature 2009; 461:747–53. https://doi.org/10.1038/nature08494.

[58] Moore JH, Asselbergs FW, Williams SM. Bioinformatics challenges for genome-wide association studies. Bioinforma Oxf Engl 2010; 26:445–55. https://doi.org/10.1093/bioinformatics/btp713.

[59] Sandoval-Motta S, Aldana M, Martínez-Romero E, Frank A. The Human Microbiome and the Missing Heritability Problem. Front Genet 2017; 8:80. https://doi.org/10.3389/fgene.2017.00080.

[60] Li C, Fu Y, Sun R, Wang Y, Wang Q. Single-Locus, and Multi-Locus Genome-Wide Association Studies in the Genetic Dissection of Fiber Quality Traits in Upland Cotton (Gossypium hirsutum L.). Front Plant Sci 2018; 9:1083. https://doi.org/10.3389/fpls.2018.01083.

[61] Zhang Y-M, Jia Z, Dunwell JM. Editorial: The Applications of New Multi-Locus GWAS Methodologies in the Genetic Dissection of Complex Traits. Front Plant Sci 2019; 10:100. https://doi.org/10.3389/fpls.2019.00100.

[62] Chu BB, Keys KL, German CA, Zhou H, Zhou JJ, Sobel E, Sinsheimer JS, Lange K. Iterative Hard Thresholding in GWAS: Generalized Linear Models, Prior Weights, and Double Sparsity. 2020. https://doi.org/10.1101/697755.

[63] Zhang Z, Ersoz E, Lai C-Q, Tod hunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, et al. Mixed linear model approach adapted for genome-wide association studies. Nat Genet 2010; 42:355–60. https://doi.org/10.1038/ng.546.

[64] Shenstone E, Cooper J, Rice B, Bohn M, Jamann TM, Lipka AE. An assessment of the performance of the logistic mixed model for analyzing binary traits in maize and sorghum diversity panels. PLOS ONE 2018; 13:e0207752. https://doi.org/10.1371/journal.pone.0207752.

[65] Wang S-B, Feng J-Y, Ren W-L, Huang B, Zhou L, Wen Y-J, Zhang J, Dunwell JM, Xu S, Zhang Y-M. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. Sci Rep 2016; 6:19444. https://doi.org/10.1038/srep19444.

[66] Zhang Y-W, Tamba CL, Wen Y-J, Li P, Ren W-L, Ni Y-L, Gao J, Zhang Y-M. MrMLM v4.0.2: An R Platform for Multi-locus Genome-wide Association Studies. Genomics Proteomics Bioinformatics 2020; 18:481–7. https://doi.org/10.1016/j.gpb.2020.06.006.

[67] Wen Y-J, Zhang Y-W, Zhang J, Feng J-Y, Dunwell JM, Zhang Y-M. An efficient multi-locus mixed model framework for the detection of small and linked QTLs in F2. Brief Bioinform 2018; 20:1913–24. https://doi.org/10.1093/bib/bby058.

[68] Wang S, Dvorkin D, Da Y. SNPEVG: a graphical tool for GWAS graphing with mouse clicks. BMC Bioinformatics 2012; 13:319. https://doi.org/10.1186/1471-2105-13-319.

[69] Turner SD. qqman: an R package for visualizing GWAS results using Q-Q and Manhattan plots. J Open Source Soft 2018; 3:731. https://doi.org/10.21105/joss.00731.

[70] Cook DE, Ryckman KR, Murray JC. Generating Manhattan Plots in Stata. Stata J 2013; 13:323–8. https://doi.org/10.1177/1536867X1301300206.

[71] Grace C, Farrall M, Watkins H, Goel A. Manhattan++: displaying genome-wide association summary statistics with multiple annotation layers. BMC Bioinformatics 2019; 20:610. https://doi.org/10.1186/s12859-019-3201-y.

[72] Haller T, Tasa T, Metspalu A. Manhattan Harvester, and Cropper: a system for GWAS peak detection. BMC Bioinformatics 2019; 20:22. https://doi.org/10.1186/s12859-019-2600-4.

[73] Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010; 26:2336–7. https://doi.org/10.1093/bioinformatics/btq419.

[74] Chittoor G, Haack K, Mehta NR, Laston S, Cole SA, Comuzzie AG, Butte NF, Voruganti VS. Genetic variation underlying renal uric acid excretion in Hispanic children: the Viva La Familia Study. BMC Med Genet 2017; 18:6. https://doi.org/10.1186/s12881-016-0366-3.

[75] Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, Hanrahan F, Pertea G, Van Tassell CP, Sonstegard TS, et al. A whole-genome assembly of the domestic cow, Bos taurus. Genome Biol 2009; 10:R42. https://doi.org/10.1186/gb-2009-10-4-r42.

[76] Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge, and principle: back to metabolism in KEGG. Nucleic Acids Res 2014; 42:D199–205. https://doi.org/10.1093/nar/gkt1076.

[77] Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4:44–57. https://doi.org/10.1038/nprot.2008.211.

[78] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13:2498–504. https://doi.org/10.1101/gr.1239303.

[79] Pickrell J, Clerget-Darpoux F, Bourgain C. Power of genome-wide association studies in the presence of interacting loci. Genet Epidemiol 2007; 31:748–62. https://doi.org/10.1002/gepi.20238.

[80] Ledur MC, Navarro N, and Pérez-Enciso M. Large-scale SNP genotyping in crosses between outbred lines: how useful is it? Heredity 2010; 105:173–82. https://doi.org/10.1038/hdy.2009.149.

[81] Schmid M, Wellmann R, Bennewitz J. Power, and precision of QTL mapping in simulated multiple porcine F2 crosses using whole-genome sequence information. BMC Genet 2018; 19:22. https://doi.org/10.1186/s12863-018-0604-0.

[82] Morin PA, Luikart G, Wayne RK, the SNP workshop group. SNPs in ecology, evolution, and conservation. Trends Ecol Evol 2004; 19:208–16. https://doi.org/10.1016/j.tree.2004.01.009.

[83] Burrow H, Mans B, Cardoso F, Birkett M, Kotze A, Hayes B, Mapholi N, Dzama K, Marufu M, Githaka N, et al. Towards a new phenotype for tick resistance in beef and dairy cattle: a review. Anim Prod Sci 2019; 59. https://doi.org/10.1071/AN18487.

[84] Mapholi NO, Maiwashe A, Matika O, Riggio V, Banga C, MacNeil MD, Muchenje V, Nephawe K, Dzama K. Genetic parameters for tick counts across months for different tick species and anatomical locations in South African Nguni cattle. Trop Anim Health Prod 2017; 49:1201–10. https://doi.org/10.1007/s11250-017-1336-2.

[85] Prayaga KC, Corbet NJ, Johnston DJ, Wolcott ML, Fordyce G, Burrow HM, Prayaga KC, Corbet NJ, Johnston DJ, Wolcott ML, et al. Genetics of adaptive traits in heifers and their relationship to growth, pubertal and carcass traits in two tropical beef cattle genotypes. Anim Prod Sci 2009; 49:413–25. https://doi.org/10.1071/EA08247.

[86] Burrow HM. Genetic Aspects of Cattle Adaptation in the Tropics, CAB International; 2014.

[87] Hong EP, Park JW. Sample size and statistical power calculation in genetic association studies. Genomics Inform 2012; 10:117–22. https://doi.org/10.5808/GI.2012.10.2.117.

[88] Kumar J, Pratap A, Solanki RK, Gupta DS, Goyal A, Chaturvedi SK, Nadarajan N, Kumar S. Genomic resources for improving food legume crops. J Agric Sci 2012; 150:289–318. https://doi.org/10.1017/S0021859611000554.

[89] Yano K, Yamamoto E, Aya K, Takeuchi H, Lo P-C, Hu L, Yamasaki M, Yoshida S, Kitano H, Hirano K, et al. Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. Nat Genet 2016; 48:927–34. https://doi.org/10.1038/ng.3596.

[90] Klein RJ. Power analysis for genome-wide association studies. BMC Genet 2007; 8:58. https://doi.org/10.1186/1471-2156-8-58.

[91] Park J-H, Wacholder S, Gail MH, Peters U, Jacobs KB, Cannock SJ, Chatterjee N. Estimation of effect size distribution from genome-wide association studies and

implications for future discoveries. Nat Genet 2010; 42:570–5. https://doi.org/10.1038/ng.610.

[92] Anderson CA, Peterson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc 2010; 5:1564–73. https://doi.org/10.1038/nprot.2010.116.

[93] Toe YY. Common statistical issues in genome-wide association studies: a review on power, data quality control, genotype calling and population structure. Curr Open Lip idol 2008; 19:133–43. https://doi.org/10.1097/MOL.0b013e3282f5dd77.

[94] Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genet Epidemiol 2000; 19 Supple 1:S36-42. https://doi.org/10.1002/1098-2272 (2000)19:1+<: AID-GEPI6>3.0.CO; 2-M.

[95] Wigginton JE, Cutler DJ, Abecasis GR. A Note on Exact Tests of Hardy-Weinberg Equilibrium. Am J Hum Genet 2005; 76:887–93?

[96] Slaking M. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. Nat Rev Genet 2008; 9:477–85. https://doi.org/10.1038/nrg2361.

[97] Bush WS, Moore JH. Chapter 11: Genome-wide association studies. Plops Compute Biol 2012; 8:e1002822. https://doi.org/10.1371/journal.pcbi.1002822.

[98] Flint-Garcia S, Thornberry J, Buckler E. Structure of Linkage Disequilibrium in Plants. Annu Rev Plant Biol 2003; 54:357–74. https://doi.org/10.1146/annurev.arplant.54.031902.134907.

[99] Falconer DS, Mackay TFC. Introduction to Quantitative Genetics. Subsequent edition. Harlow: Benjamin-Cummings Pub Co; 1996.

[100] Kolmodin R, Strindberg E, Madsen P, Jensen J, Jorjani H. Genotype by Environment Interaction in Nordic Dairy Cattle Studied Using Reaction Norms. Acta Agric Scand Sect — Anim Sci 2002; 52:11–24. https://doi.org/10.1080/09064700252806380.

[101] Moat RR, Tempelman RJ, Lopes PS, Aguilar I, Silva FF, Cardoso FF. Genotype by environment interaction for tick resistance of Hereford and Braford beef cattle using reaction norm models. Genet Sel Evol GSE 2016; 48:3. https://doi.org/10.1186/s12711-015-0178-5.

[102] Hayes BJ, Daetwyler HD, Goddard ME. Models for Genome × Environment Interaction: Examples in Livestock. Crop Sci 2016; 56:2251–9. https://doi.org/10.2135/cropsci2015.07.0451.

[103] Biegelmeyer P, Nizoli LQ, Cardoso FF, Dionello NJ. Aspectos da resistência de bovinos ao carrapato rhipicephalus (boophilus) microplus. Arch Zootec 2012; 61:1–11. https://doi.org/10.21071/az.v61i237.2954.

[104] Mulder HA, Bijma P. Effects of genotype × environment interaction on genetic gain in breeding programs1. J Anim Sci 2005; 83:49–61. https://doi.org/10.2527/2005.83149x.

[105] Sae-Lim P, Grøva L, Olesen I, Varona L. A comparison of nonlinear mixed models and response to selection of tick-infestation on lambs. PloS One 2017; 12:e0172711. https://doi.org/10.1371/journal.pone.0172711.

[106] Cooley PC, Clark RF, Folsom RE. Assessing Gene-Environment Interactions in Genome-Wide Association Studies: Statistical Approaches. Research Triangle Park (NC): RTI Press; 2014.

[107] Winham SJ, Biernacka JM. Gene-environment interactions in genome-wide association studies: current approaches and new directions. J Child Psychol Psychiatry 2013;54:1120–34. https://doi.org/10.1111/jcpp.12114.

[108] Marigorta UM, Gibson G. A simulation study of gene-by-environment interactions in GWAS implies ample hidden effects. Front Genet 2014; 5:225. https://doi.org/10.3389/fgene.2014.00225.

[109] Arnau-Soler A, Macdonald-Dunlop E, Adams MJ, Clarke T-K, MacIntyre DJ, Milburn K, Navrady L, Hayward C, McIntosh AM, Thomson PA. Genome-wide by environment interaction studies of depressive symptoms and psychosocial stress in UK Biobank and Generation Scotland. Transl Psychiatry 2019; 9:1–13. https://doi.org/10.1038/s41398-018-0360-y.

[110] Gali KK, Sackville A, Tafesse EG, Lachagari VBR, McPhee K, Hybl M, Mikić A, Smýkal P, McGee R, Burstin J, et al. Genome-Wide Association Mapping for Agronomic and Seed Quality Traits of Field Pea (Pisum sativum L.). Front Plant Sci 2019; 10:1538. https://doi.org/10.3389/fpls.2019.01538.

[111] Eltaher S, Baenziger PS, Belamkar V, Emara HA, Nower AA, Salem KFM, Alqudah AM, Sallam A. GWAS revealed effect of genotype × environment interactions for grain yield of Nebraska winter wheat. BMC Genomics 2021; 22:2. https://doi.org/10.1186/s12864-020-07308-0.

[112] Turner S, Armstrong LL, Bradford Y, Carlson CS, Crawford DC, Crenshaw AT, de Andrade M, Doheny KF, Haines JL, Hayes G, et al. Quality control procedures for genome-wide association studies. Curr Protoc Hum Genet 2011; Chapter 1:Unit1.19. https://doi.org/10.1002/0471142905.hg0119s68.

[113] Howie BN, Donnelly P, Marchini J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. PLOS Genet 2009; 5:e1000529. https://doi.org/10.1371/journal.pgen.1000529.

[114] Naj AC. Genotype Imputation in Genome-Wide Association Studies. Curr Protoc Hum Genet 2019; 102:e84. https://doi.org/10.1002/cphg.84.

[115] Spencer CCA, Su Z, Donnelly P, Marchini J. Designing Genome-Wide Association Studies: Sample Size, Power, Imputation, and the Choice of Genotyping Chip. PLOS Genet 2009; 5:e1000477. https://doi.org/10.1371/journal.pgen.1000477.

[116] García-Ruiz A, Ruiz-Lopez FJ, Wiggans GR, Van Tassell CP, Montaldo HH. Effect of reference population size and available ancestor genotypes on imputation of Mexican Holstein genotypes. J Dairy Sci 2015; 98:3478–84. https://doi.org/10.3168/jds.2014-9132.

[117] Holm H, Gudbjartsson DF, Sulem P, Masson G, Helgadottir HT, Zanon C, Magnusson OT, Helgason A, Saemundsdottir J, Gylfason A, et al. A rare variant in MYH6 is associated with high risk of sick sinus syndrome. Nat Genet 2011; 43:316–20. https://doi.org/10.1038/ng.781.

[118] Pistis G, Porcu E, Vrieze SI, Sidore C, Steri M, Danjou F, Busonero F, Mulas A, Zoledziewska M, Maschio A, et al. Rare variant genotype imputation with thousands of study-specific whole-genome sequences: implications for cost-effective study designs. Eur J Hum Genet EJHG 2015; 23:975–83. https://doi.org/10.1038/ejhg.2014.216.

[119] Quick C, Anugu P, Musani S, Weiss ST, Burchard EG, White MJ, Keys KL, Cucca F, Sidore C, Boehnke M, et al. Sequencing and imputation in GWAS: Cost-effective strategies to increase power and genomic coverage across diverse populations. Genet Epidemiol 2020; 44:537–49. https://doi.org/10.1002/gepi.22326.

[120] Sargolzaei M, Chesnais J, Schenkel F. FImpute-An efficient imputation algorithm for dairy cattle populations. J Dairy Sci 2011; 94:421.

[121] Tabor AE, Ali A, Rehman G, Rocha Garcia G, Zangirolamo AF, Malardo T, Jonsson NN. Cattle Tick Rhipicephalus microplus-Host Interface: A Review of Resistant and Susceptible Host Responses. Front Cell Infect Microbiol 2017; 7:506. https://doi.org/10.3389/fcimb.2017.00506.

[122] Turner LB, Harrison BE, Bunch RJ, Neto LRP, Li Y, Barendse W, Turner LB, Harrison BE, Bunch RJ, Neto LRP, et al. A genome-wide association study of tick burden and milk composition in cattle. Anim Prod Sci 2010; 50:235–45. https://doi.org/10.1071/AN09135.

[123] Regitano L, Prayaga K. Ticks, and tick-borne diseases in cattle. Breed Dis Resist Farm Anim Third Ed 2010:295–314.

[124] Henshall JM, Henshall JM. A genetic analysis of parasite resistance traits in a tropically adapted line of Bos taurus. Aust J Agric Res 2004; 55:1109–16. https://doi.org/10.1071/AR03085.

[125] Budeli MA, Nephawe KA, Norris D, Selapa NW, Bergh L, Maiwashe A. Genetic parameter estimates for tick resistance in Bonsmara cattle. South Afr J Anim Sci 2009; 39. https://doi.org/10.4314/sajas.v39i4.51125.

[126] Shyma KP, Gupta JP, Singh V. Breeding strategies for tick resistance in tropical cattle: a sustainable approach for tick control. J Parasit Dis 2013; 39. https://doi.org/10.1007/s12639-013-0294-5.

[127] Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. Nat Rev Genet 2009; 10:241–51. https://doi.org/10.1038/nrg2554.

[128] Aschard H, Lutz S, Maus B, Duell EJ, Fingerlin T, Chatterjee N, Kraft P, Van Steen K. Challenges and Opportunities in Genome-Wide Environmental Interaction (GWEI) studies. Hum Genet 2012; 131:1591–613. https://doi.org/10.1007/s00439-012-1192-0.

[129] Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits, and limitations of genome-wide association studies. Nat Rev Genet 2019; 20:467–84. https://doi.org/10.1038/s41576-019-0127-1.

3 Chapter 3:

Genetic parameter estimates for tick resistance in F_2 Angus x Nguni cattle artificially infested with *Amblyomma hebraeum* ticks

Abstract: This study aimed to (1) assess the effect of coat characteristics, growth traits, skin thickness, sex, and year of infestation on tick count, and (2) estimate genetic parameters for tick count, growth traits, and skin thickness in an F_2 Angus x Nguni crossbred cattle population. A total of 266 F_2 Angus x Nguni crossbred animals were produced between 2013 and 2020. The animals were artificially infested with *Amblyomma hebraeum* and assessed for tick load in batches. Data on factors that could potentially affect the level of tick load on cattle were collected and assessed for their association with tick count. Variance components and heritability were estimated by fitting the sire model using ASREML software. Live weights ($p< 0.05$) and years of infestation ($p<0.001$) had significant effects on the level of tick load. The heritability estimate for tick count was 0.21±0.22. The heritability estimates were 0.36 ± 0.26, 0.32±0.14, and 0.32±0.19 for birth weight, weaning weight, and yearling weight, respectively. The heritability estimate for skin thickness was 0.40±0.15. These heritability estimates indicate the presence of sufficient genetic variation for these traits. Therefore, genetic improvement through selective breeding is a viable option for these traits.

Keywords: genetic variation, heritability, coat characteristics, year of infestation

3.1 Introduction

Cattle production plays a key socio-economic role in many countries especially those in the developing world. It contributes to food security as a source of milk and meat. Over 80% of global cattle production occurs in regions which have high prevalence of ticks and their associated tick-borne diseases [1]. Ticks and tick-borne diseases (TTBDs) are a global problem and considered one of the major challenges to livestock production and health [2–4]. Economic losses associated to TTBDs in cattle production are estimated to be USD $20 – 30 billion per/year worldwide [4]. Economic losses due to TTBDs in the South African livestock industry has been estimated to be more than R500 million per year [3]**.** In South Africa*, Amblyomma hebraeum* is the most prevalent veterinary important tick species, which is notoriously known for infesting livestock and wildlife [5,6]. Ticks cause huge economic losses, which can be reduced by implementing effective control measures [7–9]. Presently, acaricides are used to control ticks; however, these chemicals are continuously losing efficacy because

of their intensive use and the development of resistance to chemicals by ticks [7,10]. Additionally, the traces of acaricide residuals in food products of cattle origin violates food security policies and raises public health concerns [11]. This underlines the need for alternative tick control strategies.

According to Phocas et al. [12], breeding for disease resistance is a promising route to reduce use of antibiotics, vaccines, and pesticides in animal husbandry as well as to ensure animal welfare, which is currently a growing concern. The use of the animal's natural host resistance to tick infestation, through selective breeding, has been advocated as a feasible approach to fight tick burden [13,14]. Host resistance is defined as the animal's ability to limit the level of tick burden and a defensive immune system to fight tick-borne diseases [15]. Natural host resistance to ticks is heritable and differs among different cattle breeds and involves the interaction of many factors [7,16]. Natural tick resistance is responsive to selection, therefore selective breeding for this trait could potentially reduce the costs of tick control [13,17]. Moreover, this could benefit animal welfare, reduce environmental contaminations, toxicity to other biota and the traces of acaricide residues in food products [14]. The estimation of genetic parameters related to tick count will generate information useful in establishing breeding programs. Tick count is a trait that is also affected by biological characteristics such as coat thickness, color, hair type, and weight of an animal [18]. Therefore, the influence of these factors needs to be taken into consideration when estimating the genetic parameters [14,16]. There are studies that have reported on the differences in tick load observed when animals are raised under similar environmental conditions [19,20]. Such differences can be attributed to the physiological and genetic properties of the animals.

In South Africa, most beef production systems use exotic breeds due to their superior growth performance. However, exotic breeds are known to be highly susceptible to tick infestations and vulnerable to harsh conditions associated with climate instabilities. Local indigenous breeds such as Nguni cattle are known for their tolerance to tick infestations and for withstanding harsh environmental conditions [21,22]. It has been suggested that the establishment of crossbreeding programs that will cross indigenous adapted breeds with exotic breeds could promote the sustainability of beef production in South Africa [23–25]. The Agricultural Research Council - Animal Production campus established a crossbreeding project, where they crossed Angus x Nguni breeds and assessed them for tick resistance and growth performance. Although studies have been conducted on genetic parameters for tick resistance [26,27] and growth performance [28–31], the genetic parameters of these traits in Angus x Nguni crossbreeds have not been estimated. Therefore, the objective of this study

was to estimate the genetic parameters for tick count in the F_2 Angus x Nguni cattle population. The study also estimated genetic parameters for growth traits (birth weight, weaning, and yearly weight) and skin thickness in this population.

3.2 Materials and Methods

3.2.1 Study area and period

The study was carried out from 2011 to 2020 at the South African Agricultural Research Council-Animal Production campus (ARC-AP) in Irene (25°53′59.6″S, 28°12′51.6″E). The area experiences four seasons, autumn (March to May), winter (June-August), spring (September to November) and summer (December to February).

3.2.2 Experimental population

An experimental population consisting of 5 sire families was established by crossing Nguni cows with Angus bulls using artificial insemination to establish an F_1 population in 2011 (Figure3.1). Briefly, Nguni cows were subjected to oestrus synchronization and artificially inseminated with frozen and thawed Angus semen. The F_1 population was inter-crossed to produce the F_2 generation. Prior to inter-crossing, parentage verification was conducted to prevent within-family intercrossing. The first batch of the F_2 population (Figure 3.2) was generated in 2013 and subsequent batches were produced yearly until 2020. In total, 266 F_2 animals were produced consisting of 58% females and 42% males. The F_2 population batches were subjected to artificial tick infestation and phenotypic data were recorded for tick count, gender, year of infestation, season of infestation, skin thickness, coat colour pattern, and live weight at birth, weaning and yearling.

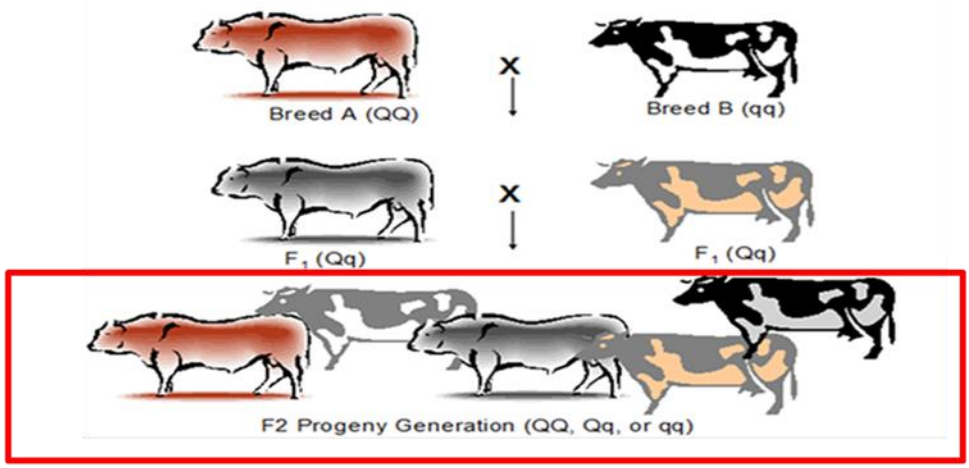

Figure 3.1: Schematic representation of the F_2 design (A= Angus and B= Nguni). *QQ, Qq and qq represent the expected genotypes for F_2 population assuming that the parental breeds were fixed for the alternative alleles (Q and q) at the tick count locus.*

Figure 3.2: Assorted colour patterns in the F_2 Angus x Nguni cattle.

3.2.3 Tick species

The tick species of interest used in the study was *Amblyomma hebraeum*. *Amblyomma hebraeum* larvae were sourced from the Agricultural Research Council - Onderstepoort Veterinary Research laboratory of South Africa. Unfed tick larvae were sourced from the laboratory, to ensure the absence of disease contamination. In preparation for tick infestation, 100 unfed larvae were housed in vials, which were then placed at room temperature (25°C) to create a normal environment for the ticks.

3.2.4 Tick infestation and counting

Prior to tick infestation, animals were not exposed to acaricides and vaccines. *Amblyomma hebraeum* unfed larvae were used to artificially infest 266 F_2 animals for a period of 10 days during different seasons. The infestation trials were conducted in batches, from year 2014 until 2020 (Table 3.1). Artificial infestation was performed using Calico bags which were attached on the dorsal region of the animal using Alcolin contact adhesive glue (Figure 3.3).

Table 3. 1: The F_2 Angus x Nguni cattle batches according to the year of infestation and season.

Batch	N	Year of infestation	Season
1	12	2014	Autumn
2	19	2015	Autumn
3	17	2016	Winter
4	29	2017	Summer
5	26	2017	Winter
6	65	2018	Autumn
7	48	2019	Autumn
8	50	2020	Winter

Figure 3.3: Schematic representation of the placement of the calico tick bags. **3.3A** shows the attachment of a bag prior to infestation and **3.3B** shows a closed attached bag after infestation in the F_2 Angus x Nguni cattle.

To attach the bag, the animals were shaved on the dorsal region using heavy-duty electric clippers (Legend®, Lister, South Africa), a day before the trial to allow the glue to dry out (Figure 3.3), and animals to cool down from the stress caused by handling. On the day of infestation, a vial containing 100 unfed tick larvae was opened and placed inside each tick

bag, which was then closed using a castration ring to ensure the ticks did not escape. Tick count was conducted on the 11th day after infestation and the animals were sprayed with acaricides to control the ticks and prevent environmental contamination. The year of infestation was used to create contemporary groups. In year 2017, two tick infestation activities were carried out in winter and spring.

3.2.5 Growth traits

The growth traits of interest were birth, weaning, and yearling weights. Cattle were weighed using a Taltec digital livestock scale at birth, weaning (at 6 months), yearling (at the age of 12 months), and upon infestation at the age of 15 months.

3.2.6 Coat characteristics

Skin thickness is a coat characteristic that can be measured in two ways that involve measuring skin biopsies and skinfold thickness using a calliper on the region of interest. The measuring of skin biopsies is the most accurate measure compared to skinfold thickness. However, it was not preferred from a welfare point of view since it subjects the animals to pain and stress. Therefore, the current study used skinfold thickness approach to assess the effect of skin thickness on the level of tick count on F_2 Angus x Nguni cattle. The skinfold thickness was measured following a method explained by Marufu et al. [13]. The skin thickness was measured on the midside area of the animal (just caudal to the 13th rib about 20 cm below the dorsal line) [13]. Skin thickness in this position is known to be uniform [32]. Raw coat colour patterns were recorded and then categorised (Table 3.2).

Table 3.2: Cattle coat patterns categories of the F_2 Angus x Nguni population.

Color pattern	**N**
Black	91
Black and other colors	82
White	1
White and other colors	38
Brown	22
Brown and other colors	23
Red	6
Red others	1

Grey	1

3.2.7 Statistical analyses

Tick count data did not follow a normal distribution (Figure 3.4A) and it was transformed using $\log_{10}(x+1)$ to confer normality (Figure 3.4B). The transformed tick count data were used as a response variable for further analysis. Analyses of variance (ANOVA) were conducted using PROC GLM procedure of SAS program [33] to assess the effect of sex, year of infestation, body weight (birth, weaning, yearling, and upon infestation) and coat characteristics using model (1). The fixed effects considered were sex and contemporary group to assess growth traits (i.e. birth, weaning, and yearly weights)

$$y_{ijkl} = \mu + Y_i + SEX_k + \beta(Wt) + e_{ikl} \quad (1)$$

y_{ijkl} is the log-transformed tick count, μ is the overall mean, Y_i is the effect of the i^{th} year of infestation, SEX_k is the effect of the k^{th} sex, β is the regression coefficient of the effect of weight at infestation, Wt is weight at infestation and e_{ikl} is the residual error.

3.2.8 Variance component estimation

The data were edited to remove outliers and the final data is presented in Table 3.3. The traits that were analysed were the log-transformed tick count, birth, weaning and yearling weights, skin thickness. For tick count and skin thickness, contemporary groups were created by concatenating year and season of infestation, while year and season of birth were concatenated to create contemporary groups for birth weight. Weaning and yearling years were considered in creating the contemporary group for weaning and yearling weights, respectively.

Table 3. 3*: Structure of the data analysed.*

Trait	Animals	Sires	Dams	Contemporary groups
Tick count	209	7	61	8
Birth weight	158	7	48	16
Weaning weight	211	7	61	7
Yearling weight	211	7	61	6
Skin thickness	209	7	61	8

Phenotypic and residual variance were estimated by fitting a sire model in ASREML software [34]. The following sire model was used for all the traits:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zs} + \mathbf{e} \quad (2)$$

Where; **y** is the vector of observations, **b** is the vector of fixed effects (e.g. contemporary group, sex, and growth trait); **s** is a vector of random sire additive genetic effects; **e** is the vector of residuals effects unique to each observation and **X** and **Z** are incidence matrices relating the fixed and random effects, respectively to **y**. A normal distribution was assumed for random effects, with animal ~N (0, $\mathbf{A}\sigma^2_s$), phenotypic variance ~N (0, $\mathbf{I}\sigma^2_p$), residual ~N (0, $\mathbf{I}\sigma^2_e$), where **A** is the numerator relationship matrix and **I** is an identity matrix, σ^2_s, σ^2_p and σ^2_e are the sire, phenotypic and residual variances, respectively. The relationship matrix was constructed using the pedigree. The narrow-sense heritability (h^2) was calculated as follows:

$$\mathbf{h^2} = (\mathbf{4\sigma_s^2})/(\mathbf{1} + \mathbf{\sigma_s^2}) \quad (3)$$

3.3 Results and discussion

3.3.1 Descriptive statistics

Tick count data were not normally distributed and most animals had tick count observation less than 50 ticks, which were semi-engorged (Figure 3.4A). The variability in tick count could be a result of differences in genetic, immunological, and physiological components of the individual animals in the study population. Figure 3.4B depicts the log-transformed tick count.

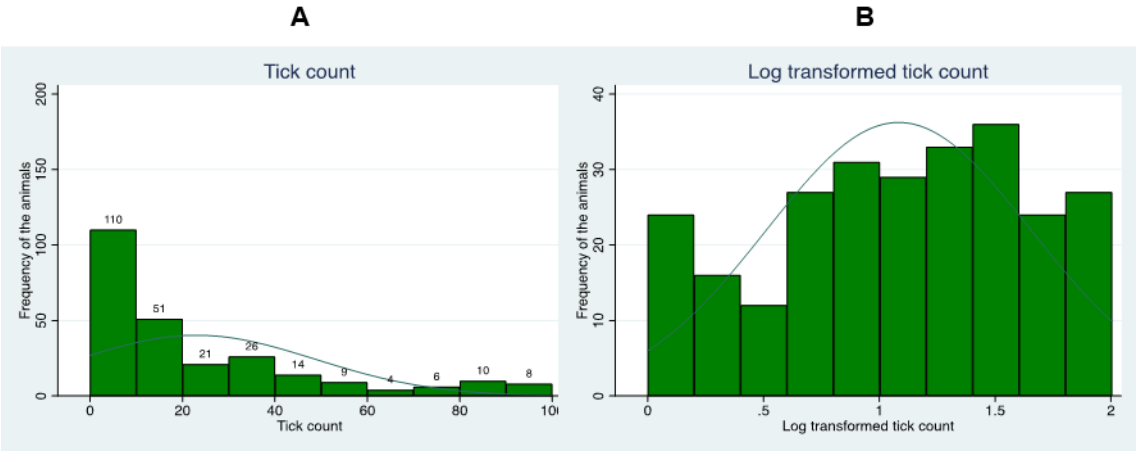

Figure 3.4 A: The raw distribution of tick count and **3.4 B** The log-transformed tick count data for the F_2 population.

Descriptive statistics of traits considered in this study are presented in Table 3.4. Furthermore, live body weight and skin thickness were subjected to normality testing and they were found to be normally distributed as depicted in Figure 3.5.

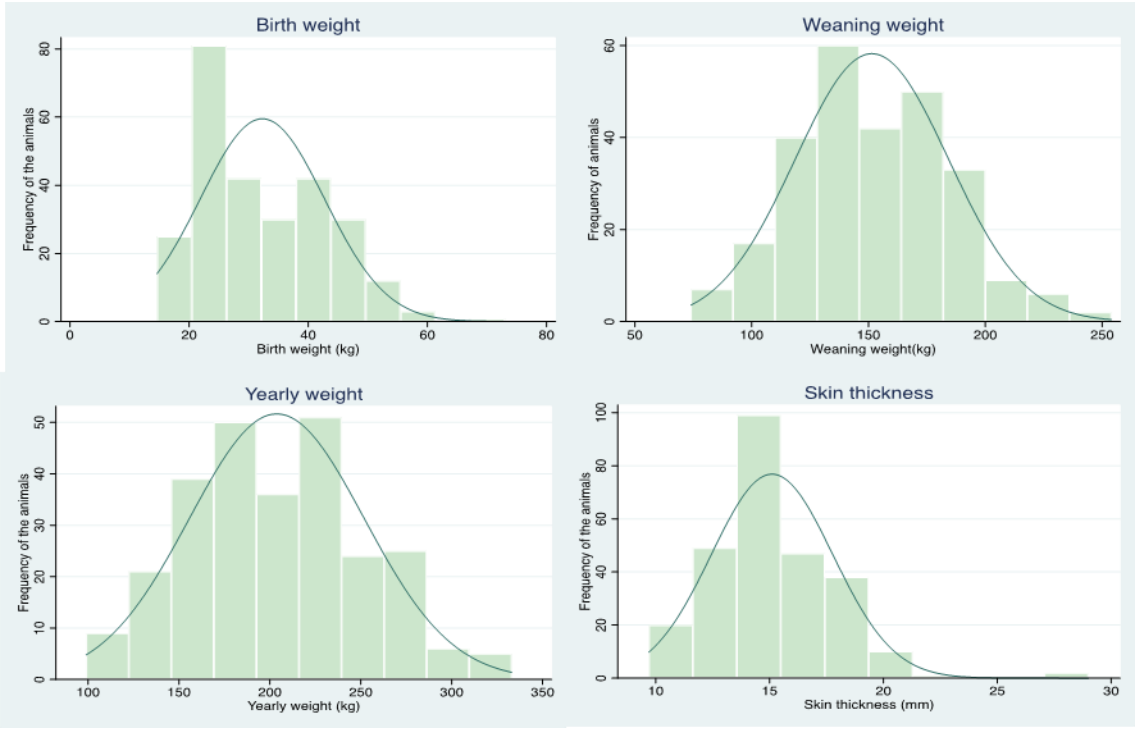

Figure 3.5: Frequency distribution for birth, weaning and yearling weights, and skin thickness for the 266 F_2 Angus x Nguni cattle.

Table 3. 4: *Descriptive statistics for the traits analysed in the F_2 Nguni X Angus cattle after some animals were removed.*

Trait	N	Mean	Min	Max	SD
Log-transformed tick count	209	1.06	0.00	2.00	0.58
Birth weight (kg)	148	28.11	17.60	40.00	6.16
Weaning weight (kg)	211	160.86	120	254	27.27
Yearling weight (kg)	211	214.84	150	333	42.53
Skin thickness(mm)	266	15.21	9.70	29.00	2.70

Table 3.5: Analysis of variance for tick count, growth traits and skin thickness.

TRAIT	Y_INF	S_INF	SEX	STHICK	CC	WWT	YWT
TCOUNT	<.0001***	NS	NS	NS	NS	<.0001***	0.0042*
STHICK	-	-	0.001***	-	-	<.0001***	0.0050*
BWT	-	-	0.001***	-	-	-	-
WWT	-	-	0.001***	-	-	-	-
YWT	-	-	0.001***	-	-	-	-

*** P<0.001; ** P<0.01; * P≤0.05; NS- Not significant, TCOUNT-tick count, BWT- birth weight, WWT- weaning weight, YWT-yearling weight, STHICK-skin thickness, Y_INF-year of infestation, S_INF-season of infestation, CC- coat colour, (-)- Not tested

3.3.2 The effect of coat characteristics

The results from the current study showed that skin thickness had no effect on the level of tick count. A similar study by Marufu et al. [13] reported no statistical significant (P > 0.05) effect of skin thickness on tick load in Bonsmara and Nguni cattle breeds. Skin thickness affected tick count in cattle in other studies [26, 28, 29]. Inconsistent findings observed across these studies may be due to different study designs, environments, tick species of interest, anatomical site where skin thickness was measured and breed differences.

Coat colour patterns did not affect the level of tick count in the current study (P>0.05). These findings agree with results observed by Nwachukwu et al. [14]. Marufu et al. [13] observed that coat colour has a statistically significant (P<0.05) effect on tick count. Although there was limited information on the effect of coat characteristics on tick count, research has shown that coat characteristics play a key role in maintaining the level of tick load on the host [35].

It is believed that light colored animals can be easily recognized by ticks under natural infestation [36]. However, when the animals are artificially infested the light colour allows easy recognition of ticks by predators such as wasps. The inconsistencies observed on the effect of coat characteristics could be due to the different methods used in different studies to measure skin thickness and to categorize coat colour patterns. Therefore, standardized

methods to measure skin thickness and to categorize coat colour patterns are required so that the effect of coat characteristics on tick resistance can be accurately evaluated in cattle [13].

3.3.3 The effect of live weight on tick count

Most studies investigating tick burden in cattle do not consider the influence of live weight on tick count [26-27]. In the current study, the influence of live weight on tick count was evaluated. Tick count was significantly affected by weaning ($P<0.05$) and yearling weights ($P<0.001$). It was observed that for every kg increase in weaning weight, tick count increased by 0.019 ± 0.006. Conversely, there was a decrease of 0.023 ± 0.006 in tick count when yearling weight increased by 1 kg. This indicates that heavier cattle at weaning are associated with higher tick loads, while animals with higher yearling weights have fewer ticks. These observations may be attributed to growth rate and nutritional status of the calves. The influence of nutritional status and growth rate on the magnitude of tick count was not assessed in the current study. According to O'kelly et al. [37], "little attention is being paid to the relationship between growth rate, nutritional status of the host, and the magnitude of tick burden", this statement applied to the current study. To date, little is known about the influence of nutrition and growth rate on the magnitude of tick count. Therefore, this knowledge gap suggest the need for future studies to investigate the relationship between nutrition, growth rate and tick count to provide better understanding, which could be used to develop strategies to control ticks in beef production.

Katiyatiya et al. [38], considered age and reported that age has a positive and significant genetic correlation with tick load. It has been observed that young animals tend to experience lower tick load as compared to older animals [38], which may be due to the differences in the effectiveness of the immunity between the two groups. Heavier animals are more compromised immunologically compared to their counterpart [39]. It has been previously reported that the weight of an animal was positively correlated with the age, which was subsequently related to the ability of the immune system of the animal to repel or allow ticks to feed [40]. In this study, age was not considered. It has been suggested that old animals emit high quantities of carbon dioxide, which enhance the attraction of ticks to the host [41], however, this does not apply to the current study because artificial infestation was used. It is recommended that animals of different ages should be assessed to evaluate the effect of age in future studies.

3.3.4 The effect of sex

Sex did not significantly affect (P>0.05) tick count. This could be due to the early age of animals at which hormonal differences were not expressed. Yessinou et al. [41] reported that males were more infested than females animals. The difference could be because their studies were conducted using cattle of different ages. The findings from this study were consistent with results reported by Chartier et al. [42], but are contrary to findings by Gharbi & Darghouth [43]. The production of carbon dioxide was regarded as the primary factor for the differences in tick load between males and females. It was suggested that males produce more carbon dioxide than females, which contributes to more attraction of ticks as noted by Yessinou et al. [41]. Testosterone hormone has also been reported to play a key role in the level of TTBDs in animals, explaining the difference in tick load between females and males. According to Hughes et al. [44], testosterone hormone reduced both innate and acquired resistance to tick feeding in rodents.

3.3.5 The effect of year of infestation

The year of infestation significantly affected (P<0.001) tick count. Tick count for the year 2014 was significantly different from other years of infestation (Table 3.5). The differences among years of infestation may be attributed to different climatic conditions during the year in which the trials were conducted. Some trials were conducted during the drought period, which could have affected the homeostasis of the animals and the behaviour of ticks. According to Monyela [45], South Africa experienced a long drought in summer of 2014/2015 and in 2015/2016. *Amblyomma hebraeum* larvae prefer cold and dry environment, therefore, the variation in climatic conditions throughout the study period could have affected the behaviour of the ticks, and tick loads [6]. The differences contributed by the year of infestation suggest the need to incorporate climate data when studying the tick count. Year and season of infestation were concatenated to create contemporary groups (Table 3.6), which significantly affected (P<0.001) tick count.

Table 3.6: Least Square Means (LSM) for year of infestation.

Year of infestation	Season	LSMeans ± SE
2014	Autumn	$1.70^{a} \pm 0.18$
2015	Autumn	$1.36^{a} \pm 0.15$
2016	Winter	$0.94^{b} \pm 0.14$
2017	Summer	$1.46^{a} \pm 0.12$
2017	Winter	$1.64^{a} \pm 0.11$
2018	Autumn	$0.69^{b} \pm 0.10$
2019	Autumn	$0.78^{b} \pm 0.09$
2020	Winter	$0.99^{b} \pm 0.09$

LSMeans with the same superscript are not significantly different

3.3.6 The effect of season of infestation

The season of infestation did not significantly (P>0.05) influence the tick count in F_2 Angus x Nguni cattle (Table 3.5). The least square means for the season of infestation also showed no significant differences (P>0.05) amongst the three seasonal periods (Table 3.7). The findings in the current study are in agreement with a study by Kemal et al. [40], which reported a non-significant effect of season on tick count in cattle of Southern Ethiopia. In the current study, the season of infestation alone did not affect tick count, however when it was concatenated with the year of infestation (Table 3.6 and 3.7), a significant effect was observed. Significant effect of season on the level of tick count has been reported by Nwachukwu et al. [14], Mekonnen et al. [46], and Mohamed et al. [47]. Furthermore, the effect of season on tick infestation in cattle have been further validated by studies which discovered different quantitative trait loci association with the level of tick infestation during wet and dry season [5,48–51]. The effect of season may be prompted by the preferred environmental conditions by the tick species of interest. *Amblyomma hebearum* is normally found in the coastal areas of South Africa, because the regions are characterized by stable temperature conditions, seasonal rainfalls and high grasslands [13,52,53]. The current study was conducted in the inland region, which exhibits different temperature conditions compared to coastal region.

However, the mode of infestation was artificial; it could have contributed to how the tick behaved upon infestation.

Table 3.7: Least Square Means for season of infestation.

Season	LSMeans ± SE
Summer	$1.06^{a} \pm 0.20$
Autumn	$0.67^{a} \pm 0.11$
Winter	$1.36^{ab} \pm 0.08$

LSMeans with the same superscript are not significantly different.

3.3.7 Heritability estimates

Heritability estimates for the log-transformed tick count, birth, weaning, and yearling weights and skin thickness are presented in Table 3.8.

Table 3.8: Heritability estimates (±SE) for tick count, birth weight, weaning weight, yearling weight and skin thickness.

Trait	Heritability estimate ± SE
Log-transformed tick count	0.21±0.22
Birth weight	0.36±0.26
Weaning weight	0.32±0.14
Yearling weight	0.32±0.19
Skin thickness	0.40±0.15

3.3.7.1 Tick count

Heritability estimate for tick count was 0.21±0.22 in the current study, indicating that this trait has low to moderate heritability. The associated standard error was high, which could be attributed to the small population size used in this study. Heritability estimate for tick count was comparable to the findings from similar studies which reported heritability estimates ranging from 0.05 to 0.58 in Bonsmara, and Nguni cattle [26,54,55]. Some of these studies were conducted in South Africa using purebred Bonsmara with heritability estimates ranging from 0.01 to 0.08 [26], and purebred Nguni cattle with heritability estimates ranging from 0 to 0.89) [54].

Heritability estimate in the current study was higher than estimates from the previous studies [56–58], which ranged from 0.09 to 0.15. Recent studies by Mapholi et al. [5,54] reported heritability estimates ranging from 0.00 to 0.89 in Nguni cattle, where animals were assessed for different tick species during different months. Furthermore, Mapholi et al. [5,54] estimated heritability for *Amblyomma hebraeum* during different months and reported heritability estimates ranging from 0.00 to 0.28. The authors concluded that the genetic variation observed was sufficient to warrant the improvement of tick resistance through selection [54].

The inconsistency in heritability estimates between the current and previous studies could be attributed to many factors, which include; breed, tick species, mode and region of infestation, lack of consistency in the protocols used to assess tick loads, population size, failure to account for the immune status of the animal, and the seasons in which the studies were conducted. It has been reported previously that sex, skin thickness, coat colour patterns and season affect the level of tick load [5,13,54] which in turn affect estimates of heritability for tick count. Despite these observations, the heritability estimate obtained in the current study indicates that this trait could be improved through selection. However, more studies should be conducted using standardized protocol to measure tick count. Additionally, animal's immune status should be considered in the statistical analysis of tick count as recommended by Porto Neto et al. [1], as it plays a key role in the expression of tick resistance.

3.3.7.2 Birth weight

In cattle, previous studies have associated birth weight with survivability, disease incidence, reproductive performance, and milk production [59,60]. The heritability estimate for birth weight in the current study was 0.36 ± 0.26 (Table 3.8). This heritability estimate is slightly lower than the estimates reported by Yin and Kong [61] (0.47), and Beyoda et al. [62] (0.44). Ossa et al. [63], and Boligon et al. [64] reported lower estimates of heritability of 0.17 and 0.25 respectively.

The differences in heritability estimates across studies could be attributed to different environmental conditions, breed, farm management strategies, and the model used for statistical analysis.

3.3.7.3 Weaning weight

The heritability estimate for weaning weight in the current study was 0.32±0.14, indicating a moderate heritability. This estimate is in agreement with findings from previous beef cattle

studies [65–67]. As has been done in previous studies, maternal heritability has been estimated in conjunction with direct heritability for growth traits. In this study, maternal heritability was not estimated because of the limited sample size, which influenced the statistical model used. It has been noted that failure to account for maternal effects on growth traits could inflate direct heritability estimates [31]. The differences observed from these studies could be attributed to factors such as herd management, breed composition, sample size, procedures to measure the phenotype, and the statistical model used.

3.3.7.4 Yearling weight

The heritability estimate for yearling weight in the current study was 0.32±0.19. This estimate is in agreement with previous studies by Neser et al. [66], van Vleck and Diop [28], Haile & Assa-Mersha [68] who reported heritability estimates of 0.23 in Brangus, 0.24 in Gobra, and 0.34 in Boran cattle, respectively. The findings from the current study are not in agreement with results reported by Pico et al. [69] and Eler et al. [70] who reported low heritability estimates of 0.14 in Brahman and 0.16 in Nellore, respectively. The differences observed amongst the above-mentioned studies could be attributed to breed, farm management, sample sizes, and statistical models used. Despite the differences, the findings from the current study indicate that yearling weight could be improved through selection in this population.

3.3.7.5 Skin thickness

The heritability estimate for skin thickness in this study was 0.40. This estimate is higher than reported in a previous study by Maiorano et al. [71], who reported a low heritability of 0.12±0.02 in Nellore cattle from Brazil. The heritability estimate obtained in the current study indicate that there is sufficient additive genetic variation for selection to increase skin thickness The skin acts as the first line of defence against pathogens and it facilitates regulation of temperature and pressure [72]. For example, animals with thick skin are better protected during tick infestation as compared to animals with a thin skin. In animals with thick skin, it is hard for ticks to bite, suck blood and introduce toxins. In cattle, skin thickness plays a role in the expression of other traits such as tick count, heat tolerance, body fat deposition, and milk production [73,74]. Skin thickness is also of interest in sheep. Moderate heritability estimates of 0.26, 0.21 and 0.21 have been reported in lambs by Ghambavani et al. [75], Tait et al.[76], and Soltanighombavani [77], respectively. The paucity of information on skin thickness in cattle elicits the need for more studies.

3.4 Conclusions

Heritability estimates for tick count, birth, weaning and yearling weights and skin thickness obtained in the current study indicate that these traits are under sufficient genetic control to warrant genetic improvement through selection. Year and season of infestation and live weight should be considered when estimating genetic parameters for tick count. Breeding for tick resistance could provide an additional and complementary approach for sustainable livestock production in tick-infested regions. However, there is a need to study the genomic and immunological components influencing tick count, to gain better understanding of the biological mechanisms underlying tick resistance in cattle.

3.5 References

[1] Porto Neto LR, Jonsson NN, D'Occhio MJ, Barendse W. Molecular genetic approaches for identifying the basis of variation in resistance to tick infestation in cattle. Vet Parasitol 2011;180:165–72. https://doi.org/10.1016/j.vetpar.2011.05.048.

[2] Cardoso FF, Matika O, Djikeng A, Mapholi N, Burrow HM, Yokoo MJI, et al. Multiple Country and Breed Genomic Prediction of Tick Resistance in Beef Cattle. Front Immunol 2021;12:620847. https://doi.org/10.3389/fimmu.2021.620847.

[3] Mapholi NO, Marufu MC, Maiwashe A, Banga CB, Muchenje V, MacNeil MD, et al. Towards a genomics approach to tick (Acari: Ixodidae) control in cattle: a review. Ticks Tick-Borne Dis 2014;5:475–83. https://doi.org/10.1016/j.ttbdis.2014.04.006.

[4] Grisi L, Leite R, Martins J, Barros AT, Andreotti R, Cançado P, et al. Reassessment of the potential economic impact of cattle parasites in Brazil Reavaliação do potencial impacto econômico de parasitos de bovinos no Brasil. Rev Bras Parasitol Vet Braz J Vet Parasitol Orgao Of Col Bras Parasitol Vet 2014;23. https://doi.org/10.1590/S1984-29612014042.

[5] Mapholi NO, Maiwashe A, Matika O, Riggio V, Bishop SC, MacNeil MD, et al. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks Tick-Borne Dis 2016;7:487–97. https://doi.org/10.1016/j.ttbdis.2016.02.005.

[6] Jongejan F, Berger L, Busser S, Deetman I, Jochems M, Leenders T, et al. Amblyomma hebraeum is the predominant tick species on goats in the Mnisi Community Area of Mpumalanga Province South Africa and is co-infected with Ehrlichia ruminantium and Rickettsia africae. Parasit Vectors 2020;13:172. https://doi.org/10.1186/s13071-020-04059-5.

[7] Rajput ZI, Hu S, Chen W, Arijo AG, Xiao C. Importance of ticks and their chemical and immunological control in livestock. J Zhejiang Univ Sci B 2006;7:912–21. https://doi.org/10.1631/jzus.2006.B0912.

[8] Olwoch JM, Reyers B, Engelbrecht FA, Erasmus BFN. Climate change and the tick-borne disease, Theileriosis (East Coast fever) in sub-Saharan Africa. J Arid Environ 2008;72:108–20. https://doi.org/10.1016/j.jaridenv.2007.04.003.

[9] Lew-Tabor AE, Rodriguez Valle M. A review of reverse vaccinology approaches for the development of vaccines against ticks and tick borne diseases. Ticks Tick-Borne Dis 2016;7:573–85. https://doi.org/10.1016/j.ttbdis.2015.12.012.

[10] Muyobela J, Nkunika POY, Mwase ET. Resistance status of ticks (Acari; Ixodidae) to amitraz and cypermethrin acaricides in Isoka District, Zambia. Trop Anim Health Prod 2015;47:1599–605. https://doi.org/10.1007/s11250-015-0906-4.

[11] De Meneghi D, Stachurski F, Adakal H. Experiences in Tick Control by Acaricide in the Traditional Cattle Sector in Zambia and Burkina Faso: Possible Environmental and Public Health Implications. Front Public Health 2016;4. https://doi.org/10.3389/fpubh.2016.00239.

[12] Phocas F, Belloc C, Bidanel J, Delaby L, Dourmad JY, Dumont B, et al. Review: Towards the agroecological management of ruminants, pigs and poultry through the development of sustainable breeding programmes: I-selection goals and criteria. Anim Int J Anim Biosci 2016;10:1749–59. https://doi.org/10.1017/S1751731116000926.

[13] Marufu MC, Qokweni L, Chimonyo M, Dzama K. Relationships between tick counts and coat characteristics in Nguni and Bonsmara cattle reared on semiarid rangelands in South Africa. Ticks Tick-Borne Dis 2011;2:172–7. https://doi.org/10.1016/j.ttbdis.2011.07.001.

[14] Nwachukwu, E.N., Ogbu, C.C., Edozie, C., Oke, U.K., Ojewola , G.S., Ekumankama, O.O. Incidence of tick infestation in mixed breeding herd of indigenous cattle in a rainforest agro-ecological zone. Livestock Research for Rural Development. Livest Res Rural Dev 2020;32.

[15] Roberts JA. Acquisition by the Host of Resistance to the Cattle Tick, Boophilus microplus (Canestrini). J Parasitol 1968;54:657–62. https://doi.org/10.2307/3277013.

[16] Tabor AE, Ali A, Rehman G, Rocha Garcia G, Zangirolamo AF, Malardo T, et al. Cattle Tick Rhipicephalus microplus-Host Interface: A Review of Resistant and Susceptible Host Responses. Front Cell Infect Microbiol 2017;7:506. https://doi.org/10.3389/fcimb.2017.00506.

[17] Burrow HM. Variances and covariances between productive and adaptive traits and temperament in a composite breed of tropical beef cattle. Livest Prod Sci 2001;3:213–33.

[18] Fourie PJ, Foster LA, Neser FWC. Score and hide-thickness, together with tick burden and body condition score, in four cattle breeds in the South-eastern Free State province of South Africa. J New Gener Sci 2013;11:66–73. https://doi.org/10.10520/EJC141432.

[19] Marufu MC. Prevalence of ticks and tick-borne diseases in cattle on communal Rangelands in the highland areas of the Eastern Cape province, South Africa. Thesis. University of Fort Hare, 2008.

[20] Nyangiwe N, Yawa M, Muchenje V. Driving forces for changes in geographic range of cattle ticks (Acari: Ixodidae) in Africa: A review. South Afr J Anim Sci 2018;48:829–41. https://doi.org/10.4314/sajas.v48i5.

[21] Mwai O, Hanotte O, Kwon Y-J, Cho S. African Indigenous Cattle: Unique Genetic Resources in a Rapidly Changing World. Asian-Australas J Anim Sci 2015;28:911–21. https://doi.org/10.5713/ajas.15.0002R.

[22] Mapholi NO, Maiwashe A, Matika O, Riggio V, Bishop SC, MacNeil MD, et al. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks Tick-Borne Dis 2016;7:487–97. https://doi.org/10.1016/j.ttbdis.2016.02.005.

[23] Scholtz MM, Theunissen A. The use of indigenous cattle in terminal cross-breeding to improve beef cattle production in Sub-Saharan Africa. Recur Genéticos Anim Una Rev Int 2010.

[24] Theunissen A, Scholtz MM, Neser FWC, MacNeil MD. Crossbreeding to increase beef production: additive and non-additive effects on weight traits. South Afr J Anim Sci 2013;43:143–52.

[25] Moyo S. Evaluation of the productivity of indigenous cattle and some exotic beef breeds and their crosses in Zimbabwe. ILCA; 1990.

[26] Budeli MA, Nephawe KA, Norris D, Selapa NW, Bergh L, Maiwashe A. Genetic parameter estimates for tick resistance in Bonsmara cattle. South Afr J Anim Sci 2009;39:321–7.

[27] Mapholi NO, Maiwashe A, Matika O, Riggio V, Banga C, MacNeil MD, et al. Genetic parameters for tick counts across months for different tick species and anatomical locations in South African Nguni cattle. Trop Anim Health Prod 2017;49:1201–10. https://doi.org/10.1007/s11250-017-1336-2.

[28] Diop M, Vleck LDV. Estimates of genetic parameters for growth traits of Gobra cattle. Anim Sci 1998;66:349–55. https://doi.org/10.1017/S1357729800009474.

[29] Muchenje V, Dzama K, Chimonyo M, Raats JG, Strydom PE. Tick susceptibility and its effects on growth performance and carcass characteristics of Nguni, Bonsmara and Angus steers raised on natural pasture. Anim Int J Anim Biosci 2008;2:298–304. https://doi.org/10.1017/S1751731107001036.

[30] Van Niekerk, M. & Neser, F.W.C., 2006. Genetic parameters for growth traits in South African Limousin cattle. S. Afr. J. Anim. Sci. 36, 5 (Suppl 1), 6-9. - https://www.ajol.info/index.php/sajas/article/view/4075. (accessed August 2, 2022).

[31] Meyer K. Variance components due to direct and maternal effects for growth traits of Australian beef cattle. Livest Prod Sci 1992;31:179–204. https://doi.org/10.1016/0301-6226(92)90017-X.

[32] Wesonga FD, Orinda GO, Ngae GN, Grootenhuis J. Comparative tick counts on game, cattle and sheep on a working game ranch in Kenya. Trop Anim Health Prod 2006;38:35–42. https://doi.org/10.1007/s11250-006-4318-3.

[33] Clark V, SAS Institute, editors. SAS/STAT 9.1: user's guide. Cary, N.C: SAS Pub; 2004.

[34] ASReml-3-User-Guide.pdf n.d.

[35] Ibelli AMG, Ribeiro ARB, Giglioti R, Regitano LCA, Alencar MM, Chagas ACS, et al. Resistance of cattle of various genetic groups to the tick Rhipicephalus microplus and the relationship with coat traits. Vet Parasitol 2012;186:425–30. https://doi.org/10.1016/j.vetpar.2011.11.019.

[36] Shyma KP, Gupta JP, Singh V. Breeding strategies for tick resistance in tropical cattle: a sustainable approach for tick control. J Parasit Dis 2013;39. https://doi.org/10.1007/s12639-013-0294-5.

[37] O'kelly JC, Seifert GW. The effects of tick (Boophilus microplus) infestations on the blood composition of Shorthorn X Hereford cattle on high and low planes of nutrition. Aust J Biol Sci 1970;23:681–90.

[38] Katiyatiya CLF, Muchenje V, Mushunje A. Seasonal variation in coat characteristics, tick loads, cortisol levels, some physiological parameters and temperature humidity index on Nguni cows raised in low- and high-input farms. Int J Biometeorol 2015;59:733–43. https://doi.org/10.1007/s00484-014-0887-4.

[39] Rocha JF, Martínez R, López-Villalobos N, Morris ST. Tick burden in Bos taurus cattle and its relationship with heat stress in three agroecological zones in the tropics of Colombia. Parasit Vectors 2019;12:73. https://doi.org/10.1186/s13071-019-3319-9.

[40] Kemal J, Tamerat N, Tuluka T. Infestation and Identification of Ixodid Tick in Cattle: The Case of Arbegona District, Southern Ethiopia. J Vet Med 2016;2016:1–8. https://doi.org/10.1155/2016/9618291.

[41] Yessinou RE, Adoligbe C, Akpo Y, Adinci J, Youssao Abdou Karim I, Farougou S. Sensitivity of Different Cattle Breeds to the Infestation of Cattle Ticks Amblyomma variegatum, Rhipicephalus microplus, and Hyalomma spp. on the Natural Pastures of Opkara Farm, Benin. J Parasitol Res 2018;2018:2570940. https://doi.org/10.1155/2018/2570940.

[42] Chartier C., Itard J., Morel P. C., Troncy P. M. Précis de parasitologie vétérinaire tropicale. Collect Univ Francoph 2000;67.

[43] Gharbi M, Darghouth MA. A review of Hyalomma scupense (Acari, Ixodidae) in the Maghreb region: from biology to control. Parasite 2014;21:2. https://doi.org/10.1051/parasite/2014002.

[44] Hughes VL, Randolph SE. Testosterone depresses innate and acquired resistance to ticks in natural rodent hosts: A force for aggregated distributions of parasites. J Parasitol 2001;87:49–54. https://doi.org/10.1645/0022-3395(2001)087[0049:TDIAAR]2.0.CO;2.

[45] Monyela BM. A two-year long drought in summer 2014/2015 and 2015/2016 over South Africa. Master Thesis. University of Cape Town, 2017.

[46] Mekonnen S, Pegram RG, Gebre S, Mekonnen A, Jobre Y, Zewde S. A synthesis review of ixodid (*Acari: Ixodidae*) and argasid (*Acari: Argasidae*) ticks in Ethiopia and their possible roles in disease transmission. Ethiop Vet J 2007;11:1–24.

[47] Mohamed B, Belay A, Hailu D. Species composition, prevalence and seasonal variations of ixodid cattle ticks in and around Haramaya town, Ethiopia. J Vet Med Anim Health 2014;6:131–7. https://doi.org/10.5897/JVMAH2014.0275.

[48] Machado MA, Azevedo ALS, Teodoro RL, Pires MA, Peixoto MGCD, de Freitas C, et al. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (Bos taurus x Bos indicus). BMC Genomics 2010;11:280. https://doi.org/10.1186/1471-2164-11-280.

[49] Martinez ML, Machado MA, Nascimento CS, Teodoro RL, Furlong J, Prata MCA, et al. Association of BoLA-DRB3.2 alleles with tick (Boophilus microplus) resistance in cattle. Genet Mol Res 2006:12.

[50] Regitano LCA, Ibelli AMG, Gasparin G, Miyata M, Azevedo ALS, Coutinho LL, et al. On the search for markers of tick resistance in bovines. Dev Biol 2008;132:225–30. https://doi.org/10.1159/000317164.

[51] Gasparin G, Miyata M, Coutinho LL, Martinez ML, Teodoro RL, Furlong J, et al. Mapping of quantitative trait loci controlling tick [Riphicephalus (Boophilus) microplus] resistance on bovine chromosomes 5, 7 and 14. Anim Genet 2007;38:453–9. https://doi.org/10.1111/j.1365-2052.2007.01634.x.

[52] Nyangiwe N, Goni S, Hervé-Claude LP, Ruddat I, Horak IG. Ticks on pastures and on two breeds of cattle in the Eastern Cape province, South Africa. Onderstepoort J Vet Res 2011;78:320. https://doi.org/10.4102/ojvr.v78i1.320.

[53] Pillay A, Nyangiwe N, Mukaratirwa S. Low genetic diversity and population structuring of Amblyomma hebraeum and Rickettsia africae from coastal and inland regions in the Eastern Cape Province of South Africa. Med Vet Entomol n.d.;n/a. https://doi.org/10.1111/mve.12629.

[54] Mapholi NO, Maiwashe A, Matika O, Riggio V, Banga C, MacNeil MD, et al. Genetic parameters for tick counts across months for different tick species and anatomical locations in South African Nguni cattle. Trop Anim Health Prod 2017;49:1201–10. https://doi.org/10.1007/s11250-017-1336-2.

[55] Burrow HM. Genetic Aspects of Cattle Adaptation in the Tropics, CAB International; 2014.

[56] Prayaga KC, Henshall JM, Prayaga KC, Henshall JM. Adaptability in tropical beef cattle: genetic parameters of growth, adaptive and temperament traits in a crossbred population. Aust J Exp Agric 2005;45:971–83. https://doi.org/10.1071/EA05045.

[57] Prayaga KC, Corbet NJ, Johnston DJ, Wolcott ML, Fordyce G, Burrow HM. Genetics of adaptive traits in heifers and their relationship to growth, pubertal and carcass traits in two tropical beef cattle genotypes. Anim Prod Sci 2009;49:413. https://doi.org/10.1071/EA08247.

[58] Porto-Neto LR, Reverter A, Prayaga KC, Chan EKF, Johnston DJ, Hawken RJ, et al. The Genetic Architecture of Climatic Adaptation of Tropical Cattle. PLoS ONE 2014;9:e113284. https://doi.org/10.1371/journal.pone.0113284.

[59] Johanson JM, Berger PJ. Birth Weight as a Predictor of Calving Ease and Perinatal Mortality in Holstein Cattle1. J Dairy Sci 2003;86:3745–55. https://doi.org/10.3168/jds.S0022-0302(03)73981-2.

[60] Nesengani LT, Nephawe KA, Sebei J, Norris D, Maiwashe A. Genetic relationship between weaning weight and milk yield in Nguni cattle. Animal 2018;12:199–204. https://doi.org/10.1017/S1751731117001483.

[61] Yin T, König S. Genetic parameters for body weight from birth to calving and associations between weights with test-day, health, and female fertility traits. J Dairy Sci 2018;101:2158–70. https://doi.org/10.3168/jds.2017-13835.

[62] Bedoya O C, Alzate R J, Ángel C J, Escobar R C, Calvo C S, Bedoya O C, et al. Genetic evaluation for weight traits in commercial Brahman cattle. Rev MVZ Córdoba 2019;24:7225–30. https://doi.org/10.21897/rmvz.1662.

[63] Ossa GA, Suárez MA, Pérez JE. Efectos del medio y la herencia sobre los pesos al nacimiento, al destete y a los 16 meses de edad en terneros de la raza criolla Romosinuano. Cienc Tecnol Agropecu 2007;8:81–92. https://doi.org/10.21930/rcta.vol8_num2_art:98.

[64] Boligon AA, Albuquerque LG de, Mercadante MEZ, Lôbo RB. Herdabilidades e correlações entre pesos do nascimento à idade adulta em rebanhos da raça Nelore. Rev Bras Zootec 2009;38:2320–6. https://doi.org/10.1590/S1516-35982009001200005.

[65] Abera H, Abegaz S, Mekasha Y. Genetic parameter estimates of pre-weaning weight of Horro (Zebu) and their crosses with Holstein Friesian and Jersey cattle breeds in Ethiopia. Int J Livest Prod 2011;2.

[66] Neser FWC, van Wyk JB, Fair MD, Lubout P, Crook BJ. Estimation of genetic parameters for growth traits in Brangus cattle. South Afr J Anim Sci 2012;42:469–73.

[67] Otto PI, Santos AL dos, Perotto D, Oliveira SN de, Granzotto F, Gobo DOR, et al. Estimation of genetic parameters for weaning and yearling weights in a composite population used to form the Purunã breed. Rev Bras Zootec 2021;50. https://doi.org/10.37496/rbz5020180224.

[68] Haile-Mariam M, Kassa-Mersha H. Estimates of direct and maternal (co)variance components of growth traits in Boran cattle. J Anim Breed Genet 1995;112:43–52. https://doi.org/10.1111/j.1439-0388.1995.tb00539.x.

[69] Pico BA. ESTIMATION OF GENETIC PARAMETERS FOR GROWTH TRAITS IN SOUTH AFRICAN BRAHMAN CATTLE n.d.:79.

[70] Eler J, Van Vleck L, Ferraz J, Lôbo R. Estimation of Variances Due to Direct and Maternal Effects for Growth Traits of Nelore Cattle. J Anim Sci 1995;73:3253–8. https://doi.org/10.2527/1995.73113253x.

[71] Maiorano AM [UNESP, Oliveira MCS, Cyrillo JNSG, Albuquerque LG [UNESP, Curi RA [UNESP, Silva JA [UNESP. Genetic study of skin thickness and its association with postweaning growth in Nellore cattle: estimation of the genetic parameters. Genet Mol Res 2016:10. https://doi.org/10.4238/gmr.15017124.

[72] Lai-Cheong JE, McGrath JA. Structure and function of skin, hair and nails. Medicine (Baltimore) 2009;37:223–6. https://doi.org/10.1016/j.mpmed.2009.03.002.

[73] Zhang H, Liu A, Li X, Xu W, Shi R, Luo H, et al. Genetic analysis of skinfold thickness and its association with body condition score and milk production traits in Chinese Holstein population. J Dairy Sci 2019;102. https://doi.org/10.3168/jds.2018-15180.

[74] Petrofsky J, Goraksh N, Alshammari F, Mohanan M, Soni J, Trivedi M, et al. The ability of the skin to absorb heat; The effect of repeated exposure and age. Med Sci Monit Int Med J Exp Clin Res 2011;17:CR1–8. https://doi.org/10.12659/MSM.881315.

[75] Soltani-Ghombavani M, Dukkipati VSR, Blair HT. Genetic association of skin thickness with lamb survival from birth to weaning, and growth and wool traits in New Zealand Romney sheep. Proc 22nd Conf Assoc Adv Anim Breed Genet AAABG Townsv Qld Aust 2-5 July 2017 2017:589–92.

[76] Tait I, Hickson R, Kenyon P, Blair H. Relationships among skin thickness, fat depth and eye muscle depth in sires and their progeny, and the breeding value for survival of their progeny. Proc N Z Soc Anim Prod n.d.;75.

[77] Soltanighombavani M. Skin thickness as a potential indirect trait for genetic improvement of lamb survival n.d.

4 Chapter 4:

Genome-wide association study for tick count, growth traits, and skin thickness in F_2 Angus x Nguni cattle

Abstract: *Amblyomma hebraeum* ticks pose substantial threat to cattle production, health, and welfare. The current tick-control methods are ineffective which underscores the need for alternative control measures. Genetic variation was explored by performing a genome-wide association study (GWAS) to identify single nucleotide polymorphism (SNP) markers associated with tick count, live body weights, and skin thickness in the F_2 Angus x Nguni population. A total of 266 F_2 Angus x Nguni animals were genotyped using Illumina BovineSNP150 assay. The genotype data was subjected to quality control, after which 115 143 SNPs were retained for downstream analysis. Genome-wide association analysis was undertaken in the R environment using the GenABEL version 1.8 package. The estimated genomic heritability was 0.12±0.15, 0.26±0.16, 0.70±0.18, 0.38±0.18, and 0.58±0.21 for tick count, birth weight, weaning weight, yearling weight, and skin thickness, respectively. SNPs associated with birth weight were observed on BTA3 and BTA12. Two SNPs were suggestively associated with tick count on BTA4 and BTA22. Furthermore, suggestive SNPs associated with weaning weight were observed on BTA2, and BTA18. Another suggestive SNP on BTA8 was associated with yearling weight. Suggestive SNPs associated with skin thickness were observed on BTA2, BTA7, BTA8, BTA9, and BTA12. Despite the modest sample size in the current study, the findings indicate that it is feasible to control tick resistance using tick count as a proxy through genomic selection**.**

Keywords: Crossbred Cattle, Genomic Heritability, SNPs, GWAS

4.1 Introduction

High levels of cattle tick infestation in beef production is a global threat to food security and animal welfare. This also has adverse economic implications. Ticks affect cattle production by sucking blood, damaging hides, and transmitting tick-borne diseases, which prominently influence increases in rates of mortalities. Reduced production translate to less availability of meat products which impacts food security for consumers. The use of acaricides and vaccines are the prevalent methods for controlling tick infestations in cattle production. However, these methods pose some challenges, which include inflated costs associated with the development of drugs, the emergence of acaricide-resistant tick strains, and environmental contamination. The emergence of acaricide-resistant strains of ticks raises animal welfare and human health

concerns. Ticks develop resistance toward acaricides through prolonged and inappropriate utilization [1–3]. This results in delayed therapy when there is an outbreak, translating to production loss and increased treatment costs. The alternative solution may include the development of new drugs; however, this may not be economically sustainable because the development of new drugs is costly, and a long period is needed for drug testing and approval. Although the drug could be approved, the sustainability of its efficacy is not guaranteed since there are possibilities that ticks will soon develop resistance toward the new drugs too. Additionally, acaricides are not environmentally friendly; previous studies have indicated that chemical residues from the acaricides remain in food products such as meat, milk, and their by-products [1,4]. This is a threat to consumers. Therefore, there is a serious need for alternative strategies to control ticks to ensure animal welfare, and sustainable production, as well as to protect consumers from harmful chemical residues.

Selection for naturally resistant animals is a promising alternative measure that can be used to control cattle ticks [5,6]. Cattle's natural resistance to tick infestations has been reported by numerous studies in different breeds [4,7,8]. The use of genetic selection to control tick infestations on cattle will reduce the prevalence of tick-borne diseases. The main challenge is that selecting naturally resistant animals requires an understanding of genetic determinants influencing the host response to tick infestation. It would be advantageous to understand phenotypic and genetic correlations between growth and reproductive traits, as this will facilitate the improvement of productivity. The role of the skin as a barrier to tick attachment may be affected by skin colour and thickness. It has been noted by Burrow et al. [9], that the magnitude of heritability estimates for tick resistance depends on the animal's growth; therefore, there is a need to study the relationship between growth traits, skin thickness, and tick count. Heavily tick-infested animals are expected to gain less weight. Skin thickness serves as a protection against tick infestation; a thick skin makes it difficult for ticks to bite [10].

Efforts have been made globally to uncover and understand genetic determinants influencing natural tick resistance in cattle. Underlying single nucleotide polymorphisms (SNPs) and genes have been discovered using molecular genetics approaches such as genome-wide association studies (GWAS). The GWAS approach utilizes phenotypic information to search for DNA markers (SNPs) associated with traits of economic interest. GWAS is regarded as an efficient method to uncover SNPs associated with complex economic traits [11]. According to Welderufael et al. [12], GWAS has the potential to identify genetic markers for use in marker-assisted selection (MAS). The information on genomic selection can potentially be used to develop breeding schemes that will facilitate breeding of cattle with increased favourable allele

frequencies for a trait of interest [13]. Using GWAS, genetic determinants associated with tick resistance have been discovered across the 29 bovine autosomal chromosomes from different cattle breeds [7,8,14]. The successful discoveries from previous tick resistance studies have generated knowledge and paved a way for more studies that need to be conducted since the information is still limited.

In Africa, commercial beef production is mostly practiced using exotic breeds, which are known to be susceptible to tick infestations, for example, the Angus breed. The lack of studies focused on uncovering genetic determinants associated with tick resistance in African indigenous breeds and their crossbreds remains a major challenge. Africa is still lagging behind in controlling tick resistance through marker-assisted selection because there is limited information to understand the genetic architecture for this trait. Although there is information from studies conducted in other regions, it is especially important to generate information on the African environment, since the environment plays a key role in influencing tick resistance.

The aim of the current study was to identify genomic regions associated with tick count in F_2 Angus x Nguni crossbreed cattle. This is the first study that has focused on unravelling genetic determinates associated with ticks in F_2 Angus x Nguni crossbreed cattle. It has been reported that the use of an F_2 crossbred population is advantageous for GWAS of polygenic traits because it improves the QTL mapping resolution, and accuracy while reducing the false discovery rate (FDR) [15]. The findings from this study will provide a better understanding of genetic determinants influencing tick count in F_2 Angus x Nguni crossbred cattle. This information could then be used to improve the trait through marker-assisted selection.

4.2 Materials and Methods

4.2.1 Phenotypic data and DNA source

Phenotypic data were collected from 266 F_2 Angus x Nguni crossbred cattle as described in Chapter 3. This includes data on tick count, skin thickness, live weights, and coat colour patterns. These traits were chosen because they play a role on the level of tick infestation count in cattle. Hair samples from the tail were collected for DNA extraction assuring enough hair roots. The collected hair samples were stored in labelled LidCat bags to avoid contamination until the day of DNA extraction. Figure 4.1 below shows a sample of the LidCat bag with hair samples.

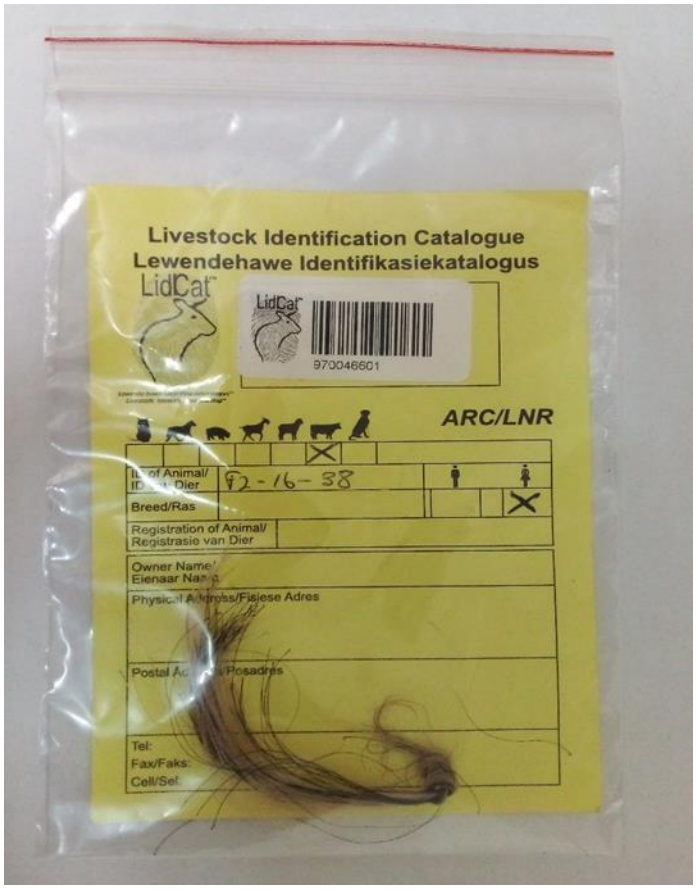

Figure 4.1: Hair samples in a LidCat bag

4.2.2 DNA extraction

Hair samples with visible roots were used to extract genomic DNA. The extraction of DNA was carried out using phenol:chloroform:isoamyl alcohol extraction and ethanol precipitation method described by Sambrook et al. [16]. The ratio of absorbance at 260 and 280 nm was used to assess DNA purity, and a ratio of between 1.8 to 1.9 was accepted as pure for the isolated DNA.

4.2.3 SNP genotyping and data management

Two hundred and sixty-six samples of F_2 Angus x Nguni cattle were genotyped using Bovine 150K SNP chip assay from Illumina at the Agricultural Research Council - Biotechnology Platform and Neogen Company. Data were received in a form of final reports, which are in text format. To create MAP and PED files, SNPConvert (https://github.com/nicolazzie/SNPConvert) was used as described by Nicolazzi et al. [17]. The SNPConvert program has different options to convert the files. For the current study, pedda_row.py program was used. This program converts Illumina ROW files into PLINK format. Two input files are required, an Illumina FinalReport file (in ROW format) and one

SNP_Map (original from Illumina). The program offers a parameter file for the users to modify accordingly before running the program. The created SNP position file (MAP) and associated genotype (PED) files allowed the resolution of ambiguous SNP before merging. Both datasets had problematic SNPs that shared the same positions, however, variants of three or more alleles were detected using the Plinkv1.9 software. These problems were solved by following the instructions depicted in the Plink1.9 software manual [18]. A brief illustration of data management is presented in Figure 4.2.

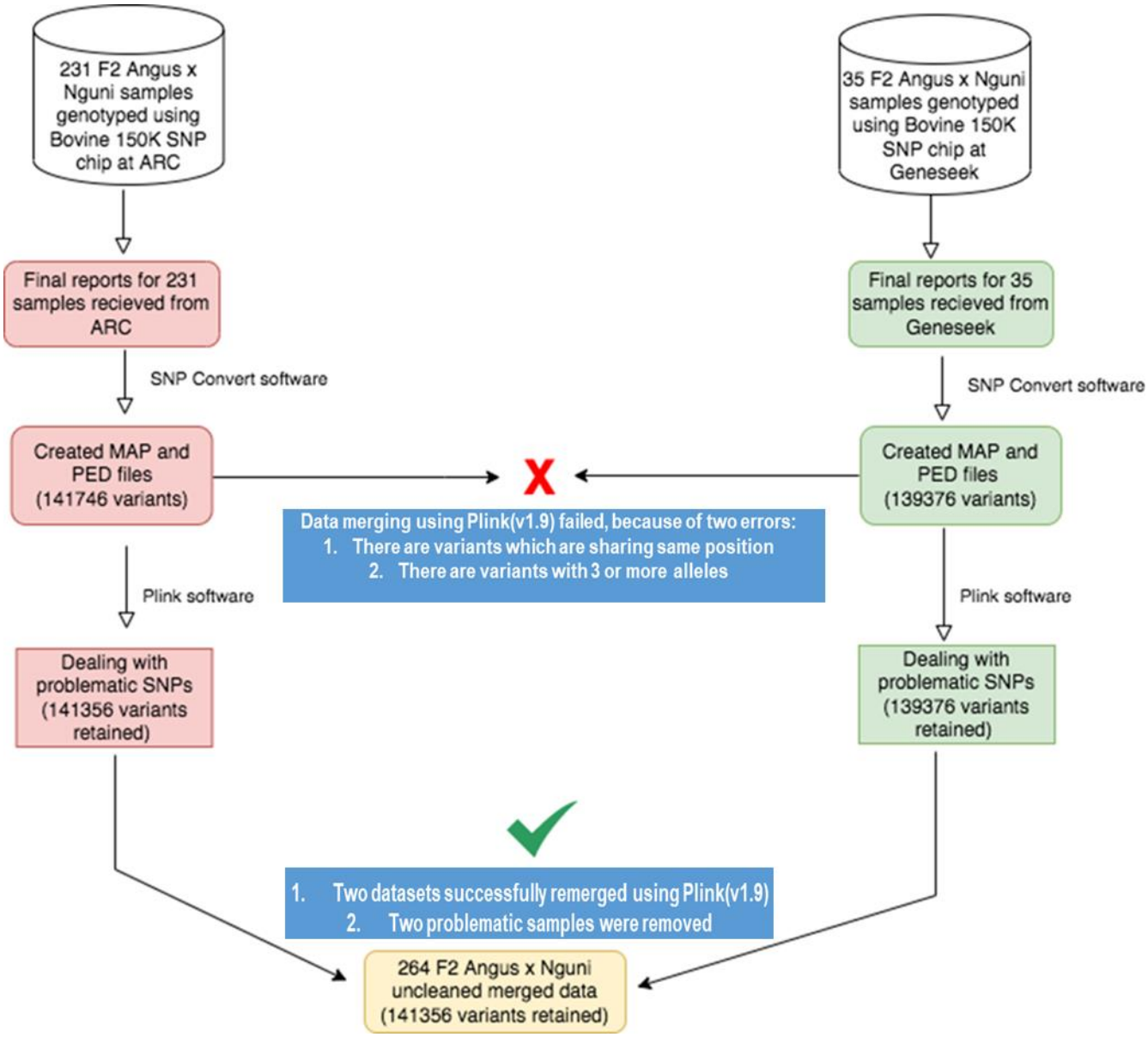

Figure 4.2: F_2 Angus x Nguni genotyped data management flowchart.

After the datasets were merged, the distribution of SNPs throughout the chromosomes were presented graphically in Figure 4.3.

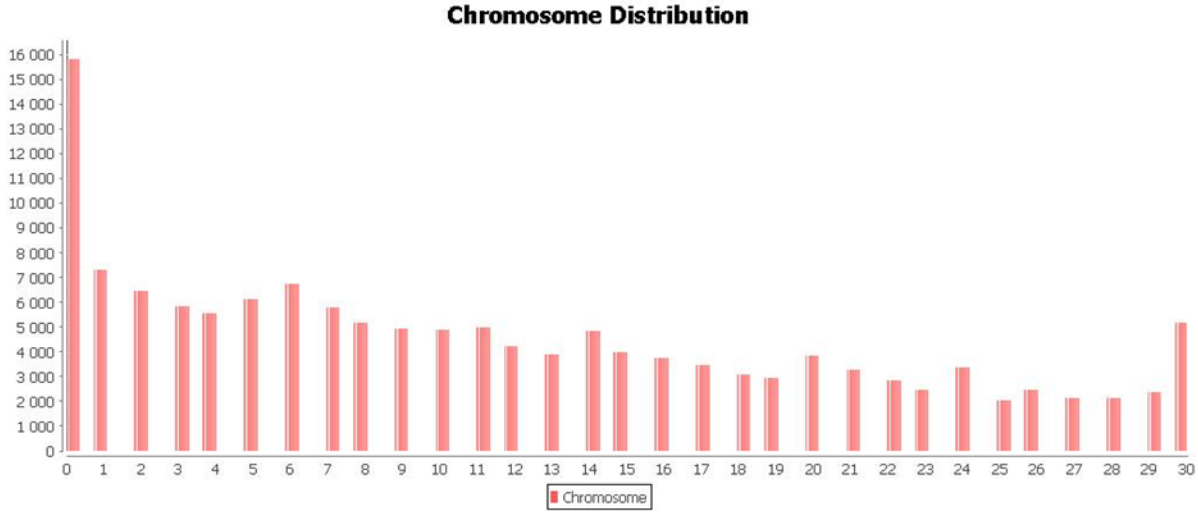

Figure 4.3: Variants distribution according to chromosomes. The y-axis represents the frequency of the SNPs, and the x-axis represents chromosomes, 0 denotes SNPs that are not assigned to any chromosomal location. Chromosomes 1 to 29 represents autosomal chromosome pairs and 30 is sex chromosome pair.

Merged data were subjected to quality control (QC) using GenABEL software in the R environment [19] using QC parameters which excluded SNPs with minor allele frequency of less than 1%, more than 10 % missing genotypes, and deviated from Hardy-Weinberg equilibrium ($p < 1.0 \times 10^{-6}$). Additionally, SNPs that were on the sex chromosome and the ones that had unmapped locations were removed. Initially, there were 141 356 variants; however, after QC only 115 143 variants were retained. The cleaned data were then used for downstream analyses, which included the estimation of linkage disequilibrium, and genome-wide association analysis.

4.2.4 Genome-wide linkage disequilibrium

Genome-wide pairwise linkage disequilibrium (r^2) over a 1000kb window was estimated using Plink (v1.9) software [18]. Average r^2 values corresponding to inter-marker distance were calculated in the R environment. The resulting r^2 values were plotted against inter-marker distance to graphically assess the linkage disequilibrium pattern for the current study.

4.2.5 Estimation of heritability using a kinship matrix

For all the traits studied, the phenotypic variance and residual variance were estimated by fitting the animal model in ASREML software [20] as described by VanRaden [21], using the following model:

$$y = x\beta + Zu + e$$

Where y is a vector of phenotypic values, $x\beta$ is the mean and e is a random error vector with variance of $\mathbf{R}\sigma_e^2$. Matrix $\mathbf{R}$ is diagonal with elements $\mathbf{R_{ii}} = (\mathbf{1}\backslash\mathbf{Rdau}) - \mathbf{1}$, where $\mathbf{Rdau}$ is the bull's reliability from daughters with parent information excluded. Moreover, vector u contains the additive genetic effects that correspond to allele substitution effects for each marker. Lastly, vector Zu represent the sum of all marker loci assumed to equal the vector of breeding values calculated.

4.2.6 Association analysis

Association analysis was performed under a null hypothesis, which assumed that there was no association between SNPs and the level of tick count. To prepare for GWAS analysis, an ANOVA was performed to identify effects associated with tick count. This analysis was broadly explained in chapter 3. The fixed effects that were significant ($p < 0.05$) in the model for tick count were used for the association analysis. Tick count data were log-transformed to approximate normality with a constant of ?? added to avoid zeros.

All association analyses were performed using a single locus model in R using the GenABEL package [19]. For all the traits, the association analysis was initially conducted using the score test model. The findings showed the presence of genetic structure. The actual presence of population stratification was investigated by calculating the genomic kinship matrix using the "ibs'' function in GenABEL. The designed kinship matrix was then used to present the genetic structure of the population in the form of a multidimensional scaling (MDS) plot showing the first two principal components. Moreover, relatedness amongst the cattle was assessed by estimating pair-wise identity by state genomic kinship matrix [7]. This was done using autosomal SNPs and a heatmap was plotted to present the relatedness. Furthermore, the single SNP association was tested using the mixed model function of the score [22], while adjusting for population stratification and relatedness [23]. The population size used for this study was too small, making it difficult to use Bonferroni corrected genome-wide significant level because it was too stringent, and is known for its potential to inflate findings in small population datasets. Therefore, a nominal threshold of $P_{nominal} = 5 \times 10^{-5}$ was applied [24]. This

meant that SNPs with a p-value smaller than 5 x 10^{-5} were considered to have nominal significance with the trait of interest. The genome-wide significance level was described as 0.05/ N, whereby N was the number of variants after QC.

The following model was used for the association analysis of all the traits of interest:

$$y = W\alpha + x\beta + Zu + e$$

where y is a vector of phenotypic values; α is a vector of fixed effects including the population mean and the first MDS component; W is the designed matrix for fixed effects; β is the marker effect; x is a vector of marker genotypes; Zu is a designed matrix for relatedness, and e is the random errors.

4.3 Results

4.3.1 Marker information

Downstream genome-wide analysis was conducted on F_2 Angus x Nguni crossbred cattle with 115 143 variants after quality control. The data consisted of variants distributed throughout the 29 bovine autosomes (Figure 4.3). As shown in Figure 4.4, LD decays as distance increases. The average LD estimate r^2 = 0.1, extended to about 400kbp in the current study.

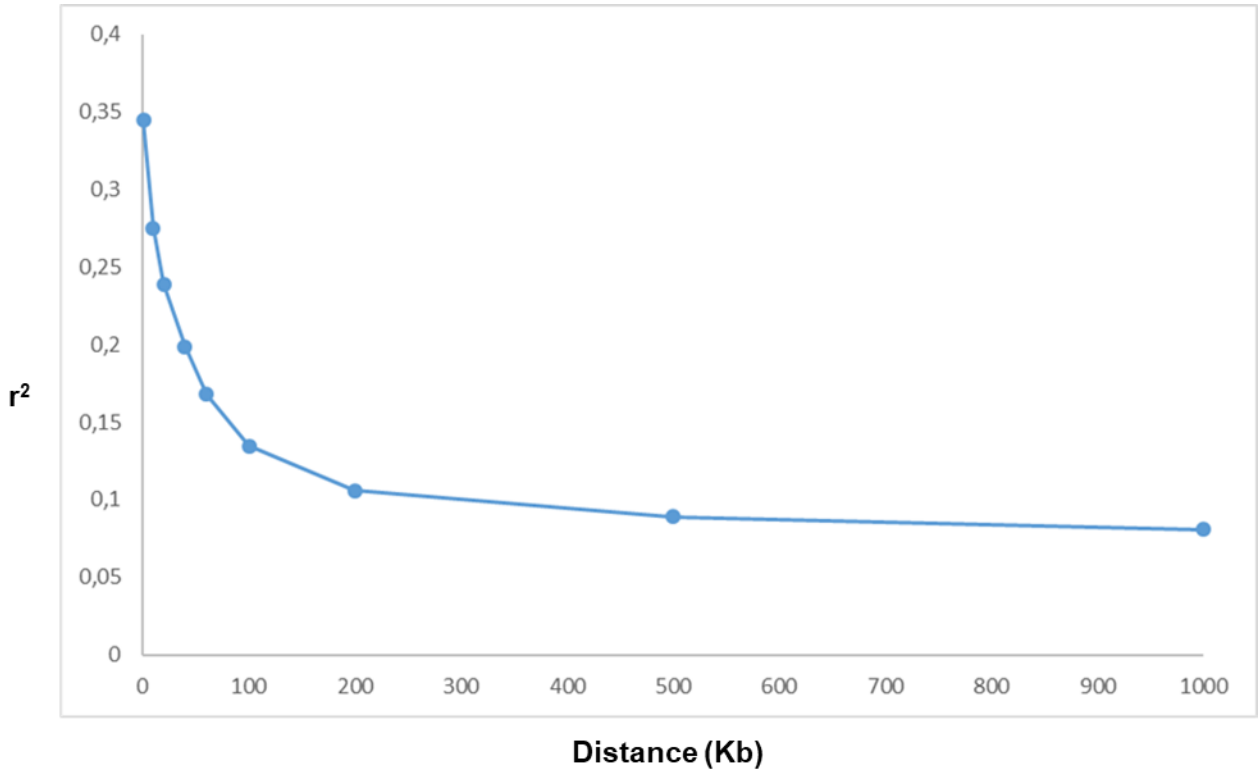

Figure 4.4: Genome-wide pairwise linkage disequilibrium in F_2 Angus x Nguni crossbred cattle.

4.3.2 Heritability estimates

Table 4.1 shows the heritability estimates obtained when the kinship matrix (computed using genotyped data) was fitted instead of a pedigree. From Table 4.1, it can be noted that the heritability estimates ranged from low to high.

Table 4.4.1: Heritability estimates obtained by fitting the kinship matrix using the animal model

Trait	Genetic variance	Phenotypic variance	Heritability (se)
BWT	4.56	17.78	0.26 (0.16)
YWT	478.70	1274.20	0.38 (0.18)
WWT	616.81	883.90	0.70 (0.18)
STHICK	3.51	6.06	0.58 (0.21)
TCOUNT	84.28	685.98	0.12 (0.15)

TCOUNT- Tick Count; BWT-Birth Weight; WWT-Weaning Weight; YWT-Yearling Weight; STHICK-Skin Thickness

4.3.3 Genome-Wide association findings

The multidimensional scaling (MSD) plot showed that the F_2 Angus x Nguni population was genetically clustering into three groups (Figure 4.5).

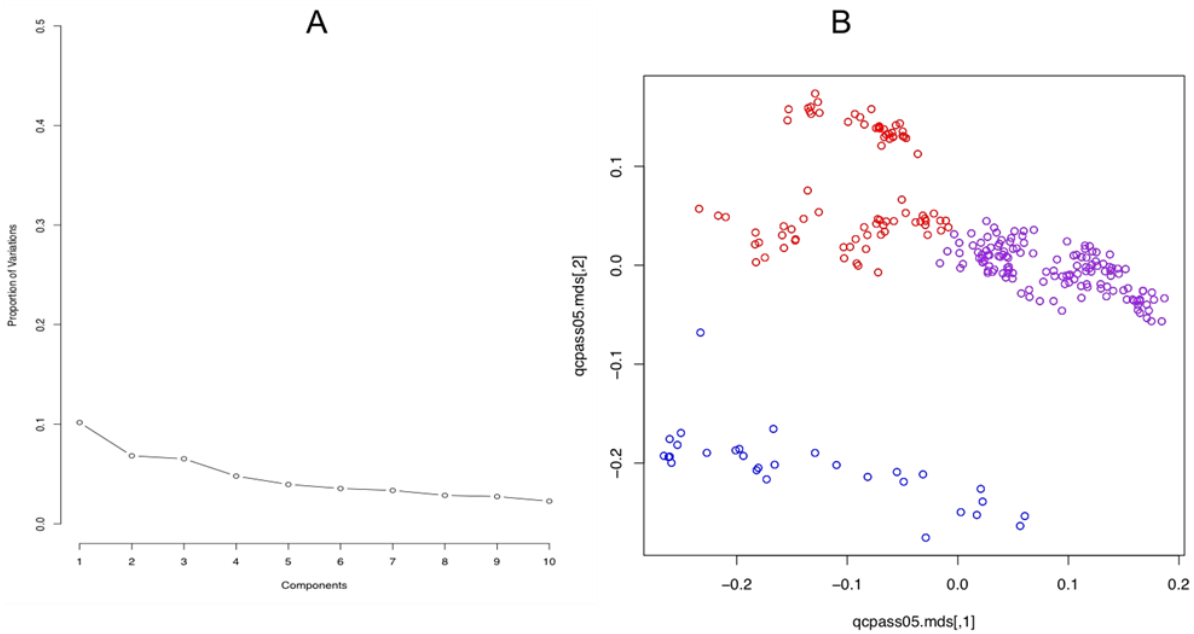

Figure 4.5: Population structure analysis. **A** is a scree plot and **B** is a multidimensional scaling plot showing the distribution of animals using the first two principal components (PCs) of the genomic relationship matrix.

Figure 4.5A is a scree plot showing the contribution to the variance of the first ten principal components to the population structure. There was not much variation attributed to the first two PCs. The most important PCs are indicated by the amount of variance they explain. Figure 4.5B, shows the clustering of animals from plots of the first two principal components. Although, the study population is genetically clustered into three groups, the red cluster does not seem to be clearly defined. The clustering indicates different proportions of Angus and Nguni crossbred animals. These observations are supported by the kinship matrix presented in the form of a heatmap in Figure 6. The kinship matrix heatmap shows the relatedness amongst the animals, using the colour code and kinship coefficient. A coefficient closer to 0 indicates less relatedness and closer to 1 indicates more relatedness. Moreover, on the left and the top of the heatmap diagram, dendrogram trees show that the study population was genetically clustered into three groups as observed on the MDS plot (Figure 4.5B).

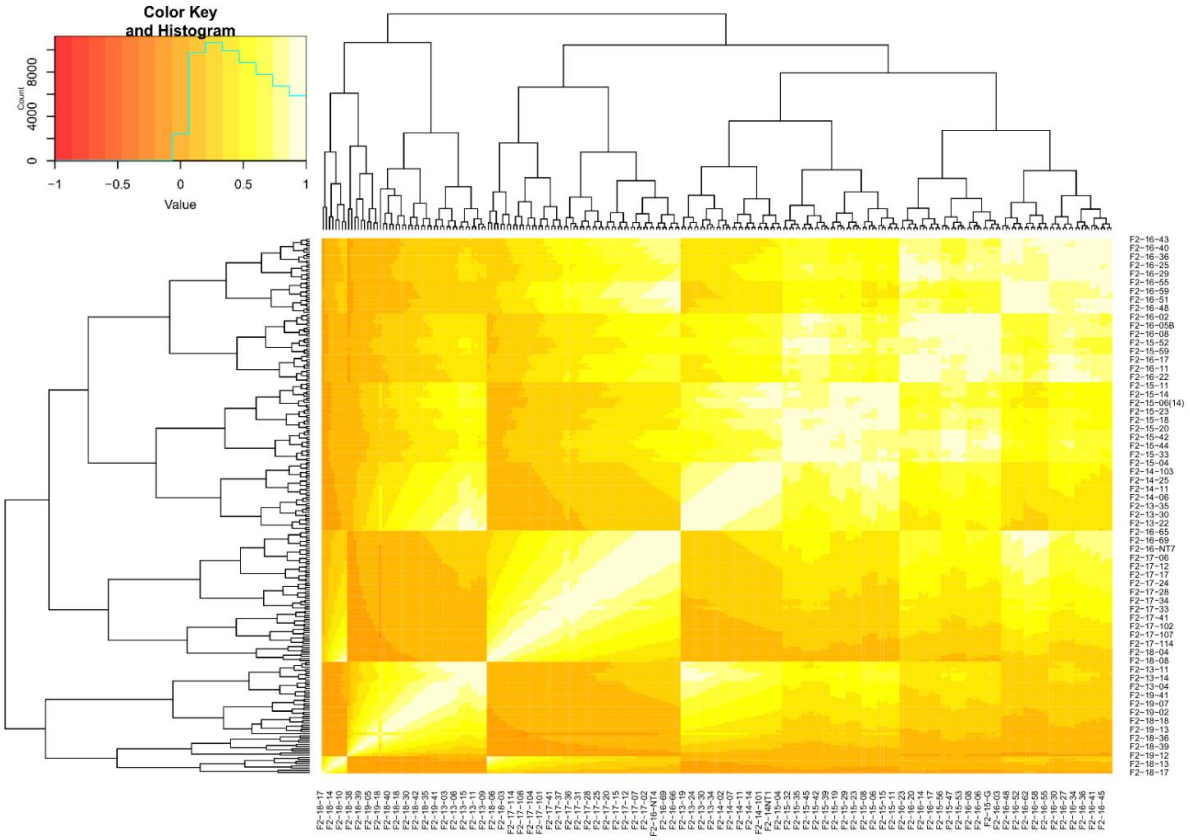

Figure 4.6: Heatmap plot and dendrogram tree of kinship matrix for F_2 Angus x Nguni cattle.

Figure 4.7A shows a Manhattan plot indicating the association results for the tick count. The genome-wide cut-off (0.05/Number of variants after quality control) and suggestive threshold ($P_{nominal}$ = 5 x 10^{-5}) are indicated by black and red dotted lines, respectively. The findings depicted in Figure 4.7A show that no SNP passed the genome-wide cut-offline, however, there were two SNPs that passed the suggestive line (nominal threshold). The two SNPs were observed on chromosome 4 (BovineHD0400032426) and 22 (BovineHD2200005710). The Manhattan plot indicates that no SNP passed the required significance thresholds. Figure 4.7B shows a quantile-quantile (q-q) plot that quantifies the extent of genomic inflation. The observed results could be a product of the Bulmer effect, which is the effect that has the potential to reduce the proportion of genetic variance explained by the marker [25].

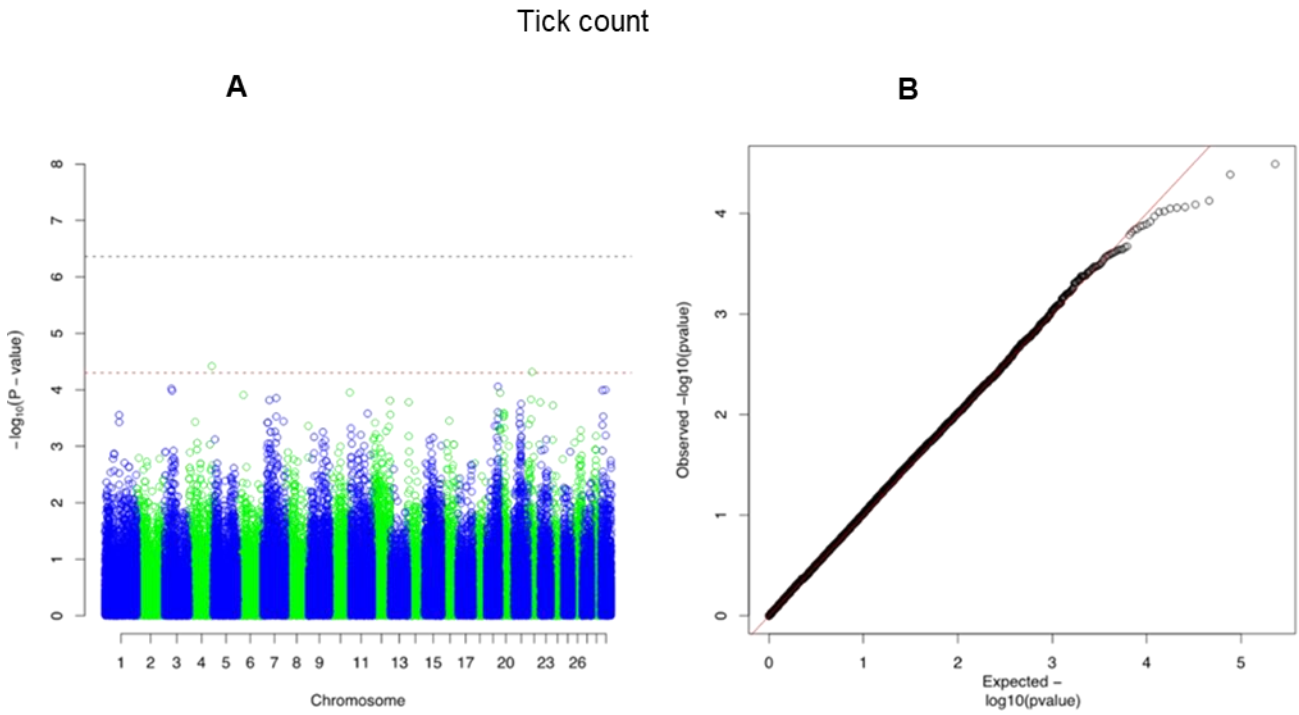

Figure 4.7: Manhattan and q-q plots showing findings for tick count trait.

Figure 4.8 shows results for growth traits and skin thickness. For birth weight, 13 SNPs were observed, two passed the genome-wide significant line (Hapmap38587-BTA-23830 and BovineHD1200022868) and 11 passed the suggestive line. For weaning and yearling weights, 6 SNPs and 1 SNP passed the suggestive threshold, respectively.

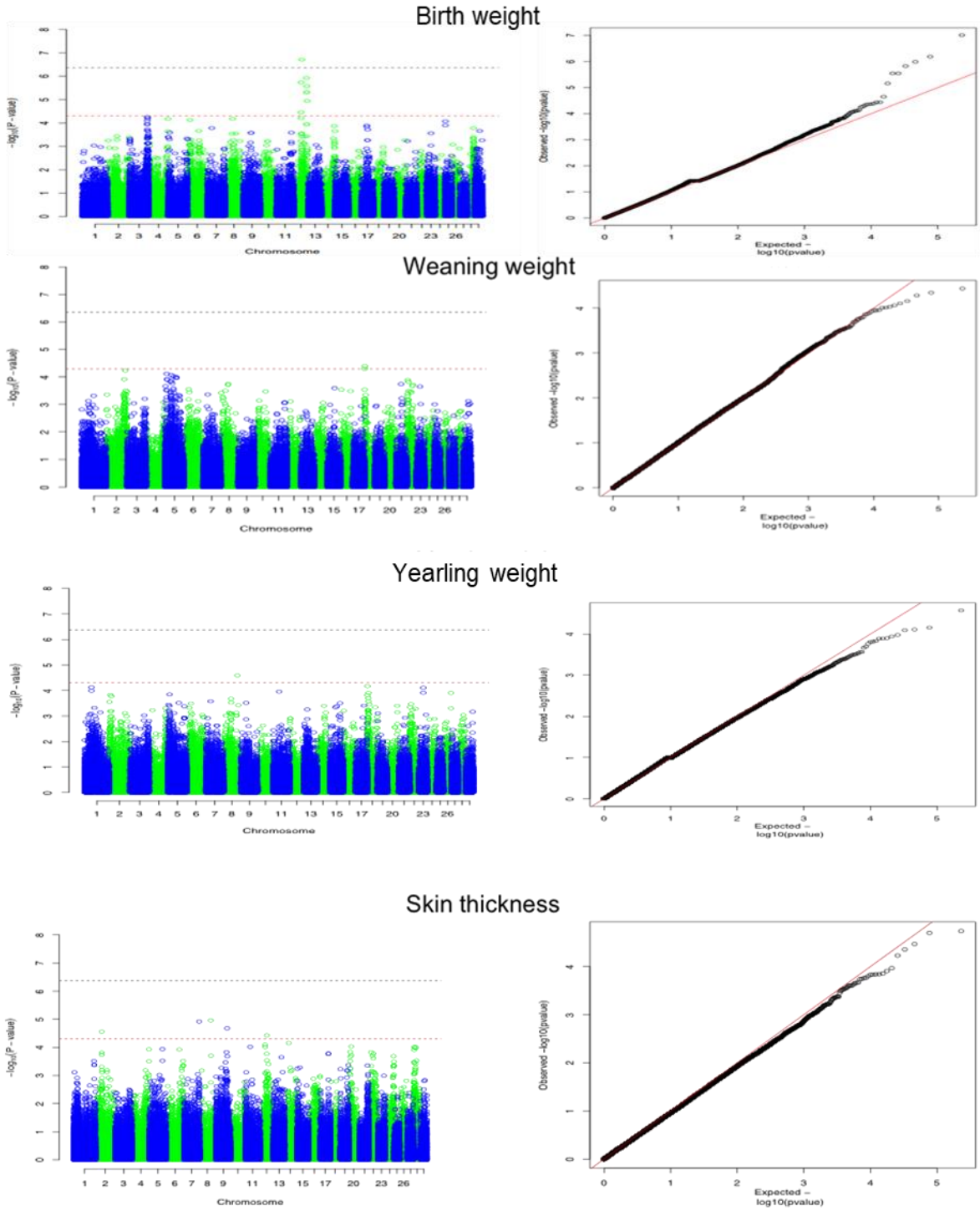

Figure 4.8: Manhattan and Quantile-Quantile plots showing findings for live weights traits and skin thickness.

More details on the SNPs of interest are presented in Table 4.2. The heritability estimate for tick count obtained in the current study was low (0.03) (Table 4.2). For growth traits, estimated genomic heritability was 0.06, 0.25, and 0.23 for birth, weaning and yearling weight,

respectively. All the heritability estimates presented in Table 4.2 were estimated using the mmscore model during genome-wide association analysis.

Table 4.4.2: Suggestive SNPs associated with tick resistance in F_2 Angus x Nguni crossbred cattle.

Trait	h_2	SNP name	Chr	Position	A1	A2	Pc1df
TCOUNT	0.03	BovineHD0400032426	4	112551883	G	A	3.80e-05
		BovineHD2200005710	22	19635553	C	T	4.81e-05
BWT	0.06	Hapmap38587-BTA-23830	12	51067836	C	A	1.95e-07
		BovineHD1200022868	12	79955065	A	G	1.19e-06
		BTA-22732-no-rs	12	46739954	C	T	1.85e-06
		ARS-BFGL-BAC-15043	12	82199690	T	G	2.64e-06
		BovineHD1200023142	12	80884150	T	C	4.89e-06
		BovineHD1200023153	12	80926254	C	T	4.89e-06
		Hapmap60388-rs29016486	12	84095269	G	A	1.14e-05
		BTA-22711-no-rs	12	45649503	G	A	3.50e-05
		BovineHD0300032874	3	113815818	C	T	5.58E-05
WWT	0.25	BovineHD1800000097	18	840780	C	T	4.06e-05
		ARS-BFGL-NGS-18307	18	156809	C	T	5.02e-05
YWT	0.23	BovineHD0800026814	8	90226480	C	T	2.56e-05
STHICK	0.37	BTB-00360130	8	78145059	G	A	1.83e-05
		Hapmap51837-BTA-49470	7	108006295	T	C	2.00e-05
		BovineHD0900022588	9	81302552	T	C	3.40e-05
		BovineHD0200012180	2	42082588	T	G	4.41e-05
		BovineHD1200012955	12	47139521	C	T	5.89e-05

TCOUNT- Tick Count; BWT-Birth Weight; WWT-Weaning Weight; YWT-Yearling Weight; STHICK-Skin Thickness; Chr- Chromosome, h^2 –Heritability

4.4 Discussion

The main benefit of genome-wide association analysis is the ability to discover genetic determinants associated with the expression of the trait of interest without prior knowledge of the location of the region on the genome and its function. The current GWAS was conducted using the F_2 Angus and Nguni crossbred population. The F_2 design is well known for being suitable to map quantitative traits loci segregating from the parental lines, especially for complex traits [26,27]. The F_2 design considered in this study was, therefore, suitable to study complex traits such as tick count, growth, and skin thickness using genome-wide association analysis. The association analysis capitalizes on the extent of linkage disequilibrium (LD) that exists within the population, population size, the genetic architecture of the trait, and environmental factors.

Findings presented in Figure 4.4 show a linkage disequilibrium that is fast decaying as the distance increases. According to Stratz et al. [28], F_2-designed populations have reduced length of LD blocks compared to pure lines, thus justifying the application of GWAS. The fast decaying of LD has the potential to positively affect the resolution of genome-wide association analysis. However, the LD decay in crossbreed cattle has not been well studied. The implementation of studies focused on investigating the nature of LD in crossbreed cattle, specifically F_2 populations is therefore important. These studies provide an understanding of LD architecture, which might be useful in developing more informed GWAS models to elucidate genetic determinants associated with traits of economic importance. Research studies focused on studying economic traits using F_2 populations have been well explored in plants and chicken research, therefore more information is available on these species as compared to cattle studies. It must be noted that developing cattle F_2 populations for research purposes may be costly and time-consuming. Such barriers may be the reason behind the limited availability of F_2 population studies in beef cattle research.

In GWAS, the use of linear mixed models is applicable for a study population made of related individuals and it is regarded as a comprehensive approach to correct for the inflation of false positives, population stratification, and polygenic background [29–32]. For this study, linear mixed models were used for all association analyses whilst adjusting for genetic structure and relatedness amongst the F_2 Angus x Nguni cattle. The presence of population stratification was observed in the current study population. This phenomenon is regarded as one of the main confounding effects that affect the success of GWAS [33]. If it is not accounted for, the presence of population stratification increases the chances of discovering false positive results and reduces the chances of discovering true associations. Therefore, it is crucial to take

population stratification and relatedness into consideration when performing GWA analyses. In this study, the assessment of population stratification revealed that the population was clustered into three groups (Figure 4.5B). The clustering of F_2 Angus x Nguni cattle indicates the three genotypes that were created through crossbreeding the F_1 cattle population. The clustering of data was further confirmed by the kinship matrix outcomes presented by the dendrogram tree (Figure 4.6), which showed three clades. The incorporation of genetic structure, kinship matrix, and fixed effects in all study association models, permitted a better fitting of GWA models for all the traits.

There is extensive information in literature noting that genotyping costs are decreasing, making it possible to conduct genomic studies [34,35]. In Africa, conducting genomic studies remains a challenge, since the cost of genotyping is still high. In this study, the genotyping of samples was conducted in batches. Genotyping data in batches gives rise to the batch effect. The batch effect is a confounding effect and is accounted for through the application of stringent quality control on genotyped data and treating batches independently before the execution of the actual GWAS analysis. The stringent quality control does not completely account for this confounding effect; instead, it reduces the number of SNPs available to carry out the association analysis. This could result in losing important SNPs influencing the trait.

4.4.1 Tick count

The genome-wide association analysis showed that no SNP passed the genome-wide significant line. However, two SNPs passed the suggestive line. The SNPs are located on BT4 (BovineHD0400032426) and BTA 22 (BovineHD2200005710). The association of these SNPs with tick count have not been reported previously in the literature. However, BTA 4 and BTA 22 have been reported to harbour genomic regions associated with the expression of the tick count in a study by Sellore et al. [14]. The two SNPs explained a small genetic variance, which indicated that the tick count is highly polygenetic. The discovery of genomic regions associated with tick count in cattle remains a challenge, due to the nature of the trait, the limited resources associated with the generation of phenotypes and genotyping, the lack of standardized methods to measure the phenotype, and the selection of appropriate association models to capture the variation explaining the trait in a population. Tick count is polygenic, meaning that its expression is influenced by various genomic determinants [7,8,36]. The GWAS power for polygenetic traits is reduced when the population is small, therefore, genetic interactions (epistasis effects) could be the reason for the observed findings. The heritability estimate obtained via the genome-wide association model was 0.03, which is considerably low when compared to the estimates obtained from an animal model fitting a kinship matrix

(Table 4.1). This difference could be due to the issue of missing heritability, which is also explained by epistasis, epigenetics, and other biological effects. The incorporation of microbiome GWAS for cattle could solve the issue related to missing heritability [37]. The inclusion of the microbiome in human GWAS showed that the microbiome is associated with many traits, and it encodes for extra genes which interact with the host genes [38]. In the current study, tick count was measured using calico bags, which is a method that has not been well explored. The method has been successfully used in a study by Marima et al. [39], which was aimed at investigating gene expression and immunological responses of the Nguni, Brahman, and Angus cattle artificially infested with *Rhipicephalus microplus* and *R. decoloratus*.

4.4.2 Growth performance traits

Growth traits such as live body weights are of economic importance in beef production [40], because of their role in determining the efficiency of a production system [41]. Live body weights are associated with meat production and are regarded as an integral part of the breeding goals. These traits are included in traditional breeding schemes and they are known to exhibit moderate to high heritability [42,43].

The availability of SNP genotype data has allowed the discovery of genetic determinants influencing growth traits using the GWAS approach. Studies have been explored in different regions using various cattle breeds such as Colombian Brahman [40], Canchim [44], Nellore [41], Braunvieh [45], and Simmental [46]. Although the information from such studies is available, it can be noted that similar studies are not common in the African region, especially on crossbred populations. Heritability estimates obtained using a genome-wide association model were 0.06, 0.25, and 0.23 for birth weight, weaning weight, and yearling weight, respectively.

The genome-wide association analysis identified nine SNPs of which one SNP passed the genome-wide significant level while eight passed the suggestive SNP level for birth weight. The nine SNPs were discovered on BTA3 and BTA12. For weaning weight, six SNPs located on BTA5, BTA8, and BTA18 were suggestively associated with the trait. BTA8 has previously been associated with weaning weight in Brahman [40]. Additionally, only one SNP on BTA 8 was suggestively associated with yearling weight.

4.4.3 Skin thickness

Skin thickness serves as a barrier to protect the host from tick bites. In cattle, if the skin is thick, it is believed to give the animals protection by making it difficult for ticks to bite. In this study, skin thickness showed no genetic associated with the level of tick load. It has been observed that skin thickness plays a key role in controlling the level of tick load in cattle [47,48]. Therefore, it is important to study the genetic basis of skin thickness so that this trait can be considered in the breeding goal to control tick infestation in cattle.

Information on heritability estimates and genetic determinants influencing this trait is extremely limited.The estimated genomic heritability observed for skin thickness was 0.37 in the currebnt study. Moreover, genome-wide association analysis revealed 5 SNPs that were suggestively associated with skin thickness in F_2 Angus x Nguni cattle. The observed SNPs were located on BTA2, BTA7, BTA8, BTA9, and BTA12. Further studies are required to validate the identified genomic regions for skin thickness in cattle.

4.5 Conclusions

The GWA analysis identified SNPs associated with tick count, growth traits, and skin thickness in F_2 Angus x Nguni population. These results provide baseline information that could be used to inform future research aimed at generating information, which will inform breeding programs to improve tick resistance. The small population size, exclusion of G x E interaction in the model, the size of the SNP genotyping platform used, and the nature of the traits could be the plausible explanation for non-significant genome-wide SNPs for tick count, weaning weight, yearling weight, and skin thickness. Additionally, there is a need for more studies with large population sizes. There is a need for post-GWAS analysis to understand the biological relevance of identified SNPs on the expression of tick count, growth traits, and skin thickness.

4.6 References

[1] Castro-Janer E, Martins JR, Mendes MC, Namindome A, Klafke GM, Schumaker TTS. Diagnoses of fipronil resistance in Brazilian cattle ticks (Rhipicephalus (Boophilus) microplus) using in vitro larval bioassays. Vet Parasitol 2010;173:300–6. https://doi.org/10.1016/j.vetpar.2010.06.036.

[2] Porto Neto LR, Bunch RJ, Harrison BE, Barendse W. DNA variation in the gene ELTD1 is associated with tick burden in cattle. Anim Genet 2011;42:50–5. https://doi.org/10.1111/j.1365-2052.2010.02120.x.

[3] Mondal DB, Sarma K, Saravanan M. Upcoming of the integrated tick control program of ruminants with special emphasis on livestock farming system in India. Ticks Tick-Borne Dis 2013;4:1–10. https://doi.org/10.1016/j.ttbdis.2012.05.006.

[4] Machado MA, S Azevedo AL, Teodoro RL, Pires MA, CD Peixoto MG, de Freitas C, et al. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (Bos taurus × Bos indicus). BMC Genomics 2010;11:280. https://doi.org/10.1186/1471-2164-11-280.

[5] Frisch JE, O'Neill CJ. Comparative evaluation of beef cattle breeds of African, European and Indian origins. 1. Live weights and heterosis at birth, weaning and 18 months. Anim Sci 1998;67:27–38. https://doi.org/10.1017/S1357729800009760.

[6] Budeli MA, Nephawe KA, Norris D, Selapa NW, Bergh L, Maiwashe A. Genetic parameter estimates for tick resistance in Bonsmara cattle. South Afr J Anim Sci 2009;39:321–7. https://doi.org/10.10520/EJC94629.

[7] Mapholi NO, Maiwashe A, Matika O, Riggio V, Bishop SC, MacNeil MD, et al. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks Tick-Borne Dis 2016;7:487–97. https://doi.org/10.1016/j.ttbdis.2016.02.005.

[8] Otto PI, Guimarães SEF, Verardo LL, Azevedo ALS, Vandenplas J, Soares ACC, et al. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J Dairy Sci 2018;101:11020–32. https://doi.org/10.3168/jds.2017-14223.

[9] Burrow HM, Gulbransen B, Johnson SK, Davis GP, Shorthose WR, Elliott RF. Consequences of selection for growth and heatresistance on growth, feed conversion efficiency, commercialcarcass traits and meat quality of zebu crossbred cattle. Aust JAgric Res 1991;43:1373—138.

[10] Carrera JPB. Description: Genetic parameters for resistance to ticks, gastrointestinal helminths and Eimeria spp. and perspectives on the use of selection in Nelore cattle 2013. http://bdtd.ibict.br/vufind/Record/UFMG_bb85858d4bac1021a17bfbfb9a58acf1 (accessed February 17, 2022).

[11] Alqudah AM, Sallam A, Stephen Baenziger P, Börner A. GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley – A review. J Adv Res 2020;22:119–35. https://doi.org/10.1016/j.jare.2019.10.013.

[12] Welderufael BG, Løvendahl P, de Koning D-J, Janss LLG, Fikse WF. Genome-Wide Association Study for Susceptibility to and Recoverability From Mastitis in Danish Holstein Cows. Front Genet 2018;9:141. https://doi.org/10.3389/fgene.2018.00141.

[13] Tiezzi F, Parker-Gaddis KL, Cole JB, Clay JS, Maltecca C. A genome-wide association study for clinical mastitis in first parity US Holstein cows using single-step approach and

genomic matrix re-weighting procedure. PloS One 2015;10:e0114919. https://doi.org/10.1371/journal.pone.0114919.

[14] Sollero BP, Junqueira VS, Gomes CCG, Caetano AR, Cardoso FF. Tag SNP selection for prediction of tick resistance in Brazilian Braford and Hereford cattle breeds using Bayesian methods. Genet Sel Evol 2017;49:49. https://doi.org/10.1186/s12711-017-0325-2.

[15] Ledur MC, Navarro N, Pérez-Enciso M. Large-scale SNP genotyping in crosses between outbred lines: how useful is it? Heredity 2010;105:173–82. https://doi.org/10.1038/hdy.2009.149.

[16] Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory; 1989.

[17] Nicolazzi EL, Marras G, Stella A. SNPConvert: SNP Array Standardization and Integration in Livestock Species. Microarrays 2016;5:17. https://doi.org/10.3390/microarrays5020017.

[18] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet 2007;81:559–75.

[19] Aulchenko YS, de Koning D-J, Haley C. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. Genetics 2007;177:577–85. https://doi.org/10.1534/genetics.107.075614.

[20] ASReml-3-User-Guide.pdf n.d.

[21] VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci 2008;91:4414–23. https://doi.org/10.3168/jds.2007-0980.

[22] Chen W-M, Abecasis GR. Family-based association tests for genomewide association scans. Am J Hum Genet 2007;81:913–26. https://doi.org/10.1086/521580.

[23] Guo Y, Huang Y, Hou L, Ma J, Chen C, Ai H, et al. Genome-wide detection of genetic markers associated with growth and fatness in four pig populations using four approaches. Genet Sel Evol 2017;49:21. https://doi.org/10.1186/s12711-017-0295-4.

[24] Abanda B, Schmid M, Paguem A, Iffland H, Preuß S, Renz A, et al. Genetic Analyses and Genome-Wide Association Studies on Pathogen Resistance of Bos taurus and Bos indicus Cattle Breeds in Cameroon. Genes 2021;12:976. https://doi.org/10.3390/genes12070976.

[25] Van Grevenhof EM, Van Arendonk JA, Bijma P. Response to genomic selection: The Bulmer effect and the potential of genomic selection when the number of phenotypic records is limiting. Genet Sel Evol 2012;44:26. https://doi.org/10.1186/1297-9686-44-26.

[26] Edwards DB, Ernst CW, Tempelman RJ, Rosa GJM, Raney NE, Hoge MD, et al. Quantitative trait loci mapping in an F2 Duroc x Pietrain resource population: I. Growth traits. J Anim Sci 2008;86:241–53. https://doi.org/10.2527/jas.2006-625.

[27] Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 2013;9:29. https://doi.org/10.1186/1746-4811-9-29.

[28] Stratz P, Schmid M, Wellmann R, Preuß S, Blaj I, Tetens J, et al. Linkage disequilibrium pattern and genome-wide association mapping for meat traits in multiple porcine F2 crosses. Anim Genet 2018;49:403–12. https://doi.org/10.1111/age.12684.

[29] Chen H, Wang C, Conomos MP, Stilp AM, Li Z, Sofer T, et al. Control for Population Structure and Relatedness for Binary Traits in Genetic Association Studies via Logistic Mixed Models. Am J Hum Genet 2016;98:653–66. https://doi.org/10.1016/j.ajhg.2016.02.012.

[30] Price AL, Zaitlen NA, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. Nat Rev Genet 2010;11:459–63. https://doi.org/10.1038/nrg2813.

[31] Zhang Z, Ersoz E, Lai C-Q, Todhunter RJ, Tiwari HK, Gore MA, et al. Mixed linear model approach adapted for genome-wide association studies. Nat Genet 2010;42:355–60. https://doi.org/10.1038/ng.546.

[32] Wang S-B, Feng J-Y, Ren W-L, Huang B, Zhou L, Wen Y-J, et al. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. Sci Rep 2016;6:19444. https://doi.org/10.1038/srep19444.

[33] Wacholder S, Rothman N, Caporaso N. Population Stratification in Epidemiologic Studies of Common Genetic Variants and Cancer: Quantification of Bias. JNCI J Natl Cancer Inst 2000;92:1151–8. https://doi.org/10.1093/jnci/92.14.1151.

[34] Kumar S, Banks TW, Cloutier S. SNP Discovery through Next-Generation Sequencing and Its Applications. Int J Plant Genomics 2012;2012:e831460. https://doi.org/10.1155/2012/831460.

[35] Koopaee HK, Koshkoiyeh AE. SNPs genotyping technologies and their applications in farm animals breeding programs: review. Braz Arch Biol Technol 2014;57:87–95. https://doi.org/10.1590/S1516-89132014000100013.

[36] Shyma KP, Gupta JP, Singh V. Breeding strategies for tick resistance in tropical cattle: a sustainable approach for tick control. J Parasit Dis Off Organ Indian Soc Parasitol 2015;39:1–6. https://doi.org/10.1007/s12639-013-0294-5.

[37] Mkize N, Maiwashe A, Dzama K, Dube B, Mapholi N. Suitability of GWAS as a Tool to Discover SNPs Associated with Tick Resistance in Cattle: A Review. Pathogens 2021;10:1604. https://doi.org/10.3390/pathogens10121604.

[38] Sandoval-Motta S, Aldana M, Martínez-Romero E, Frank A. The Human Microbiome and the Missing Heritability Problem. Front Genet 2017;8:80. https://doi.org/10.3389/fgene.2017.00080.

[39] Marima JK, Nel CL, Marufu MC, Jonsson NN, Dube B, Dzama K. A genetic and immunological comparison of tick-resistance in beef cattle following artificial infestation with Rhipicephalus ticks. Exp Appl Acarol 2020;80:569–90. https://doi.org/10.1007/s10493-020-00480-8.

[40] Martínez R, Bejarano D, Gómez Y, Dasoneville R, Jiménez A, Even G, et al. Genome-wide association study for birth, weaning and yearling weight in Colombian Brahman cattle. Genet Mol Biol 2017;40:453–9. https://doi.org/10.1590/1678-4685-GMB-2016-0017.

[41] Terakado APN, Costa RB, de Camargo GMF, Irano N, Bresolin T, Takada L, et al. Genome-wide association study for growth traits in Nelore cattle. Animal 2018;12:1358–62. https://doi.org/10.1017/S1751731117003068.

[42] Frizzas OG, Grossi DA, Buzanskas ME, Paz CCP, Bezerra LAF, Lôbo RB, et al. Heritability estimates and genetic correlations for body weight and scrotal circumference adjusted to 12 and 18 months of age for male Nellore cattle. Animal 2009;3:347–51. https://doi.org/10.1017/S175173110800373X.

[43] Baldi F, Albuquerque LG, Alencar MM. Random regression models on Legendre polynomials to estimate genetic parameters for weights from birth to adult age in Canchim cattle. J Anim Breed Genet Z Tierzuchtung Zuchtungsbiologie 2010;127:289–99. https://doi.org/10.1111/j.1439-0388.2010.00853.x.

[44] Buzanskas ME, Grossi DA, Ventura RV, Schenkel FS, Sargolzaei M, Meirelles SLC, et al. Genome-Wide Association for Growth Traits in Canchim Beef Cattle. PLOS ONE 2014;9:e94802. https://doi.org/10.1371/journal.pone.0094802.

[45] Zepeda-Batista JL, Núñez-Domínguez R, Ramírez-Valverde R, Jahuey-Martínez FJ, Herrera-Ojeda JB, Parra-Bracamonte GM. Discovering of Genomic Variations Associated to Growth Traits by GWAS in Braunvieh Cattle. Genes 2021;12:1666. https://doi.org/10.3390/genes12111666.

[46] Braz CU, Rowan TN, Schnabel RD, Decker JE. Genome-wide association analyses identify genotype-by-environment interactions of growth traits in Simmental cattle. Sci Rep 2021;11:13335. https://doi.org/10.1038/s41598-021-92455-x.

[47] Foster L, Fourie P, Neser F, Fair M. Differences in physical traits such as coat score and hide-thickness together with tick burdens and body condition score in four breeds in the Southern Free State 2013.

[48] Fourie PJ, Foster LA, Neser FWC. Score and hide-thickness, together with tick burden and body condition score, in four cattle breeds in the South-eastern Free State province of South Africa. J New Gener Sci 2013;11:66–73. https://doi.org/10.10520/EJC141432.

[49] Maiorano AM, Oliveira MCS, Cyrillo JNSG, Albuquerque LG, Curi RA, Silva JAI. Genetic study of skin thickness and its association with postweaning growth in Nellore cattle: estimation of the genetic parameters. Genet Mol Res GMR 2016;15. https://doi.org/10.4238/gmr.15017124.

[50] Zhang H, Liu A, Li X, Xu W, Shi R, Luo H, et al. Genetic analysis of skinfold thickness and its association with body condition score and milk production traits in Chinese Holstein population. J Dairy Sci 2019;102:2347–52. https://doi.org/10.3168/jds.2018-15180.

[51] Hamid MA, Hussain SMI, Khan MKI, Islam MN, Biswas M a. A. Skin thickness in relation to milk production of crossbred cows. Pak J Biol Sci Pak 2000.

5 Chapter 5:
Post-GWAS analysis to find genes associated with tick count, growth traits, and skin thickness

Abstract: Genome-wide association studies lack the capacity to explain the biological mechanisms influencing economic traits in livestock, therefore post-GWAS analysis are required. The elucidation of the biological mechanisms underlying tick resistance, growth traits, and skin thickness in cattle is crucial for the improvement of these traits through genetic selection. Therefore, the current study performed a post-GWAS functional analysis to understand biological mechanisms underlying tick count, growth traits, and skin thickness in F_2 Angus X Nguni cattle. Functional annotation, pathway and cluster analysis were formed using DAVID and ShinyGo bioinformatics resources. Fisher's exact test with False discovery rate adjustment was used and the statistical significance was considered at p (DFR) <0.05.

The gene ontology (GO) terms that were enriched by the genes associated with tick count included, cellular processes, regulation of biological processes, and response to a stimulus. Consequently, genes *ZNF746*, *GIMAP8*, and *RARRES2* were identified as potential biomarkers for tick count control in cattle. For growth traits, the enriched GO terms and pathways related to ubiquinol-cytochrome-c reductase activity, plasma membrane raft, DNA ligase, and ATP dependent. These categories are associated with cell differentiation, skeletal muscle development, and metabolism regulation. Therefore, genes *TRPM8*, *SPP2*, *UQCRFS1*, *MB*, *TMEFF1*, *CAVIN4*, and *MSANTD3* were identified as candidate genes controlling growth traits. For skin thickness, the enriched GO terms and pathways related to cell differentiation and transmembrane receptor tyrosine-protein kinase. The categories were enriched by *FER* and *NTRK2* genes, which were regarded as potential candidate genes for skin thickness in the study population.

These findings need to be validated through fine mapping and the implementation of more studies. To conclude, biological mechanisms uncovered from this study hold promise for the genetic improvement of the traits of interest through selective breeding.

Keywords: immune response, development of skeletal muscles, cell differentiation, metabolism

5.1 Introduction

Genome-wide association studies (GWAS) have been instrumental in the discovery of SNPs relating to the expression of economic traits in beef cattle. Nevertheless, information on the biological processes underlying the expression of these traits is still limited [1]. Genome-wide association analyses produce summary statistics that do not provide insight into the biological mechanisms underlying the economic traits of interest. Additionally, the genome-wide association analyses do not take into consideration that the expression of traits is a byproduct of gene networks [2]. The other components limiting the elucidation of complex traits through GWAS include the issue of epistasis and the bi-allelic nature of the SNPs which could attribute to the failure of capturing multi-allelic quantitative trait loci (QTL) [2]. Therefore, post-GWAS analyses such as gene-set enrichment, pathway, and gene network analyses are required as downstream analyses to elucidate the biological information relating to the complex nature of traits [3–8].

A post-GWAS analyses are a crucial step for mapping causal variants and providing a better understanding of biological processes underlying the trait beyond the SNPs [9]. Amongst the bioinformatics platforms which are available to perform gene set enrichment and pathways analyses, DAVID (The Database for Annotation, Visualization, and Integrated Discovery) is the most preferred. DAVID is a web-based platform that enables researchers to generate biological context for large gene lists associated with traits in a broad range of species [10–12]. The upper hand for this tool over the others is its ability to derive information from diverse sources such as UniProt, KEGG, REACTOME, and NCBI [13].

Through post-genome-wide association analyses, previous studies have uncovered candidate genes and gene ontology (GO) terms associated with tick resistance in cattle. A study by Otto et al. [14] discovered genes (*MYOSA, TREML1, CD83, PRSS16, TREM1,* and *TREM2*) relating to tick resistance in the F_2 Gir × Holstein population. Moreover, the study discovered gene network interactions that were associated with the immune system. This warranted the need to validate the identified genes for their effects on the tick resistance phenotype in cattle. Another study by Mapholi et al. [15] uncovered genes (*KCNQ4, TRPM8, CSN1S2, PRkG1, and GPR142*) relating to tick resistance in the Nguni breed, and concluded that their findings provide the potential for marker-assisted selection in Nguni cattle. A systematic review by dos Santos et al. [1] presented a summary of post-GWAS analyses for tick resistance in cattle. The review provided consolidated information on candidate genes and biological evidence relating to tick resistance in cattle from numerous studies. Currently, the available biological

information relating to tick resistance is based on the study of pure-breeds as compared to crossbred animals. Studies on crossbred animals are limited.

Growth traits and skin thickness form part of the factors affecting tick resistance in cattle. The role of growth traits and skin thickness in tick resistance necessitates the need to study the biological mechanisms influencing these traits. Skin thickness is a trait not well studied; therefore, knowledge of the biological mechanisms influencing this trait in cattle is scarce. Candidate genes and biological networks associated with the expression of growth traits in cattle have been uncovered in previous studies. In their studies Buzanskas et al. [16] (*DPP6, MANEA, LOC783932*); and Du et al. [17] (*FGF4, ITGA5, PLA2G4A, ANGPT4*) found candidate genes that were associated with birth weight in cattle. Candidate genes associated with weaning weight were identified by Buzanksas et al. [16] (*FARSB, RALGDS, GTF3C5*), and Smith et al.[18] (*LCORL, SLIT2, GRD2, FAM1848, MOS, CCSER10*). Furthermore, in their study Buzanskas et al. [16] observed that *MARCH3, PHAX, ALDA7A1, GRAMD3, MIR2458* and *LOC1008484523* were associated with yearly weight.

The discoveries from such studies hold the potential to inform the development of alternative tick-control strategies and improvement of tick count, growth, and skin thickness traits through selective breeding. Despite the availability of knowledge from previous studies, information related to the biological mechanisms controlling tick count, growth traits, and skin thickness is still limited. Therefore, an improved understanding of the mechanisms underlying tick count, growth traits and skin thickness at a cellular, molecular, and biological level could contribute positively to the establishment of schemes aimed to improve and monitor these economic traits.

Therefore, this chapter aimed to perform a post-GWAS analysis to map candidate genes closest to the SNPs of interest and to perform functional annotation analyses for genes related to tick count, growth traits, and skin thickness in F_2 Angus x Nguni cattle.

5.2 Materials and methods

5.2.1 Bioinformatics

Suggestive SNPs from the GWAS analyses were used as foundational information to find candidate genes associated with the traits of interest. The Ensembl Genome Browser 107.12 [19] was used to locate genes nearest to the suggestive SNPs using the *Bos taurus* reference genome (Assembly ARS-UCD1.2) [20]. The genes located nearest to the SNPs of interest

within a 1Mb region were considered in creating the gene list for all the traits of interest independently.

5.2.2 Gene functional annotation

The created gene lists were then used for functional, pathway, and cluster analyses in DAVID bioinformatics resources version 2021 (https://david.ncifcrf.gov/summary.jsp) using *Bos taurus* as a background [21]. Fisher's exact test with False discovery rate (FDR) adjustment was used and statistical significance was considered at p (FDR) <0.05. The bioinformatics resources used to investigate previous relations of genes to the traits of interest were UniprotK (https://www.uniprot.org/uniprotkb) [22], NCBI (https://www.ncbi.nlm.nih.gov/) [23], gene card (https://www.genecards.org) [24], and Animal QTLdb (https://www.animalgenome.org/cgi-bin/QTLdb/index) [25]. ShinyGO v0.741 (http://bioinformatics.sdstate.edu/go74/)was also used for functional annotation for all the traits. On this platform, the *Bos taurus* assembly that was used is STRING.9913. *Bos* and default settings were used as background information and p (FDR) <0.05 considered as significant.

5.3 Results

5.3.1 Tick count

For the tick count, the focus was on two suggestive SNPs located on BTA4 and BTA22. The suggestive position on BTA4 was associated with 27 genes and the BTA22 position was associated with the two nearest genes. Figure 5.1 shows a novel transcripts in BTA4 and BTA22, showing candidate genes closest to the suggestive SNPs associated with tick count in F_2 Angus x Nguni. The description of the genes is presented in Table 5.1. Seven genes were uncharacterized, while genes belonging to the zinc finger protein were abundant, followed by the RNA family, GTPase, IMAP family member.

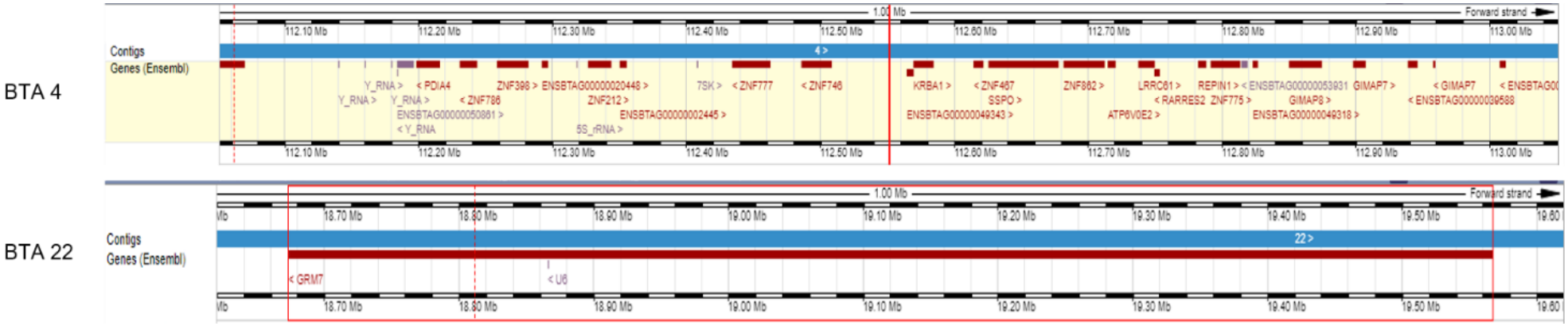

Figure 5.1: A screenshot showing the nearest candidate genes within the 1 Mb region surrounding the suggestive SNPs for tick count in BTA4 and 22.

Table 5.1: List of identified genes within the 1 Mb region surrounding the suggestive SNPs for tick count.

BTA	Position	Ensembl ID	Nearest gene	Description
4	112551883	ENSBTAG00000042081	*Y_RNA*	Y RNA
		ENSBTAG00000017143	*PDIA4*	Protein disulfide isomerase family A member 4
		ENSBTAG00000020445	*ZNF398*	Zinc finger protein 398
		ENSBTAG00000014768	*ZNF786*	Zinc finger protein 786
		ENSBTAG00000000389	*ZNF212*	Zinc finger protein 212
		ENSBTAG00000050861	Unknown	Uncharacterized
		ENSBTAG00000055284	*5S_rRNA*	5S ribosomal RNA
		ENSBTAG00000020448	Unknown	Uncharacterized
		ENSBTAG00000051861	*7SK*	7SK RNA
		ENSBTAG00000002810	*ZNF777*	Zinc finger protein 777
		ENSBTAG00000031106	*ZNF746*	Zinc finger protein 746
		ENSBTAG0000002445	Unknown	Uncharacterized
		ENSBTAG00000010219	*KRBA1*	KRAB-A domain containing 1
		ENSBTAG00000008542	*SSPO*	SCO-spondin
		ENSBTAG00000019844	*ZNF467*	Zinc finger protein 467
		ENSBTAG00000039594	*ZNF862*	Zinc finger protein 862
		ENSBTAG00000049343	Unknown	Uncharacterized
		ENSBTAG00000031059	*LRRC61*	Leucine-rich repeat containing 61
		ENSBTAG00000038241	*REPIN1*	Replication initiator 1
		ENSBTAG00000006022	*ATP6V0E2*	ATPase H+ transporting V0 subunit e2
		ENSBTAG00000004215	*RARRES2*	Retinoic acid receptor responder 2
		ENSBTAG00000007983	*ZNF775*	Zinc finger protein 775
		ENSBTAG00000053931	Unknown	Uncharacterized
		ENSBTAG00000030940	*GIMAP7*	GTPase, IMAP family member 7
		ENSBTAG00000014402	*GIMAP8*	GTPase, IMAP family member 8
		ENSBTAG00000049318	Unknown	Uncharacterized
		ENSBTAG00000039588	Unknown	Uncharacterized
22	19635553	ENSBTAG00000013047	*GRM7-201*	Glutamate metabotropic receptor 7
		ENSBTAT00000059985.2	*U6-201*	U6 spliceosomal RNA

The DAVID functional annotation for these genes uncovered four Gene Ontology terms and 13 pathways significantly enriched for tick count (Table 5.2). The observed GO terms and

pathways were sorted according to the most significant p (FDR) value. Amongst the GO terms enriched, three belonged to molecular function (GO: 0005525, GO: 0003700, GO: 0000978), and one belonged to biological process (GO: 0006357).

Table 5.2: Gene Ontology (GO) terms and pathways significantly enriched using genes associated with a tick count

Category	Term	Genes	Count	PValue	FE	FDR
INTERPRO	IPR006703:AIG1	*GIMAP7, GIMAP8*	5	2.39E-10	438.25	6.46E-09
UP_SEQ_FEATURE	DOMAIN: AIG1-type G	*GIMAP7, GIMAP8*	5	3.04E-10	416.30	3.07E-08
INTERPRO	IPR022137: Protein of unknown function DUF3669, zinc finger protein	*ZNF746,ZNF777, ZNF212,ZNF398*	4	7.56E-09	818.07	1.02E-07
SMART	SM00355: ZnF_C2H2	*REPIN1,ZNF746, ZNF777,ZNF77,ZNF467,ZNF786, ZNF212,ZNF398*	8	2.26E-07	12.49	2.71E-06
INTERPRO	IPR013087: Zinc finger C2H2-type/integrase DNA-binding domain	*REPIN1,ZNF746, ZNF777,ZNF775, ZNF467,ZNF786, ZNF212,ZNF398*	8	2.43E-07	15.41	2.18E-06
GOTERM_BP_DIRECT	GO:0006357~regulation of transcription from RNA polymerase II promoter	*REPIN1,ZNF746, ZNF777ZNF775, ZNF467,ZNF786, ZNF212,ZNF398*	8	1.10E-05	7.74	3.85E-04
INTERPRO	IPR001909: Krueppel-associated box	*ZNF746,ZNF777, ZNF862,ZNF212, ZNF398*	5	2.49E-05	26.56	1.68E-04
UP_SEQ_FEATURE	DOMAIN: KRAB	*ZNF746,ZNF777, ZNF862,ZNF212, ZNF398*	5	2.89E-05	25.79	0.00
SMART	SM00349: KRAB	*ZNF746,ZNF777, ZNF862,ZNF212, ZNF398*	5	3.73E-05	21.44	0.00
KEGG_PATHWAY	bta05168: Herpes simplex virus 1 infection	*ZNF746,ZNF777, ZNF786,ZNF212, ZNF398*	5	6.23E-05	15.48	0.00
UP_SEQ_FEATURE	DOMAIN:C2H2-type	*ZNF777,ZNF775, ZNF467,ZNF786, ZNF212,ZNF398*	6	9.39E-05	11.60	0.00
GOTERM_MF_DIRECT	GO:0005525~GTP binding	*GIMAP7, GIMAP8*	5	4.11E-04	12.84	0.01
GOTERM_MF_DIRECT	GO:0003700~transcription factor activity, sequence-specific DNA binding	*ZNF746,ZNF467, ZNF212, ZNF398*	4	0.00	16.60	0.01
GOTERM_MF_DIRECT	GO:0000978~RNA polymerase II core promoter proximal region sequence-specific DNA binding	*ZNF746,ZNF775, ZNF467,ZNF786, ZNF212,ZNF398*	6	0.00	5.90	0.01
UP_KW_LIGAND	KW-0342~GTP-binding	*GIMAP7,GIMAP8*	5	0.00	6.77	0.02
INTERPRO	IPR027417:P-loop containing nucleoside triphosphate hydrolase	*GIMAP7,GIMAP8*	5	0.01	6.49	0.03
INTERPRO	IPR003655:Krueppel-associated box-related	*ZNF777, ZNF398*	2	0.01	306.78	0.03

GO- gene ontology, FDR- False discovery rate, FE-Fold enrichment

The clustering analysis grouped the enriched terms into two groups (Figure 5.2). A full description of the clusters is presented in Annexure 1.1. In Figure 5.2, the green color represents the positively associated gene terms within a group, and black represents unknown relationships. The enrichment score informs the order of annotation terms within the cluster. For cluster 1, the enriched GO terms are commonly related annotations for GTPase, and IMAP Family member genes as compared to Zinc finger protein. Cluster 2 shows a scattered pattern of green and black; this symbolizes functional differences amongst the enriched GO terms within the group.

GO terms and pathways enriched through using the ShinyGO bioinformatics source are shown in Table 5.3. The enriched terms are presented in the form of a network as presented in Figure 5.3, where each node depicts an enriched pathway. The size of the node corresponds to the number of genes involved. Additionally, lines connect the nodes, and the thickness of the line indicates the percentage of overlapping genes. Table 5.4 shows genes grouped by functional categories defined by high-level GO terms for tick count. The high-level GO category observed was the cellular process characterized by a group of six genes, and the lowest being one.

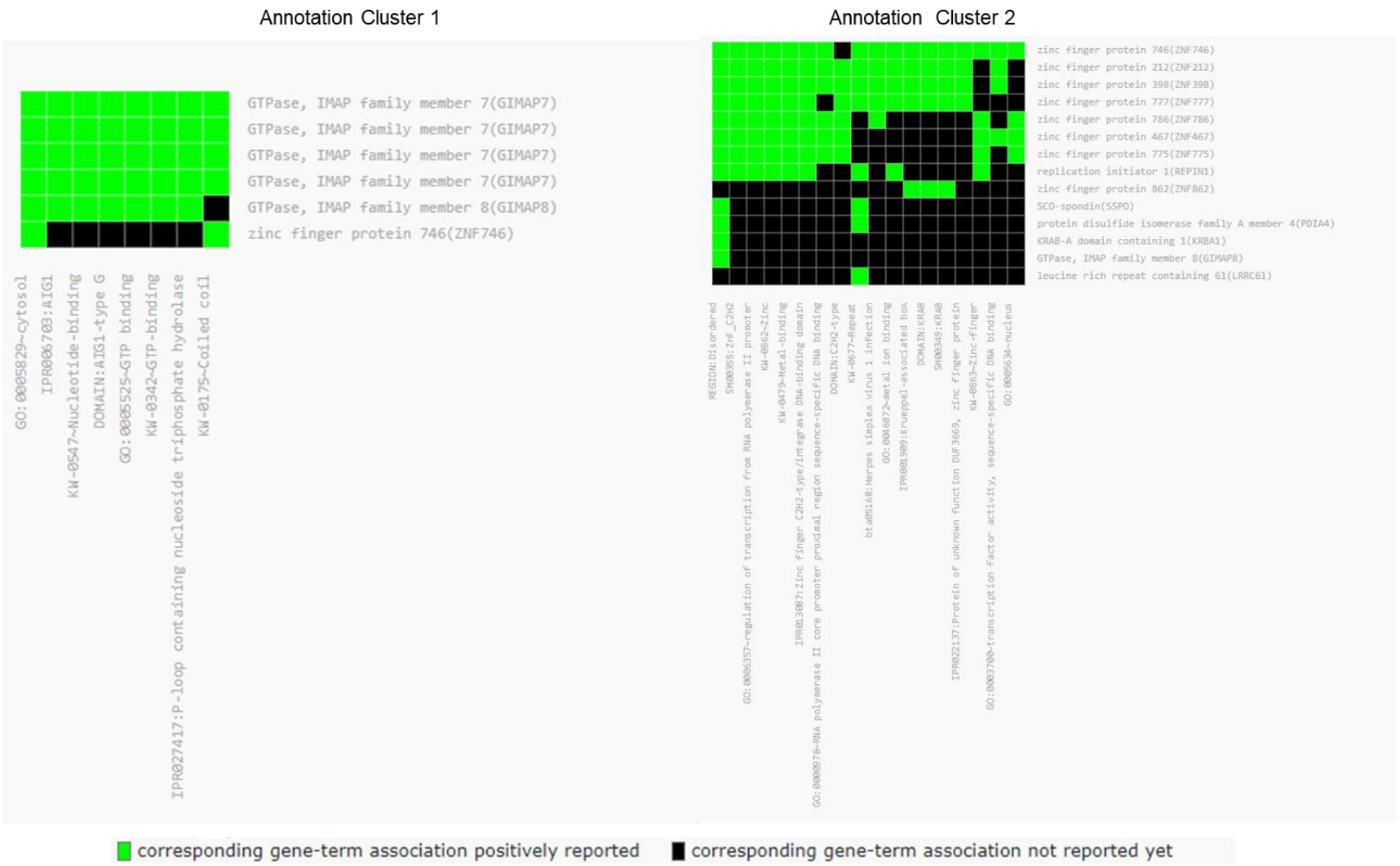

Figure 5.2. Gene to terms 2D heat map showing all related genes and their associated terms for annotation cluster 1 and cluster 2 genes for tick count using DAVID

Table 5.3: Enriched GO terms and pathways for tick count using ShinyGO bioinformatics source

Pathway	Genes	nGenes	Pathway Genes	Enrichment FDR	FE
Zinc finger protein	*ZNF212 ZNF777 ZNF746 ZNF398 ZNF282*	5	6	1.77E-12	833
Protein of unknown function DUF3669, zinc finger protein	*ZNF212 ZNF777 ZNF746 ZNF398*	4	5	9.22E-10	799.68
Protein of unknown function DUF3669, zinc finger protein, and Iguana/Dzip1-like DAZ-interacting protein N-terminal	*ZNF212 ZNF777 ZNF746 ZNF398*	4	6	1.84E-09	666.4
Zinc finger, C2H2 type	*ZNF212 ZNF777 ZNF746 ZNF786 ZNF467 ZNF398 ZNF775 REPIN1 ZNF282*	9	476	3.85E-08	18.9
KRAB box	*ZNF212 ZNF777 ZNF746 ZNF786 ZNF398 ZNF862 ZNF282*	7	209	7.20E-08	33.480
Krueppel associated box	*ZNF212 ZNF777 ZNF746 ZNF786 ZNF398 ZNF862 ZNF282*	7	206	7.20E-08	33.967
Zinc finger	*ZNF212 ZNF777 ZNF746 ZNF786 ZNF467 ZNF398 ZNF775 REPIN1 ZNF282*	9	561	9.35E-08	16.0367
Mixed, incl. protein of unknown function duf3669, zinc finger protein, and phosphatase and actin regulator 1	*ZNF212 ZNF777 ZNF746 ZNF398*	4	17	1.09E-07	235.200
Zinc finger C2H2 superfamily	*ZNF212 ZNF777 ZNF746 ZNF786 ZNF467 ZNF398 ZNF775 REPIN1*	8	444	3.12E-07	18.011
Mixed, incl. krba1 family repeat, and replication initiator 1	*KRBA1 LRRC61 REPIN1*	3	5	4.34E-07	599.760
Mostly uncharacterized, incl. protein of unknown function duf3669, zinc finger protein, and phosphatase and actin regulator 1	*ZNF212 ZNF777 ZNF746 ZNF398*	4	26	4.90E-07	153.785
Zinc finger C2H2-type	*ZNF212 ZNF777 ZNF746 ZNF786 ZNF467 ZNF398 ZNF775 REPIN1*	8	488	4.90E-07	16.387
KRAB domain superfamily	*ZNF212 ZNF777 ZNF746 ZNF398 ZNF862*	5	125	4.95E-06	39.984
Krueppel-associated box	*ZNF212 ZNF777 ZNF746 ZNF398 ZNF862*	5	131	5.80E-06	38.153
AIG1-type guanine nucleotide-binding (G) domain	*ENSBTAG00000030940 LOC511617 GIMAP8*	3	13	8.22E-06	230.677
AIG1 family	*ENSBTAG00000030940 LOC511617 GIMAP8*	3	14	9.81E-06	214.200
Mostly uncharacterized, incl. dbird complex, and trim37, math domain	*KRBA1 LRRC61 REPIN1*	3	16	1.42E-05	187.425

Mostly uncharacterized, incl. protein of unknown function duf3669, zinc finger protein, and transport and Golgi organization 2	*ZNF212 ZNF777 ZNF746 ZNF398*	4	75	2.38E-05	53.312
C2H2-type zinc finger	*ZNF777 ZNF746 ZNF467 ZNF775 REPIN1*	5	251	1.063E-05	19.912
Zinc-finger	*ZNF212 ZNF777 ZNF746 ZNF786 ZNF467 ZNF398 ZNF775 REPIN1*	8	1202	2.778E-04	6.653
Mostly uncharacterized, incl. class I histocompatibility antigen, nkg2d ligand, domains 1 and 2, and glyco_18	*KRBA1 LRRC61 REPIN1*	3	45	2.856E-04	66.640
Krueppel-associated box-related	*ZNF777 ZNF398*	2	6	3.279E-04	333.200
Mostly uncharacterized, incl. class I histocompatibility antigen, nkg2d ligand, domains 1 and 2, and armadillo-like	*KRBA1 LRRC61 REPIN1*	3	75	1.217E-033	39.984
Zinc	*ZNF212 ZNF777 ZNF746 ZNF786 ZNF467 ZNF398 ZNF775 REPIN1*	8	1566	1.574E-03	5.107
Repeat	*ZNF212 ZNF777 ZNF746 PDIA4 ZNF398 LRRC61 SSPO REPIN1*	8	1707	2.784E-03	4.685
C2H2-type zinc finger	*ZNF467 ZNF775 REPIN1 ZNF282*	4	286	3.234E-04	13.980
50S ribosome-binding GTPase	*ENSBTAG00000030940 LOC511617 GIMAP8*	3	117	3.889E-03	25.631
Zinc-finger of C2H2 type	*ZNF746 REPIN1 ZNF282*	3	125	4.557E-03	23.990
C2H2-type zinc-finger domain	*ZNF786 ZNF467*	2	37	1.084E-02	54.032
Metal-binding	*ZNF212 ZNF777 ZNF746 ZNF786 ZNF467 ZNF398 ZNF775 REPIN1*	8	2286	1.726E-02	3.498

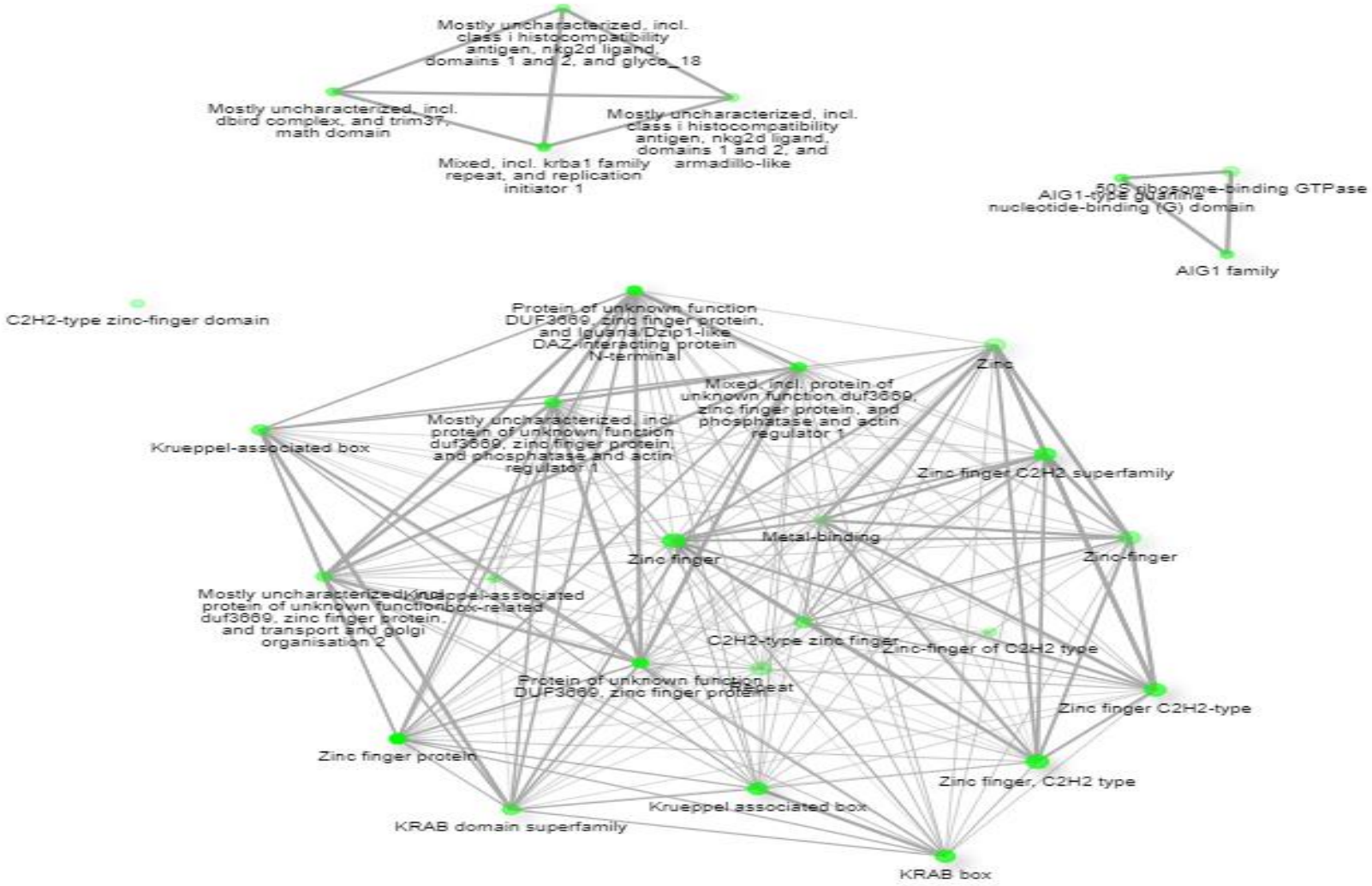

Figure 5.3: Network for tick count enriched GO terms and pathways using ShinyGo bioinformatics platform

Table 5.4: Tick count-related genes grouped by functional categories defined by high-level GO terms.

N	High-level GO category	Genes
6	Cellular processes	*RARRES2 ATP6V0E2 PDIA4 LRRC61 SSPO REPIN1*
4	Regulation of biological process	*RARRES2 ZNF746 PDIA4 SSPO*
4	Biological regulation	*RARRES2 ZNF746 PDIA4 SSPO*
4	Regulation of cellular process	*RARRES2 ZNF746 PDIA4 SSPO*
3	Binding	*RARRES2 ZNF746 REPIN1*
3	Organelle	*ZNF746 PDIA4 REPIN1*
3	Intracellular	*ZNF746 PDIA4 REPIN1*
2	Extracellular region	*RARRES2 SSPO*
2	Response to stimulus	*RARRES2 PDIA4*
2	Protein binding	*RARRES2 ZNF746*
2	Extracellular space	*RARRES2 SSPO*
2	Response to stress	*RARRES2 PDIA4*
2	Ion binding	*ZNF746 REPIN1*
2	Organelle lumen	*PDIA4 REPIN1*
2	Cellular metabolic process	*RARRES2 REPIN1*
2	Positive regulation of the biological process	*RARRES2 ZNF746*
2	Negative regulation of the biological process	*ZNF746 SSPO*
1	Immune response	*RARRES2*
1	Cell adhesion	*SSPO*
1	Biosynthetic process	*REPIN1*
1	Response to external stimulus	*RARRES2*
1	Response to biotic stimulus	*RARRES2*

5.3.2 Growth traits

The description for genes mapped to the nearest SNPs of interest is presented in Annexure 1.2. Functional annotation findings obtained through DAVID revealed only two GO terms significantly enriched for birth weight and yearling weight (Table 5.5) and none for weaning weight. Conversely, enrichment analyses using ShinyGo v0.741 revealed several GO terms enriched for birth, weaning, and yearly weight (Table 5.6, 5.7, and 5.8 respectively). The enriched categories for birth, weaning, and yearly weight are further depicted graphically in Figure 5.4. The enriched GO terms interactions for birth, weaning, and yearly weight are exhibited in Figures 5.5, 5.6, and 5.7, respectively. Few genes characterized birth weight

enrichment; therefore, the grouping of related genes according to the enrichment was not possible. Weaning weight-related genes were grouped according to functional categories defined by high-level GO terms (Table 5.9). Lastly, few genes characterized yearly weight; therefore, functional enrichment for this trait was not possible through both bioinformatics sources used.

Table 5.5: Enriched GO term for growth traits using DAVID

Trait	Category	Term	Genes	Count	%	PValue	FE
BWT	GOTERM_CC_DIRECT	GO:0005654~nucleoplasm	*LMO7, HJRP, LIG4, UCHL3*	4	40	0.03	4.80
YWT	GOTERM_BP_DIRECT	GO:0030154~cell differentiation	*TMEFF1, CAVIN4*	2	66.67	0.045	29.073

Table 5.6: Enriched GO terms and pathways terms for birth weight using ShinyGo

Pathway	Genes	Pathway Genes	Enrichment FDR	FE
DNA ligase, ATP-dependent	*LIG4*	3	0.031	555.333
Peptidase C12, ubiquitin carboxyl-terminal hydrolase	*UCHL3*	4	0.031	416.500
DNA ligase, ATP-dependent, N-terminal	*LIG4*	3	0.031	555.333
DNA ligase, ATP-dependent, C-terminal	*LIG4*	3	0.031	555.333
DNA ligase, ATP-dependent, central	*LIG4*	3	0.031	555.333
DNA ligase, ATP-dependent, conserved site	*LIG4*	3	0.031	555.333
Holliday junction regulator protein family C-terminal	*HJURP*	4	0.031	416.500
The domain of unknown function DUF4757	*LMO7*	2	0.031	833.000
DNA ligase, ATP-dependent, N-terminal domain superfamily	*LIG4*	3	0.031	555.333
Peptidase C12, ubiquitin carboxyl-terminal hydrolase superfamily	*UCHL3*	4	0.031	416.500
Neuronal tyrosine-phosphorylated phosphoinositide-3-kinase adapter, N-terminal	*MYO16*	3	0.031	555.333
ATP-dependent DNA ligase domain	*LIG4*	3	0.031	555.333
Ubiquitin carboxyl-terminal hydrolase, family 1	*UCHL3*	4	0.031	416.500
DNA ligase N terminus	*LIG4*	3	0.031	555.333
ATP dependent DNA ligase C terminal region	*LIG4*	3	0.031	555.333
Holliday junction regulator protein family C-terminal repeat	*HJURP*	4	0.031	416.500
Neuronal tyrosine-phosphorylated phosphoinositide-3-kinase adapter	*MYO16*	3	0.031	555.333
The domain of the unknown function (DUF4757)	*LMO7*	2	0.031	833.000
Mixed, incl. baff-r, tall-1 binding, and taci, cysteine-rich domain	*LOC107131142*	5	0.032	333.200

Mixed, incl. armadillo-type fold	*MROH2A*	5	0.032	333.200
Non-homologous end joining	*LIG4*	5	0.032	333.200
Mixed, incl. attractin-like, c-type lectin-like domain, and granular component	*DNAJB3*	5	0.032	333.200
Mixed, incl. p2x3 purinoceptor, and transient receptor potential cation channel subfamily v member 1	*TRPM8*	6	0.035	277.667
TRPM, SLOG domain	*TRPM8*	6	0.035	277.667
Mixed, incl. serine aminopeptidase, s33, and glucuronoside catabolic process	*ABHD13*	7	0.036	238.000
Serine aminopeptidase, S33	*ABHD13*	7	0.036	238.000
Secreted phosphoprotein 24 (Spp-24) cystatin-like domain	*SPP2*	7	0.037	238.000
BRCT domain, a BRCA1 C-terminus domain	*LIG4*	8	0.040	208.250
Smooth muscle protein/calponin	*LMO7*	9	0.044	185.111
Mixed, incl. epithelial sodium channel, conserved site, and piezo family	*TRPM8*	11	0.047	151.454

Table 5.7: Enriched GO terms and pathways terms for weaning weight using ShinyGo

Pathway	**Genes**	**Pathway Genes**	**Enrichment FDR**	**FE**
Ubiquinol-cytochrome-c reductase activity	*UQCRFS1*	5	0.043	444.267
Oxidoreductase activity, acting on diphenols and related substances as donors	*UQCRFS1*	5	0.043	444.267
Oxidoreductase activity, acting on diphenols and related substances as donors, cytochrome as acceptor	*UQCRFS1*	5	0.043	444.267
Oxygen carrier activity	*MB*	9	0.046	246.815
Oxygen binding	*MB*	9	0.046	246.815

Table 5.8: Enriched GO terms for yearling weight using ShinyGo

Pathway	**Genes**	**Pathway Genes**	**Enrichment FDR**	**FE**
Caveola	*CAVIN4*	27	0.034	246.815
Myofibril	*CAVIN4*	66	0.034	100.970
Sarcomere	*CAVIN4*	61	0.034	109.246
Z disc	*CAVIN4*	40	0.034	166.600
I band	*CAVIN4*	42	0.034	158.667
Sarcolemma	*CAVIN4*	39	0.034	170.872
Contractile fiber	*CAVIN4*	67	0.034	99.463
Plasma membrane raft	*CAVIN4*	33	0.034	201.939
Membrane raft	*CAVIN4*	87	0.034	76.598
Membrane region	*CAVIN4*	92	0.034	72.435

Membrane microdomain	*CAVIN4*	87	0.034	76.598

Table 5.9: Weaning weight-related genes grouped by functional categories defined by high-level GO terms.

N	High-level GO category	Genes
4	Binding	*MB MCM5 RBFOX2 UQCRFS1*
4	Cellular process	*MB MCM5 RBFOX2 UQCRFS1*
3	Metabolic process	*MCM5 RBFOX2 UQCRFS1*
3	Organelle	*MCM5 RBFOX2 UQCRFS1*
3	Response to stimulus	*MB MCM5 RBFOX2*
3	Intracellular	*MCM5 RBFOX2 UQCRFS1*
3	Ion binding	*MB MCM5 UQCRFS1*
3	Membrane-bounded organelle	*MCM5 RBFOX2 UQCRFS1*
3	Intracellular organelle	*MCM5 RBFOX2 UQCRFS1*
3	Cellular metabolic process	*MCM5 RBFOX2 UQCRFS1*
3	Organic cyclic compound binding	*MB MCM5 RBFOX2*
3	Heterocyclic compound binding	*MB MCM5 RBFOX2*
2	Catalytic activity	*MCM5 UQCRFS1*
2	Multicellular organismal process	*MB RBFOX2*
2	Developmental process	*MB RBFOX2*
2	Protein-containing complex	*MCM5 UQCRFS1*
2	Biological regulation	*MB RBFOX2*
2	Nitrogen compound metabolic process	*MCM5 RBFOX2*
2	Response to stress	*MB MCM5*
2	Primary metabolic process	*MCM5 RBFOX2*
2	Anatomical structure development	*MB RBFOX2*
2	Cellular response to stimulus	*MCM5 RBFOX2*
2	Organic substance metabolic process	*MCM5 RBFOX2*
1	Immune system process	*MB*
1	Transporter activity	*MB*
1	Signaling	*RBFOX2*
1	Regulation of biological process	*RBFOX2*
1	Localization	*MB*

1	Cellular component organization or biogenesis	*MCM5*
1	Molecular function regulator	*RBFOX2*
1	Myeloid cell homeostasis	*MB*
1	Immune system development	*MB*

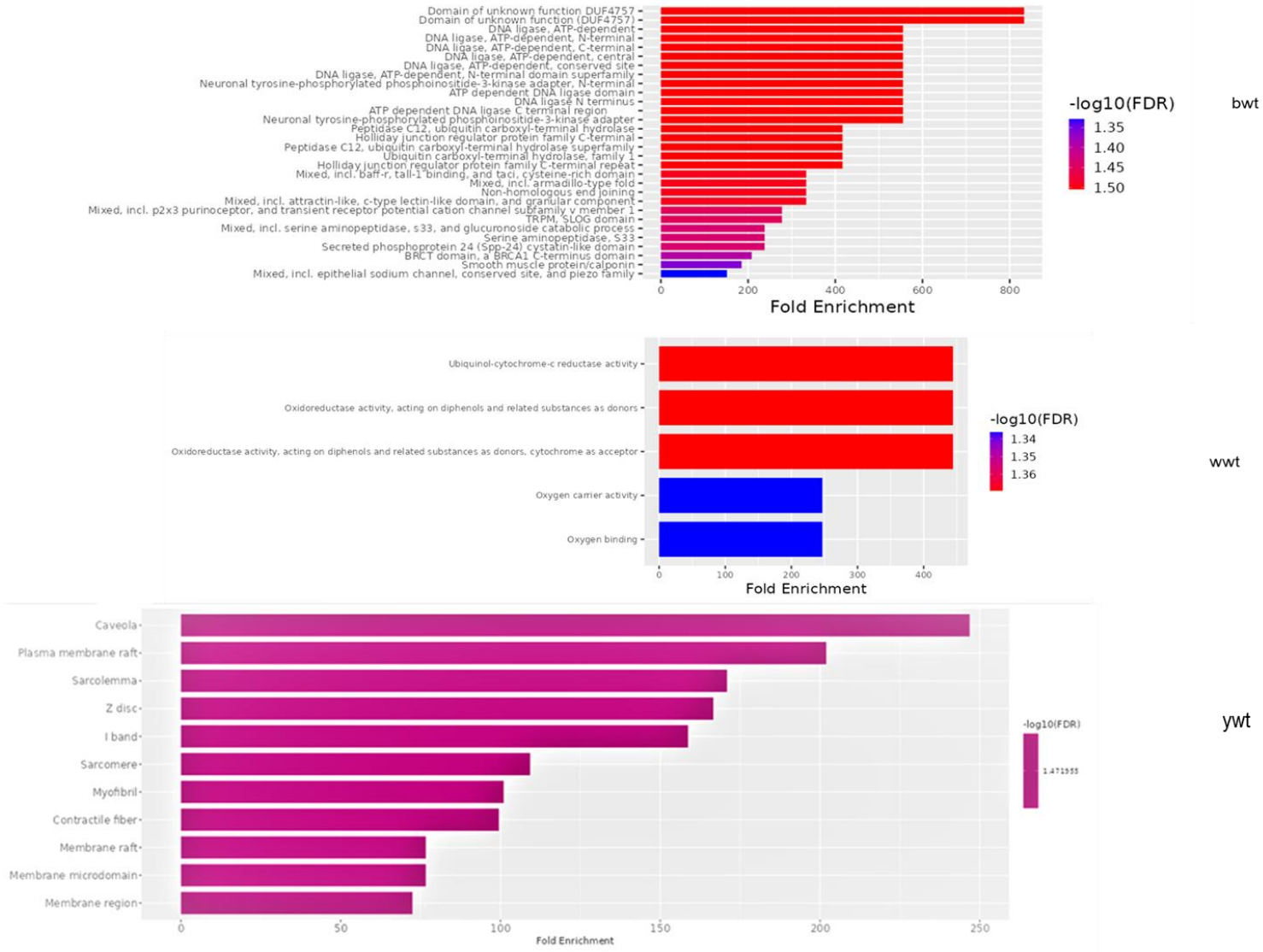

Figure 5.4: Enriched GO terms and pathways for birth weight (bwt), weaning (wwt), and yearling weight (ywt) generated by ShinyGo

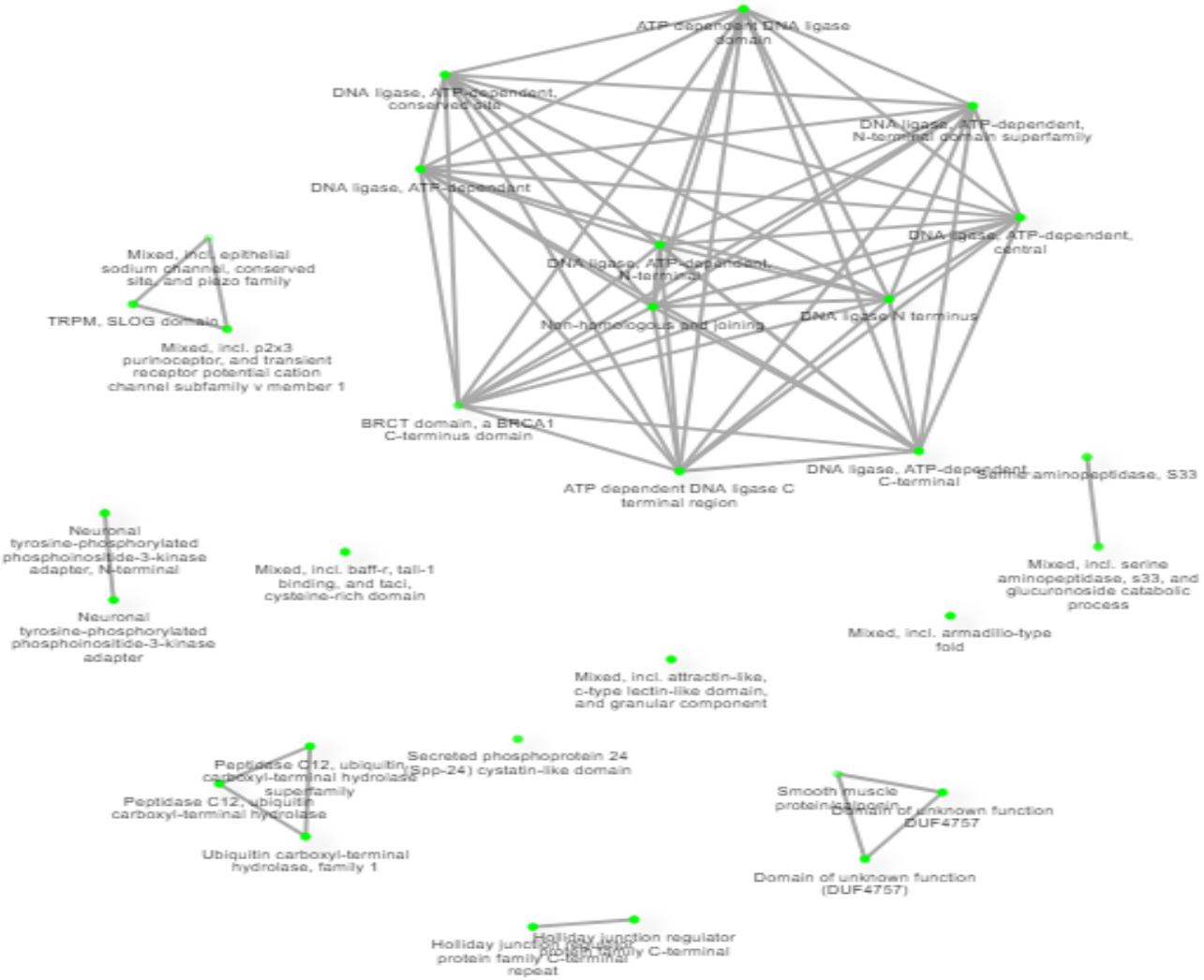

Figure 5.5: Networks for birth weight enriched pathways using ShinyGo bioinformatics platform

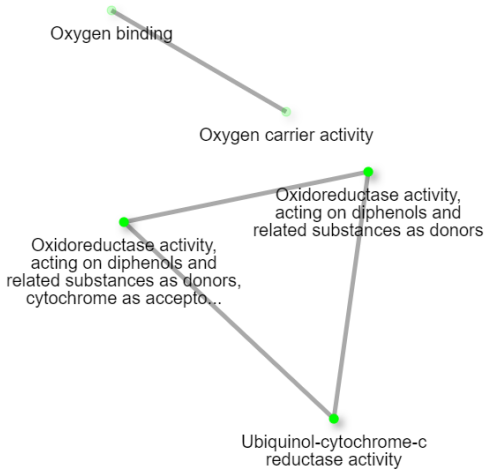

Figure 5.6: Network for weaning weight enriched pathways using ShinyGo bioinformatics platform

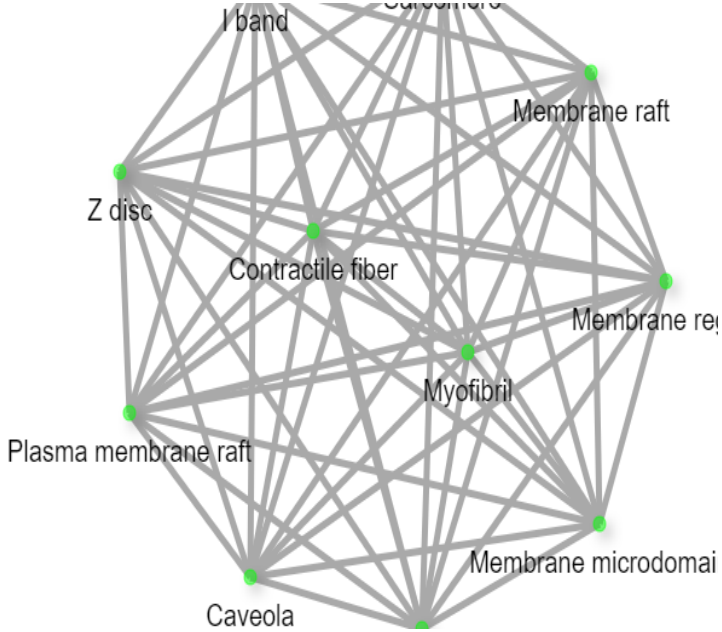

Figure 5.7: Network for yearling weight enriched pathways using ShinyGo bioinformatics platform

5.3.3 Skin thickness

The mapped genes for skin thickness are presented in Annexure 2. Furthermore, GO terms enriched through the DAVID bioinformatics resource are presented in Table 5.10. Clustering through DAVID was not a possibility for this trait because the related genes were very few. Furthermore, the enrichment analyses conducted through ShinyGo revealed enriched GO terms, that are described in Table 5.11 and further presented graphically in Figure 5.8. The interaction for the enriched GO terms is exhibited in Figure 5.9.

Table 5.10: Gene ontology terms and pathways significantly enriched using DAVID for skin thickness

Category	Term	Genes	%	PV	FE
INTERPRO	IPR020635:Tyrosine-protein kinase, catalytic domain	*FER, NTRK2*	66.67	0.01	188.79
INTERPRO	IPR008266:Tyrosine-protein kinase, active site	*FER, NTRK2*	66.67	0.01	160.06
GOTERM_MF_DIRECT	GO:0004714~transmembrane receptor protein tyrosine kinase activity	*FER, NTRK2*	66.67	0.01	137.12
GOTERM_BP_DIRECT	GO:0007169~transmembrane receptor protein tyrosine kinase signaling pathway	*FER, NTRK2*	66.67	0.01	115.69
INTERPRO	IPR001245:Serine-threonine/tyrosine-protein kinase catalytic domain	*FER, NTRK2*	66.67	0.01	109.08
SMART	SM00219:TyrKc	*FER, NTRK2*	66.67	0.02	87.9
INTERPRO	IPR017441:Protein kinase, ATP binding site	*FER, NTRK2*	66.67	0.03	38.75
UP_SEQ_FEATURE	ACT_SITE:Proton acceptor	*FER, NTRK2*	66.67	0.04	35.66
GOTERM_BP_DIRECT	GO:0045087~innate immune response	*FER, NTRK2*	66.67	0.04	32.1
INTERPRO	IPR000719:Protein kinase, catalytic domain	*FER, NTRK2*	66.67	0.04	29.75
GOTERM_BP_DIRECT	GO:0030154~cell differentiation	*FER, NTRK2*	66.67	0.05	29.07
INTERPRO	IPR011009:Protein kinase-like domain	*FER, NTRK2*	66.67	0.05	27.32
UP_KW_MOLECULAR_FUNCTION	KW-0418~Kinase	*FER, NTRK2*	66.67	0.05	18.96

PV- pvalue, FE- Fold enrichment

Table 5.11: Enriched GO terms and pathways for skin thickness using ShinyGo

Pathway	Genes	Pathway Genes	Enrichment FDR	FE
Tyrosine-protein kinase, active site	*FER, NTRK2*	88	8.02E-04	151.455
Tyrosine-protein kinase, catalytic domain	*FER, NTRK2*	76	8.02E-04	175.368
Serine-threonine/tyrosine-protein kinase, catalytic domain	*FER, NTRK2*	126	1.10E-03	105.778
Tyrosine-protein kinase, Fes/Fps type	*FER*	2	1.68E-03	3332.000
Fes/Fps/Fer, SH2 domain	*FER*	2	1.68E-03	3332.000
Tyrosine-protein kinase, neurotrophic receptor	*NTRK2*	3	1.80E-03	2221.333
Tyrosine-protein kinase receptor NTRK, C2-Ig-like domain	*NTRK2*	3	1.80E-03	2221.333
Protein kinase, ATP binding site	*FER, NTRK2*	354	3.25E-03	37.650
Tyrosine-protein kinase, receptor class II, conserved site	*NTRK2*	9	4.20E-03	740.444
Protein kinase domain	*FER, NTRK2*	454	4.26E-03	29.357
Protein kinase-like domain superfamily	*FER NTRK2*	493	4.56E-03	27.034
FCH domain	*FER*	19	6.65E-03	350.737
F-BAR domain	*FER*	21	6.78E-03	317.333
Leucine-rich repeat N-terminal domain	*NTRK2*	51	0.015	130.667
AH/BAR domain superfamily	*FER*	69	0.019	96.580
The cysteine-rich flanking region, C-terminal	*NTRK2*	74	0.019	90.054
SH2 domain	*FER*	104	0.025	64.077
SH2 domain superfamily	*FER*	106	0.025	62.868
Immunoglobulin I-set	*NTRK2*	126	0.028	52.889
Immunoglobulin subtype 2	*NTRK2*	196	0.041	34.000
Leucine-rich repeat	*NTRK2*	231	0.046	28.848
Zinc finger, RING-type	*PJA2*	258	0.049	25.829

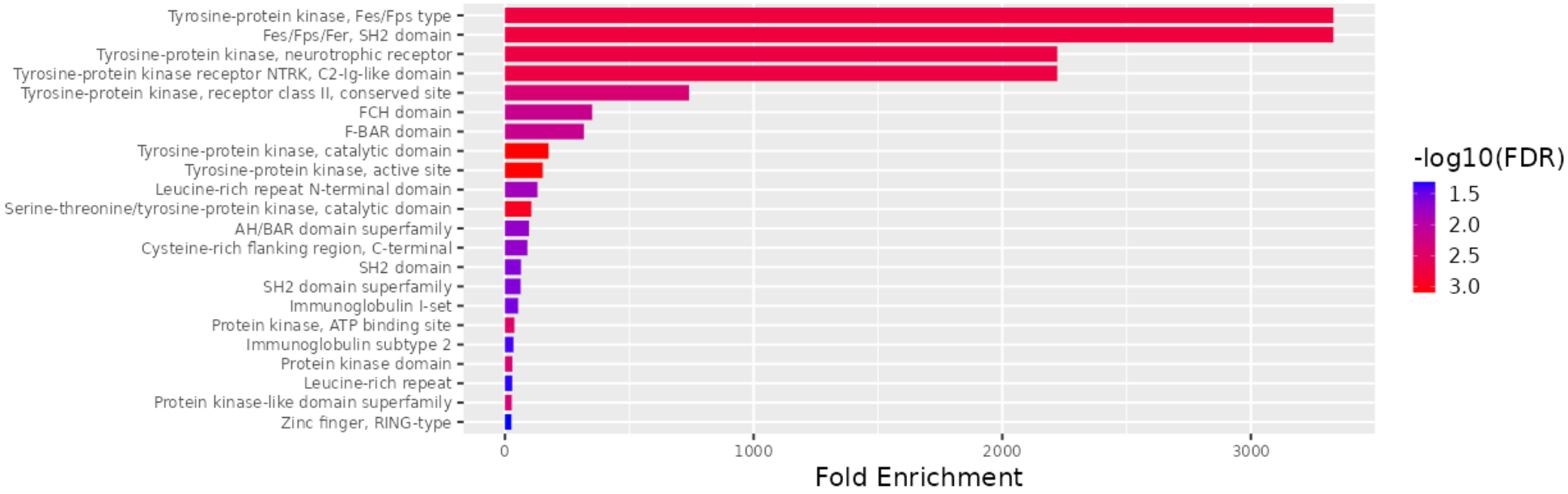

Figure 5.8: Enriched pathways for the skin thickness

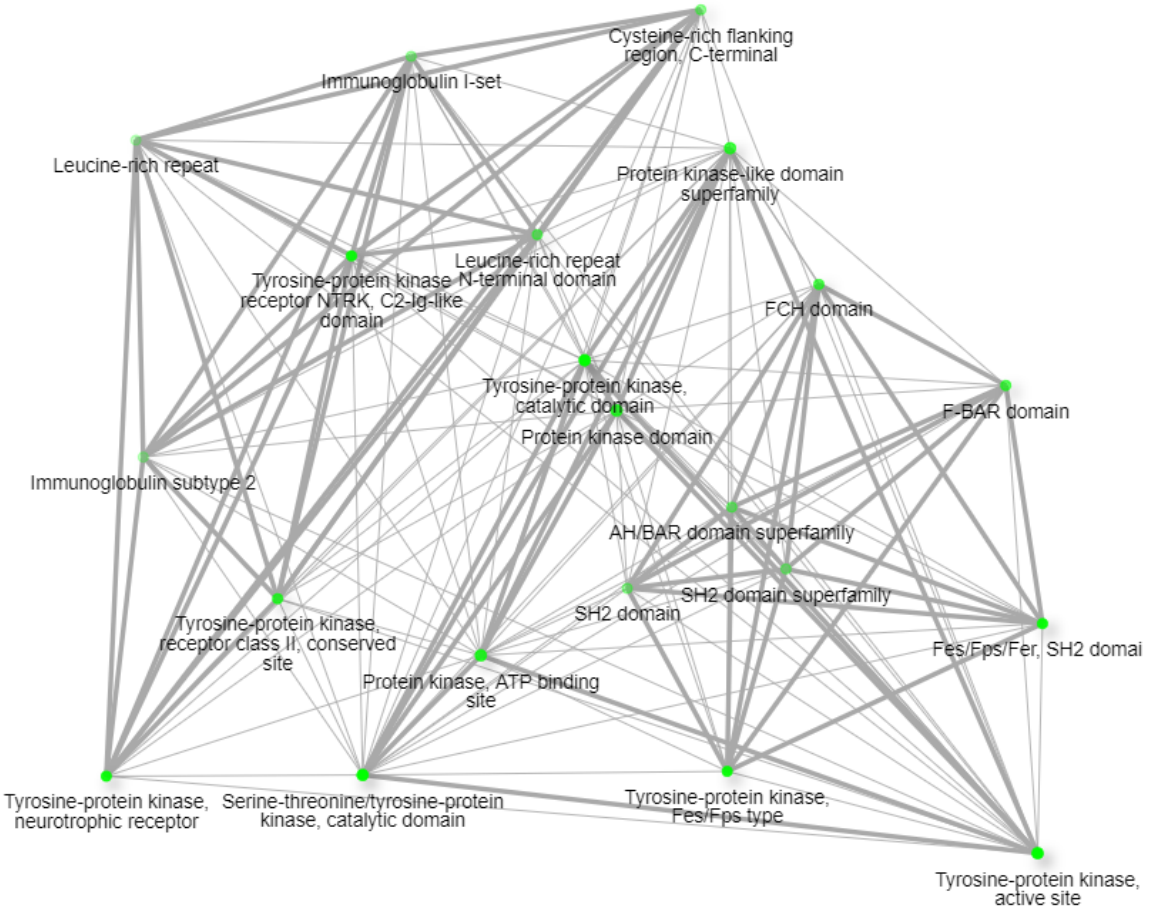

Figure 5.9: Networks for skin thickness enriched pathways using ShinyGo bioinformatics platform

5.4 Discussion

5.4.1 Tick count

The presence of uncharacterized genes could be attributed to the fact that the whole bovine genome assembly is still incomplete. The abundantly identified genes were from the zinc finger protein family (*ZNF398, ZNF789, ZNF212, ZNF777, ZNF746, ZNF467, ZNF775,* and *ZNF862*). The region at SNP BovineHD2200005710 on BTA22 was characterized by two genes*, U6* and *GRM7*. These genes have not been associated with tick count in cattle previously. However, *Metabotropic Glutamate Receptor 7*(*GRM7)* has been suspected to play a role in parasite resistance in cattle [26,27]. The enrichment findings excluded the two genes related to BTA22.

Zinc finger proteins found on BTA4 are the largest protein families in eukaryotes, which control a variety of cellular, molecular, and biological processes [28–30]. No previous study has reported the association of zinc finger genes with tick count in cattle. Based on functional annotation analysis conducted in the current study, zinc finger genes were involved in three significantly enriched GO terms. The three significantly enriched GO terms were related to biological process (GO:0006357~regulation of transcription from RNA polymerase II promoters), and molecular functions (GO:0003700~transcription factor activity, sequence-specific DNA binding; GO:0000978~RNA polymerase II core promoter proximal region sequence-specific DNA binding). RNA polymerase II promoter is regulated in response to internal and external stimuli [31]. Additionally, these zinc finger genes were significantly enriched in a variety of pathways using DAVID and ShinyGo.

The significantly enriched pathways also included IPR001909:Krueppel-associated box, SM00349:KRAB, and IPR003655:Krueppel-associated box-related. These genes are involved in the modulation of adaptive immunity and are responsible for responding to stimuli in mice and human [32]. In cattle, the zinc finger genes involved in Krueppel-associated box (KRAB) have been associated with the development of embryos [33], and the synthesis of milk in dairy cattle [34]. Additionally, a GWA study by Wen et a. [35] found that KRAB genes (*ZNF862* and *ZNF775)* had an association with supernumerary teats in Holstein cattle. In sheep, there was an association of KRAB genes with host resilience to small ruminant Lentivirus in two populations [36]. In chickens, the zinc finger gene has been considered as a potential gene for host resistance to Avian influenza, due to its role in the immune response during viral infection [37]. In plants, zinc finger domains have been associated with disease resistance [38], and were reported to regulate resistance mechanisms for various biotic and abiotic

stresses [39][40]. Therefore, the effect of zinc finger proteins on the expression of tick count in cattle requires further investigation.

In the present study, the two clusters were generated by the functional annotation clustering analysis. These were mostly comprised of GTPase, IMAP family member (*GIMAP7* and *GIMAP8)*, and zinc finger domain. In cluster 1, the GTPase, IMAP family member genes were positively associated with most of the terms that were enriched. The least was a zinc finger protein (*ZNF746*), which was positively associated with two terms related to GO:0005829~cytosol and KW-0175~coiled coil. On the other hand, cluster 2 exhibited the *ZNF746* gene as the most positively associated with enriched terms, followed by genes *ZNFF 212, ZNF 398,* and *ZNF777.*The overall enrichment analysis for tick count implicated *the ZNF746* gene. This gene is a protein-coding gene located in the cytosol and nucleus. It plays a pivotal role in the positive regulation of transcription by RNA polymerase II [41]. The ShinyGo analysis related the genes to functions such as regulation of biological processes, the regulation of cellular processes, and protein binding. These functions are part of the important mechanisms for the host's response to tick infestation. Therefore, the Zinc Finger Protein 746 (*ZNF746)* gene may be considered as a candidate for tick count in cattle.

Another potential candidate gene underlying tick count expression in the current study is GTPase, IMAP Family Member 8 (*GIMAP8).* This gene has been previously reported by Santo et al. [1], as a candidate gene for tick resistance in cattle. It plays a role in the modulation of the immune system during the response to infections [42]. Berg et al. [43] postulated that *GIMAP8* is part of the GTPase IMAP family members, which are intensively expressed in the last stage of B and T cell development. The role of B cells in conferring tick resistance in cattle has been demonstrated by Robbertse et al. [44]. They observed that in tick resistant breeds, B cells are important mediators of immune response due to their influx and proliferation in CD3+ T cells at tick fixation areas. Another candidate gene for tick count expression that was found in the current study is *Retinoic Acid Receptor Responder 2* (*RARRES2).* This is a protein-coding gene that encodes a secreted chemotactic protein, which initiates chemotaxis response to chemical stimuli [45]. Furthermore, Bondue et al. [46] confirmed that *RARRES2* is a chemotactic factor for leukocytes during a response to stimuli. According to Geering et al [47], granulocyte chemotaxis forms part of the essential elements of innate immune response to a stimulus. Chemotaxis is the movement of cells in response to stimulus, and in this context of tick infestation events, important cells are moved to the area of irritation or tick fixation [48–50]. Furthermore, a review by Tabor et al. [51] highlighted distinctive differences observed amongst the skin of tick-resistant and tick-susceptible cattle. The skin of resistant *Bos indicus*

cattle was intensively invaded by granulocytes, such as, eosinophils and basophils at the point of infestation, which was not the case in their susceptible counterparts. Therefore, the role of *RARRES2* in controlling the level of tick count in cattle needs to be studied further.

The region associated with tick count in BTA4 can be considered a hot spot for the trait's expression in F_2 Angus x Nguni cattle. To gain more insight into the underlying biological architecture and mechanisms of this region related to tick resistance, more intensive studies should be conducted. The genes *ZNF746*, *GIMAP8, and RARRES2* are potential biomarkers that can be used in the marker-assisted selection for tick count.

5.4.2 Growth traits

The analysis of birth weight showed genes that were located to the BTA 3, and these included *TRPM8, SPP2, HJRP, MROH2A, DNAJB3, ENSBTAG0000005483, ENSBTAG00000051335, ENSBTAG0000004331, and ENSBTAG00000053153* (Annexure 2). The *transient receptor potential cation channel subfamily M member 8* (*TRPM8*) is a protein-coding gene involved in calcium ion transmembrane transport and positive regulation of cold-induced thermogenesis [52]. *Maestro heat like repeat family member 2A* (*MROH2A)* is a protein-coding protein, and the function of the encoded protein has been characterized [53]. DnaJ Heat Shock Protein Family (Hsp40) Member B3 (*DNAJB3*) is a pseudogene [24], that has been associated with metabolic dysfunction in mice [54]. Four genes associated with BTA 3 for birth weight were uncharacterized; therefore, it was not possible to attain biological relevance of these genes on birth weight.

The proteins encoded by *TRPM8* are involved in brown adipocytes facilitating the regulation of body temperature, and the survival of new-borns in mammals [55]. An important paralog for *TRPM8*, *TRPM2* is known to play a role in the innate immune system and ion channel transport [56]. *TRPM8* facilitates calcium ion transmembrane transport, which allows calcium ions to regulate biological function involved in embryo formation and development in mammals [57,58]. In mice, this gene influences energy balance and fine-tunes eating behaviour [59]. The deficiency of this gene results in metabolic dysfunctions [59]. In humans, the potential use of *TRPM8* agonists includes the reduction of glucose levels, weight gain, and treatment of metabolic syndromes, obesity, and type 2 diabetes [60]. This gene has been implicated in cattle as a candidate gene associated with environmental adaptation and metabolic homeostasis [61].

The *TRPM8* gene is also linked with *SPP2* (*Secreted Phosphoprotein 2*), a bone morphogen [62]. The role of bone morphogen proteins includes the influence on skeletal development and growth during early life in mammals [63]. A study by Dzhalginsky et al. [64] investigated the correlation between the expression of *SPP2* and live weight in sheep. Despite the significantly low expression of *SPP2*, they concluded that the correlation between the expressions of the gene with live weight exists. They considered this gene as a potential genetic marker to control live weight in sheep. Moreover, *TRPM8* and *SPP2* were identified as part of the genes responsible for reproduction and growth in sheep [65]. Studies on the relationship for *TRPM8* and *SPP2* genes with live weights in cattle are limited.

The genes that were grouped under the significantly enriched GO term (GO:0005654~nucleoplasm) related to cellular components, and these are *LMO7*, *HJRP*, *LIG4*, and *UCHL3* . The role of nucleoplasm is to provide a conducive environment for cell activities that occur inside the nucleus [66]. Growth and metabolism are part of the biological mechanisms which are controlled and regulated by the nucleus [67]. Furthermore, these genes (*LMO7*, *HJRP*, *LIG4*, *UCHL3,* and *MYO16*) were significantly enriched for pathways generated by the ShinyGo bioinformatics resource. The enriched pathways were related to DNA ligase, and ATP-dependent. The network analysis showed the connections among some of the enriched pathways.

Three chromosomal regions were uncovered for weaning weight and were associated with 11 genes (**BTA 5**: *TOMI*, *MCM5*, *IXS*, *RASD2, RBF0X2, ENSBTAG0000004906,* and *MB;* **BTA18**: *UQCRFS*, *VSTM2B*, *OR4P4*, *6*, and *HMGXB4*). For BTA18 two SNPs were suggestively associated with weaning weight. No GO terms were significantly enriched for weaning weight using the DAVID. However, ShinyGo successfully generated enriched pathways relating to the molecular function category for this trait. The enriched pathways included ubiquinol-cytochrome-c reduction activity, oxygen carrier activity, and oxygen binding, among other. The enriched pathways involved *MB* and *UQCRFS1*, which suggests that a connection among the enriched pathways exists. Myoglobin (*MB*) is a protein-coding gene that is expressed in skeletal and cardiac muscles [68]. In addition, myoglobin plays a role in muscle oxidation metabolism [69,70], and affects the colour of meat in pigs, sheep, and cows [71–73]. Ubiquinol-Cytochrome C Reductase, Rieske Iron-Sulfur Polypeptide 1 (*UQCRFS1)* is a protein coding gene involved in mitochondrial respiratory chain complex III assembly and respiratory electron transport chain [74]. It is associated with feed efficiency and energy balance in dairy cattle [75]. The gene set enrichment performed using ShinyGo showed that most genes related to weaning weight belonged to binding, cellular process, and

metabolic process. These functions are related to growth and development, therefore this information provides insight into biological mechanisms underlying weaning in cattle.

The chromosomal region BTA8 was uncovered for yearling weight and was associated with 7 genes (*MSANTD3*, *TMEFF1*, *CAVIN4*, *CDK20*, *ENSBTAG00000021235*, *ENSBTAG00000025756*, and *ENSBTAG00000054632*). Out of these genes, three were uncharacterized. One biological processed GO term (GO:0030154~cell differentiation) was enriched for yearling weight using the DAVID bioinformatics resource. Cell differentiation is a process in which cells become specialized [76], informed by a changed structural morphology and functional characteristics [77]. The role of cell differentiation is irreversible. This facilitates selective expression of the genome, whereby the expression of different genes is turned on or off, to produce proteins [77]. Du et al. [78] noted manipulation of progenitor cell differentiation through nutrition management, selective breeding or fetal programming as promising tool that can be used to improve cattle performance and carcass value. The inducers of cell differentiation in relation to yearling weight in F_2 Angus x Nguni cattle were *TMEFF1* and *CAVIN4.*

Transmembrane Protein with EGF-Like and One Follistatin-Like Domain (*TMEFF1)* is a protein-coding gene predicted to be involved in animal organ morphogenesis, neuron projection development, and tissue development [79]. Furthermore, previous evidence shows that *TMEFF1* and *CAVIN4* play a significant role in the biological mechanisms regulating energy metabolism and skeletal muscle development in cattle [80,81]. Caveolae-associated protein 4 (*CAVIN4)* is a protein-coding gene, which promotes signaling in cardiac muscle cells, and may facilitate myofibrillar organization [82]. This gene has been identified as a novel muscle disease candidate caveolar protein in humans [83]. The gene enrichment analysis using the ShinyGO revealed the association of *CAVIN4* with significantly enriched pathways for the current study. Part of the enriched pathways included Plasma membrane raft, myofibril, caveola, and contractile fiber, where the interaction among the pathways exists. Additionally, Myb/SANT-Like DNA-Binding Domain-Containing Protein 4 (*MSANTD3*) and Cyclin-Dependent Kinase 20 (*CDK20*) have been regarded as candidate genes related to cell growth and metabolism in chicken [84] and cattle [85], respectively.

Overall, most uncovered candidate genes in the current study showed biological relevance in skeletal muscle development and energy metabolism contributing to the expression of growth traits in cattle. This holds a promise that these genes can be used in selection schemes to improve growth performance in cattle. However, it should be noted that the biological

relevance of some of the uncovered genes for growth traits was unknown. Therefore, there is a need to study these genes further.

5.4.3 Skin thickness

The skin is a large, complex organ with a wide range of functions, which include sensation, protection, and thermoregulation. Although the skin is thin at birth, its thickness in dairy and beef cattle varies due to genetic effects, age, sex, breed, nutritional status, and environmental conditions. Skin thickness is a particularly important trait in cattle as it is associated with immune response [86], heat tolerance [87], tick resistance [88], body score condition, milk production [89], and reproductive traits [90].

In this study, five genes were mapped on BTA7 and BTA8 for skin thickness. The mapped genes included *NTRK2*, *bta-m-2465*, *ENSBTAG00000051829*, *FER*, and *PJA2* (Annexure 2). The relation of these genes with skin thickness has not been reported in previous studies. This may be attributed to the fact that there are extremely few investigations that have studied the genetic components of skin thickness in cattle. Neurotrophic receptor tyrosine kinase 2 (*NTRK2*) is a protein-coding membrane-bound receptor that phosphorylates itself upon neutrophil binding [24]. The signalling involved in this gene leads to cell differentiation. In humans, this gene is associated with body weight [91], and a mutation in this gene results in obesity [92]. Furthermore, it has been implicated as a molecular signature for human epidermal melanocytes [93]. A study by Schleger & Bean [94] on the melanocyte system of cattle skin, postulated that the melanocyte system might have an important surveillance role in the defense mechanisms of the skin. In sheep, it has been confirmed as a novel candidate gene for litter size [95]. This kinase harbours the deferential methylation region associated with fibroblasts in Holstein and Angus cattle [96]. Fibroblasts play a key role in regulating skin physiology and cutaneous wound repair [97]. The role of fibroblast has been observed in the process of healing the full-skin thickness wound in rats [98].

FER tyrosine kinase (*FER*) is a protein-coding gene, which plays a role in the regulation of the actin cytoskeleton, cell migration, and chemotaxis [24]. The actin cytoskeleton has an important cellular function in mammals, and it is implicated as a biological factor influencing the biology of the skin [99–101]. Additionally, the *FER* gene is known for its contribution to regulating innate and adaptive immunity [100,102]. In sheep, this gene has been associated with *Haemonchus contortus* resistance in the Marada Nova breed [103]. It was also noted that the *FER* gene is involved in leukocyte recruitment, responsible for the response to bacteria

[104]. Mapholi et al. [15] noted that this gene was associated with tick resistance in the Nguni breed. The existing relationship between skin thickness and tick resistance in cattle has been reported by Foster et al. [105] and Fourie et al. [106]. The relationship between skin thickness and tick count was not established in this study. However, there is a need to further investigate the biological influence *FER* gene and its contribution to the control of the two traits.

Functional annotation uncovered enriched GO biological processes terms related to innate immune response (GO:00450087), cell differentiation (GO:00330154), and transmembrane receptor protein (GO:0007169). One molecular function GO related to a transmembrane receptor protein (GO:0004714) was also enriched, together with various pathways which include tyrosine-protein kinase, catalytic (IPR020365), tyrosine-protein kinase, and active site (IPR008266). The significantly enriched categories for this trait only involved *NTRK2 and FER* genes in both bioinformatics resources used for functional annotation. Therefore, *NTRK2 and FER* are noted as important genes influencing the expression of skin thickness in F_2 Angus x Nguni cattle. However, the fine mapping of these genes is required to validate their biological contribution to skin thickness.

5.5 Conclusions

The current study uncovered novel and previous genes related to tick count, growth traits, and skin thickness in F_2 Angus x Nguni cattle. Biological mechanisms influencing the traits of interest were also uncovered. For tick count, the region of BTA4 was associated with tick count; thus, it could be responsible for the expression of this trait in the F_2 Angus x Nguni cattle. Improvement of tick count could be achieved through marker-assisted selection, with the genes located on BTA4 as potential biomarkers. The genes uncovered for growth traits were associated with skeletal muscle development and energy metabolism. Growth performance can thus be improved if these genes are incorporated into selection schemes for growth performance in this population. The genes uncovered on the skin were *NTRK2 and FER* that were associated with immunity in the F_2 Angus x Nguni cattle. This study contributed to the understanding of the biological mechanisms underlying tick count, growth traits, and skin thickness in cattle. Furthermore, future studies are recommended to validate the identified regions which could be used to facilitate the improvement of the traits of interest through selective breeding.

5.6 References

[1] dos Santos CG, Sousa MF, Vieira JIG, de Morais LR, Fernandes AAS, de Oliveira Littiere T, et al. Candidate genes for tick resistance in cattle: a systematic review combining post-GWAS analyses with sequencing data. J Appl Anim Res 2022;50:460–70. https://doi.org/10.1080/09712119.2022.2096035.

[2] Dadousis C, Pegolo S, Rosa GJM, Gianola D, Bittante G, Cecchinato A. Pathway-based genome-wide association analysis of milk coagulation properties, curd firmness, cheese yield, and curd nutrient recovery in dairy cattle. J Dairy Sci 2017;100:1223–31. https://doi.org/10.3168/jds.2016-11587.

[3] Srikanth K, Lee S-H, Chung K-Y, Park J-E, Jang G-W, Park M-R, et al. A Gene-Set Enrichment and Protein-Protein Interaction Network-Based GWAS with Regulatory SNPs Identifies Candidate Genes and Pathways Associated with Carcass Traits in Hanwoo Cattle. Genes 2020;11:E316. https://doi.org/10.3390/genes11030316.

[4] Kopke G, Anklam K, Kulow M, Baker L, Swalve HH, Lopes FB, et al. The identification of gene ontologies and candidate genes for digital dermatitis in beef cattle from a genome-wide association study. Int J Vet Sci Res 2020;6:027–37. https://doi.org/10.17352/ijvsr.000050.

[5] Clancey E, Kiser JN, Moraes JGN, Dalton JC, Spencer TE, Neibergs HL. Genome-wide association analysis and gene set enrichment analysis with SNP data identify genes associated with 305-day milk yield in Holstein dairy cows. Anim Genet 2019;50:254–8. https://doi.org/10.1111/age.12792.

[6] Hirschhorn JN. Genomewide association studies--illuminating biologic pathways. N Engl J Med 2009;360:1699–701. https://doi.org/10.1056/NEJMp0808934.

[7] Otto PI, Guimarães SEF, Calus MPL, Vandenplas J, Machado MA, Panetto JCC, et al. Single-step genome-wide association studies (GWAS) and post-GWAS analyses to identify genomic regions and candidate genes for milk yield in Brazilian Girolando cattle. J Dairy Sci 2020;103:10347–60. https://doi.org/10.3168/jds.2019-17890.

[8] Marina H, Pelayo R, Suárez-Vega A, Gutiérrez-Gil B, Esteban-Blanco C, Arranz JJ. Genome-wide association studies (GWAS) and post-GWAS analyses for technological traits in Assaf and Churra dairy breeds. J Dairy Sci 2021;104:11850–66. https://doi.org/10.3168/jds.2021-20510.

[9] Adam Y, Samtal C, Brandenburg J, Falola O, Adebiyi E. Performing post-genome-wide association study analysis: overview, challenges and recommendations. F1000Research 2021;10:1002. https://doi.org/10.12688/f1000research.53962.1.

[10] Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44–57. https://doi.org/10.1038/nprot.2008.211.

[11] Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009;37:1–13. https://doi.org/10.1093/nar/gkn923.

[12] The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists | Genome Biology | Full Text n.d. https://genomebiology.biomedcentral.com/articles/10.1186/gb-2007-8-9-r183 (accessed November 13, 2022).

[13] Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics 2020;36:2628–9. https://doi.org/10.1093/bioinformatics/btz931.

[14] Otto PI, Guimarães SEF, Verardo LL, Azevedo ALS, Vandenplas J, Soares ACC, et al. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J Dairy Sci 2018;101:11020–32. https://doi.org/10.3168/jds.2017-14223.

[15] Mapholi NO, Maiwashe A, Matika O, Riggio V, Bishop SC, MacNeil MD, et al. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks Tick-Borne Dis 2016;7:487–97. https://doi.org/10.1016/j.ttbdis.2016.02.005.

[16] Buzanskas ME, Grossi DA, Ventura RV, Schenkel FS, Sargolzaei M, Meirelles SLC, et al. Genome-Wide Association for Growth Traits in Canchim Beef Cattle. PLoS ONE 2014;9:e94802. https://doi.org/10.1371/journal.pone.0094802.

[17] Du L, Duan X, An B, Chang T, Liang M, Xu L, et al. Genome-Wide Association Study Based on Random Regression Model Reveals Candidate Genes Associated with Longitudinal Data in Chinese Simmental Beef Cattle. Anim Open Access J MDPI 2021;11:2524. https://doi.org/10.3390/ani11092524.

[18] Smith JL, Wilson ML, Nilson SM, Rowan TN, Schnabel RD, Decker JE, et al. Genome-wide association and genotype by environment interactions for growth traits in U.S. Red Angus cattle. BMC Genomics 2022;23:517. https://doi.org/10.1186/s12864-022-08667-6.

[19] Ensembl genome browser 108 n.d. https://www.ensembl.org/index.html (accessed November 21, 2022).

[20] ARS-UCD1.2 - bosTau9 - Genome - Assembly - NCBI n.d. https://www.ncbi.nlm.nih.gov/assembly/GCF_002263795.1/ (accessed November 21, 2022).

[21] DAVID Functional Annotation Bioinformatics Microarray Analysis n.d. https://david.ncifcrf.gov/ (accessed November 21, 2022).

[22] * in UniProtKB search (230496503) | UniProt n.d. https://www.uniprot.org/uniprotkb?query=* (accessed November 21, 2022).

[23] National Center for Biotechnology Information n.d. https://www.ncbi.nlm.nih.gov/ (accessed November 21, 2022).

[24] GeneCards - Human Genes | Gene Database | Gene Search n.d. https://www.genecards.org/ (accessed November 21, 2022).

[25] Animal QTL Database n.d. https://www.animalgenome.org/cgi-bin/QTLdb/index (accessed November 21, 2022).

[26] Xu L, Hou Y, Bickhart DM, Zhou Y, Hay EH abdel, Song J, et al. Population-genetic properties of differentiated copy number variations in cattle. Sci Rep 2016;6:23161. https://doi.org/10.1038/srep23161.

[27] Vajana E, Barbato M, Colli L, Milanesi M, Rochat E, Fabrizi E, et al. Combining Landscape Genomics and Ecological Modelling to Investigate Local Adaptation of Indigenous Ugandan Cattle to East Coast Fever. Front Genet 2018;9:385. https://doi.org/10.3389/fgene.2018.00385.

[28] Zhang M. DETERMINATION OF THE CONSENSUS TARGET SEQUENCES RECOGNIZED BY ZNFO, A NOVEL OOCYTE-SPECIFIC ZINC FINGER TRANSCRIPTION FACTOR IN CATTLE. MS. West Virginia University Libraries, 2018. https://doi.org/10.33915/etd.7286.

[29] Ganss B, Jheon A. Zinc Finger Transcription Factors in Skeletal Development. Crit Rev Oral Biol Med 2004;15:282–97. https://doi.org/10.1177/154411130401500504.

[30] Sobocińska J, Molenda S, Machnik M, Oleksiewicz U. KRAB-ZFP Transcriptional Regulators Acting as Oncogenes and Tumor Suppressors: An Overview. Int J Mol Sci 2021;22:2212. https://doi.org/10.3390/ijms22042212.

[31] Muniz L, Nicolas E, Trouche D. RNA polymerase II speed: a key player in controlling and adapting transcriptome composition. EMBO J 2021;40:e105740. https://doi.org/10.15252/embj.2020105740.

[32] Santoni de Sio F. Kruppel-associated box (KRAB) proteins in the adaptive immune system. Nucl Austin Tex 2014;5. https://doi.org/10.4161/nucl.28738.

[33] M. Zhang, H. Baldwin, J. Current, J. Yao. 35 Characterization of the promoter region of ZNFO, an oocyte-specific gene in cattle. Reprod Fertil Dev 2021;33:125–125. https://doi.org/10.1071/RDv33n2Ab35.

[34] Beckett L, Xie S, Thimmapuram J, Tucker HA, Donkin SS, Casey T. Mammary transcriptome reveals cell maintenance and protein turnover support milk synthesis in early-lactation cows. Physiol Genomics 2020;52:435–50. https://doi.org/10.1152/physiolgenomics.00046.2020.

[35] Wen H, Luo H, Yang M, Augustino SMA, Wang D, Mi S, et al. Genetic parameters and weighted single-step genome-wide association study for supernumerary teats in Holstein cattle. J Dairy Sci 2021;104:11867–77. https://doi.org/10.3168/jds.2020-19943.

[36] Massa AT, Mousel MR, Durfee CJ, Herndon MK, Hemmerling KM, Taylor JB, et al. A DNA Regulatory Element Haplotype at Zinc Finger Genes Is Associated with Host Resilience to Small Ruminant Lentivirus in Two Sheep Populations. Animals 2021;11:1907. https://doi.org/10.3390/ani11071907.

[37] Drobik-Czwarno W, Wolc A, Fulton JE, Arango J, Jankowski T, O'Sullivan NP, et al. Identifying the genetic basis for resistance to avian influenza in commercial egg layer chickens. Animal 2018;12:1363–71. https://doi.org/10.1017/S1751731117002889.

[38] Gupta SK, Rai AK, Kanwar SS, Sharma TR. Comparative Analysis of Zinc Finger Proteins Involved in Plant Disease Resistance. PLoS ONE 2012;7:e42578. https://doi.org/10.1371/journal.pone.0042578.

[39] Giri J, Vij S, Dansana PK, Tyagi AK. Rice A20/AN1 zinc-finger containing stress-associated proteins (SAP1/11) and a receptor-like cytoplasmic kinase (OsRLCK253) interact via A20 zinc-finger and confer abiotic stress tolerance in transgenic Arabidopsis plants. New Phytol 2011;191:721–32. https://doi.org/10.1111/j.1469-8137.2011.03740.x.

[40] Feurtado JA, Huang D, Wicki-Stordeur L, Hemstock LE, Potentier MS, Tsang EWT, et al. The Arabidopsis C2H2 zinc finger INDETERMINATE DOMAIN1/ENHYDROUS promotes the transition to germination by regulating light and hormonal signaling during seed maturation. Plant Cell 2011;23:1772–94. https://doi.org/10.1105/tpc.111.085134.

[41] ZNF746 zinc finger protein 746 [Homo sapiens (human)] - Gene - NCBI n.d. https://www.ncbi.nlm.nih.gov/gene/155061#summary (accessed November 27, 2022).

[42] GIMAP8 - GTPase IMAP family member 8 - Homo sapiens (Human) | UniProtKB | UniProt n.d. https://www.uniprot.org/uniprotkb/Q8ND71/entry#function (accessed November 27, 2022).

[43] Berg HE, Blackburn PR, Baughn LB, Ketterling RP, Xu X, Greipp PT, et al. Identification of a novel KMT2A/GIMAP8 gene fusion in a pediatric patient with acute undifferentiated leukemia. Genes Chromosomes Cancer 2021;60:108–11. https://doi.org/10.1002/gcc.22902.

[44] Robbertse L, Richards SA, Maritz-Olivier C. Bovine Immune Factors Underlying Tick Resistance: Integration and Future Directions. Front Cell Infect Microbiol 2017;7. https://doi.org/10.3389/fcimb.2017.00522.

[45] RARRES2 retinoic acid receptor responder 2 [Homo sapiens (human)] - Gene - NCBI n.d. https://www.ncbi.nlm.nih.gov/gene/5919#summary (accessed November 27, 2022).

[46] Bondue B, Wittamer V, Parmentier M. Chemerin and its receptors in leukocyte trafficking, inflammation and metabolism. Cytokine Growth Factor Rev 2011;22:331–8. https://doi.org/10.1016/j.cytogfr.2011.11.004.

[47] Geering B, Stoeckle C, Rozman S, Oberson K, Benarafa C, Simon H-U. DAPK2 positively regulates motility of neutrophils and eosinophils in response to intermediary chemoattractants. J Leukoc Biol 2014;95:293–303. https://doi.org/10.1189/jlb.0813462.

[48] Kemp DH, Bourne A. Boophilus microplus: the effect of histamine on the attachment of cattle-tick larvae--studies in vivo and in vitro. Parasitology 1980;80:487–96. https://doi.org/10.1017/s0031182000000950.

[49] Schleger AV, Lincoln DT, McKenna RV, Kemp DH, Roberts JA. Boophilus microplus: cellular responses to larval attachment and their relationship to host resistance. Aust J Biol Sci 1976;29:499–512. https://doi.org/10.1071/bi9760499.

[50] Marufu MC, Chimonyo M, Mans BJ, Dzama K. Cutaneous hypersensitivity responses to Rhipicephalus tick larval antigens in pre-sensitized cattle. Ticks Tick-Borne Dis 2013;4:311–6. https://doi.org/10.1016/j.ttbdis.2012.12.001.

[51] Tabor AE, Ali A, Rehman G, Rocha Garcia G, Zangirolamo AF, Malardo T, et al. Cattle Tick Rhipicephalus microplus-Host Interface: A Review of Resistant and Susceptible Host Responses. Front Cell Infect Microbiol 2017;7. https://doi.org/10.3389/fcimb.2017.00506.

[52] TRPM8 transient receptor potential cation channel subfamily M member 8 [Homo sapiens (human)] - Gene - NCBI n.d. https://www.ncbi.nlm.nih.gov/gene/79054#summary (accessed November 25, 2022).

[53] MROH2A maestro heat like repeat family member 2A [Homo sapiens (human)] - Gene - NCBI n.d. https://www.ncbi.nlm.nih.gov/gene/339766#summary (accessed November 25, 2022).

[54] Nejat S, Dehbi M, Thornalley P, Moustaid-Moussa N, Scoggin S, Menikdiwela K. The Effect of DNAJB3 Deficiency on Metabolic Dysfunctions in Diet-induced Obese Mice. Curr Dev Nutr 2022;6:1078. https://doi.org/10.1093/cdn/nzac070.037.

[55] Yamashita H, Wang Z, Wang Y, Furuyama T, Kontani Y, Sato Y, et al. Impaired basal thermal homeostasis in rats lacking capsaicin-sensitive peripheral small sensory neurons. J Biochem (Tokyo) 2008;143:385–93. https://doi.org/10.1093/jb/mvm233.

[56] Feske S, Wulff H, Skolnik EY. Ion Channels in Innate and Adaptive Immunity. Annu Rev Immunol 2015;33:291–353. https://doi.org/10.1146/annurev-immunol-032414-112212.

[57] Zhang P, Zheng C-B, Chen Z, Liu X-Y. Editorial: The Role of Calcium Channels in Human Health and Disease. Front Mol Biosci 2022;9:834108. https://doi.org/10.3389/fmolb.2022.834108.

[58] Paudel S, Sindelar R, Saha M. Calcium Signaling in Vertebrate Development and Its Role in Disease. Int J Mol Sci 2018;19:3390. https://doi.org/10.3390/ijms19113390.

[59] Reimúndez Dubra A, Fernández-Peña C, García G, Fernandez Caloto R, Ordás P, Gallego R, et al. Deletion of the Cold Thermoreceptor TRPM8 Increases Heat Loss and Food Intake Leading to Reduced Body Temperature and Obesity in Mice. J Neurosci 2018;38:3002–17. https://doi.org/10.1523/JNEUROSCI.3002-17.2018.

[60] González-Muñiz R, Bonache MA, Martín-Escura C, Gómez-Monterrey I. Recent Progress in TRPM8 Modulation: An Update. Int J Mol Sci 2019;20:2618. https://doi.org/10.3390/ijms20112618.

[61] Tian R, Nanaie HA, Wang X, Zhao M, Wang F, Li H, et al. Genomic adaptation to extreme climate conditions in beef cattle as a consequence of cross-breeding progra 2022. https://doi.org/10.21203/rs.3.rs-1971517/v1.

[62] Majhi RK, Saha S, Kumar A, Ghosh A, Swain N, Goswami L, et al. Expression of temperature-sensitive ion channel TRPM8 in sperm cells correlates with vertebrate evolution. PeerJ 2015;3:e1310. https://doi.org/10.7717/peerj.1310.

[63] Reddi AH. Bone morphogenetic proteins and skeletal development: the kidney-bone connection. Pediatr Nephrol 2000;14:598–601. https://doi.org/10.1007/s004670000364.

[64] Trukhachev V, Skripkin V, Kvochko A, Kulichenko A, Kovalev D, Pisarenko S, et al. Correlation between gene expression profiles in muscle and live weight in Dzhalginsky Merino sheep. Rev Colomb Cienc Pecu 2016;29:188–98. https://doi.org/10.17533/udea.rccp.v29n3a04.

[65] Snyman G, Süllwald S, Olivier W, Visser C. Identification of Genes Targeted in South African Merino and Afrino Sheep Populations Under Long-Term Selection for Reproduction and Body Weight. 2020. https://doi.org/10.21203/rs.3.rs-128951/v1.

[66] Nucleoplasm - an overview | ScienceDirect Topics n.d. https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/nucleoplasm (accessed November 25, 2022).

[67] Nucleus | Definition, Function, Structure, & Facts | Britannica n.d. https://www.britannica.com/science/nucleus-biology (accessed November 25, 2022).

[68] MB myoglobin [Homo sapiens (human)] - Gene - NCBI n.d. https://www.ncbi.nlm.nih.gov/gene/4151#summary (accessed November 25, 2022).

[69] Moore LG, Zamudio S, Zhuang J, Droma T, Shohet RV. Analysis of the Myoglobin Gene in Tibetans Living at High Altitude. High Alt Med Biol 2002;3:39–47. https://doi.org/10.1089/152702902753639531.

[70] Cassar-Malek I, Passelaigue F, Bernard C, Léger J, Hocquette J-F. Target genes of myostatin loss-of-function in muscles of late bovine fetuses. BMC Genomics 2007;8:63. https://doi.org/10.1186/1471-2164-8-63.

[71] Mortimer SI, Fogarty NM, van der Werf JHJ, Brown DJ, Swan AA, Jacob RH, et al. Genetic correlations between meat quality traits and growth and carcass traits in Merino sheep1. J Anim Sci 2018;96:3582–98. https://doi.org/10.1093/jas/sky232.

[72] King DA, Shackelford SD, Kuehn LA, Kemp CM, Rodriguez AB, Thallman RM, et al. Contribution of genetic influences to animal-to-animal variation in myoglobin content and beef lean color stability1. J Anim Sci 2010;88:1160–7. https://doi.org/10.2527/jas.2009-2544.

[73] Cross AJ, King DA, Shackelford SD, Wheeler TL, Nonneman DJ, Keel BN, et al. Genome-Wide Association of Myoglobin Concentrations in Pork Loins. Meat Muscle Biol 2018;2. https://doi.org/10.22175/mmb2017.08.0042.

[74] PubChem. UQCRFS1 - ubiquinol-cytochrome c reductase, Rieske iron-sulfur polypeptide 1 (human) n.d. https://pubchem.ncbi.nlm.nih.gov/gene/UQCRFS1/human (accessed November 25, 2022).

[75] Dorji J, MacLeod IM, Chamberlain AJ, Vander Jagt CJ, Ho PN, Khansefid M, et al. Mitochondrial protein gene expression and the oxidative phosphorylation pathway associated with feed efficiency and energy balance in dairy cattle. J Dairy Sci 2021;104:575–87. https://doi.org/10.3168/jds.2020-18503.

[76] Cell Differentiation - an overview | ScienceDirect Topics n.d. https://www.sciencedirect.com/topics/medicine-and-dentistry/cell-differentiation (accessed November 26, 2022).

[77] Cell Differentiation - Creative Diagnostics n.d. https://www.creative-diagnostics.com/cell-differentiation.htm (accessed November 26, 2022).

[78] Du M, Huang Y, Das AK, Yang Q, Duarte MS, Dodson MV, et al. MEAT SCIENCE AND MUSCLE BIOLOGY SYMPOSIUM: Manipulating mesenchymal progenitor cell differentiation to optimize performance and carcass value of beef cattle1,2. J Anim Sci 2013;91:1419–27. https://doi.org/10.2527/jas.2012-5670.

[79] TMEFF1 transmembrane protein with EGF like and two follistatin like domains 1 [Homo sapiens (human)] - Gene - NCBI n.d. https://www.ncbi.nlm.nih.gov/gene/8577#summary (accessed November 26, 2022).

[80] Lo H, Lim Y-W, Xiong Z, Martel N, Ferguson C, Ariotti N, et al. Cavin4 interacts with Bin1 to promote T-tubule formation and stability in developing skeletal muscle. J Cell Biol 2021;220. https://doi.org/10.1083/jcb.201905065.

[81] Leal-Gutiérrez JD, Rezende FM, Elzo MA, Johnson D, Peñagaricano F, Mateescu RG. Structural Equation Modeling and Whole-Genome Scans Uncover Chromosome Regions and Enriched Pathways for Carcass and Meat Quality in Beef. Front Genet 2018;9:532. https://doi.org/10.3389/fgene.2018.00532.

[82] CAVIN4 caveolae associated protein 4 [Homo sapiens (human)] - Gene - NCBI n.d. https://www.ncbi.nlm.nih.gov/gene/347273#summary (accessed November 26, 2022).

[83] Bastiani M, Liu L, Hill MM, Jedrychowski MP, Nixon SJ, Lo HP, et al. MURC/Cavin-4 and cavin family members form tissue-specific caveolar complexes. J Cell Biol 2009;185:1259–73. https://doi.org/10.1083/jcb.200903053.

[84] Kang H, Zhao D, Xiang H, Li J, Zhao G, Li H. Large-scale transcriptome sequencing in broiler chickens to identify candidate genes for breast muscle weight and intramuscular fat content. Genet Sel Evol 2021;53:66. https://doi.org/10.1186/s12711-021-00656-9.

[85] Pan C, Lei Z, Wang S, Wang X, Wei D, Cai X, et al. Genome-wide identification of cyclin-dependent kinase (CDK) genes affecting adipocyte differentiation in cattle. BMC Genomics 2021;22:532. https://doi.org/10.1186/s12864-021-07653-8.

[86] Ghebremariam MK, Rutten VPMG, Vernooij JCM, Uqbazghi K, Tesfaalem T, Butsuamlak T, et al. Prevalence and risk factors of bovine tuberculosis in dairy cattle in Eritrea. BMC Vet Res 2016;12:80. https://doi.org/10.1186/s12917-016-0705-9.

[87] Alfonzo EPM, Barbosa da Silva MVG, dos Santos Daltro D, Stumpf MT, Dalcin VC, Kolling G, et al. Relationship between physical attributes and heat stress in dairy cattle from different genetic groups. Int J Biometeorol 2016;60:245–53. https://doi.org/10.1007/s00484-015-1021-y.

[88] Maiorano AM, Oliveira MCS, Cyrillo JNSG, Albuquerque LG, Curi RA, Silva J a. I. Genetic study of skin thickness and its association with postweaning growth in Nellore cattle: estimation of the genetic parameters. Genet Mol Res 2016;15. https://doi.org/10.4238/gmr.15017124.

[89] Hamid: Skin thickness in relation to milk production... - Google Scholar n.d. https://scholar.google.com/scholar_lookup?title=Skin%20thickness%20in%20relation%20to%20milk%20production%20of%20crossbred%20cows&publication_year=2000&author=M.A.%20Hamid&author=S.M.I.%20Husain&author=M.K.I.%20Khan&author=M.N.%20Islam&author=M.A.%20Biswas (accessed November 22, 2022).

[90] Zhang H, Liu A, Li X, Xu W, Shi R, Luo H, et al. Genetic analysis of skinfold thickness and its association with body condition score and milk production traits in Chinese Holstein population. J Dairy Sci 2019;102:2347–52. https://doi.org/10.3168/jds.2018-15180.

[91] Kichaev G, Bhatia G, Loh P-R, Gazal S, Burch K, Freund MK, et al. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. Am J Hum Genet 2019;104:65–75. https://doi.org/10.1016/j.ajhg.2018.11.008.

[92] Yeo GSH, Connie Hung C-C, Rochford J, Keogh J, Gray J, Sivaramakrishnan S, et al. A de novo mutation affecting human TrkB associated with severe obesity and developmental delay. Nat Neurosci 2004;7:1187–9. https://doi.org/10.1038/nn1336.

[93] Belote RL, Le D, Maynard A, Lang UE, Sinclair A, Planells-Palop V, et al. Single cell analysis of human site-specific melanocyte differentiation and the decoding of

developmental programs in melanoma 2020:2020.05.25.115287. https://doi.org/10.1101/2020.05.25.115287.

[94] Schleger A, Bean K. The Melanocyte System of Cattle Skin II. Melanotic Melanocytes of Epidermis and Dermis. Aust J Biol Sci 1973;26:985. https://doi.org/10.1071/BI9730985.

[95] Esmaeili-Fard SM, Gholizadeh M, Hafezian SH, Abdollahi-Arpanahi R. Genome-wide association study and pathway analysis identify NTRK2 as a novel candidate gene for litter size in sheep. PloS One 2021;16:e0244408. https://doi.org/10.1371/journal.pone.0244408.

[96] Benjamin AL, Green BB, Crooker BA, McKay SD, Kerr DE. Differential responsiveness of Holstein and Angus dermal fibroblasts to LPS challenge occurs without major differences in the methylome. BMC Genomics 2016;17:258. https://doi.org/10.1186/s12864-016-2565-x.

[97] Dermal Fibroblast - an overview | ScienceDirect Topics n.d. https://www.sciencedirect.com/topics/engineering/dermal-fibroblast (accessed November 22, 2022).

[98] Kwon A-H, Qiu Z, Hirao Y. Topical application of plasma fibronectin in full-thickness skin wound healing in rats. Exp Biol Med Maywood NJ 2007;232:935–41.

[99] Kim MO, Ryu JM, Suh HN, Park SH, Oh Y-M, Lee SH, et al. cAMP Promotes Cell Migration Through Cell Junctional Complex Dynamics and Actin Cytoskeleton Remodeling: Implications in Skin Wound Healing. Stem Cells Dev 2015;24:2513–24. https://doi.org/10.1089/scd.2015.0130.

[100] Kopecki Z, Yang GN, Arkell RM, Jackson JE, Melville E, Iwata H, et al. Flightless I over-expression impairs skin barrier development, function and recovery following skin blistering. J Pathol 2014;232:541–52. https://doi.org/10.1002/path.4323.

[101] Leduc C, Sobilo L, Toumi H, Mondon P, Lespessailles E, Ossant F, et al. TGF-beta-induced early gene-1 overexpression promotes oxidative stress protection and actin cytoskeleton rearrangement in human skin fibroblasts. Biochim Biophys Acta 2016;1860:1071–8. https://doi.org/10.1016/j.bbagen.2016.02.009.

[102] McCafferty D-M, Craig AWB, Senis YA, Greer PA. Absence of Fer protein-tyrosine kinase exacerbates leukocyte recruitment in response to endotoxin. J Immunol Baltim Md 1950 2002;168:4930–5. https://doi.org/10.4049/jimmunol.168.10.4930.

[103] Niciura SCM, Benavides MV, Okino CH, Ibelli AMG, Minho AP, Esteves SN, et al. Genome-Wide Association Study for Haemonchus contortus Resistance in Morada Nova Sheep. Pathog Basel Switz 2022;11:939. https://doi.org/10.3390/pathogens11080939.

[104] Qi W, Ebbert KVJ, Craig AWB, Greer PA, McCafferty D-M. Absence of Fer protein tyrosine kinase exacerbates endotoxin induced intestinal epithelial barrier dysfunction in vivo. Gut 2005;54:1091–7. https://doi.org/10.1136/gut.2004.061887.

[105] Foster LA, Fourie PJ, Neser FWC, Fair MD. Differences in physical traits such as coat score and hide-thickness together with tick burdens and body condition score in four breeds in the Southern Free State n.d.:4.

[106] Fourie PJ, Foster LA, Neser FWC. Score and hide-thickness, together with tick burden and body condition score, in four cattle breeds in the South-eastern Free State province of South Africa n.d.;11:8.

6 CHAPTER 6:
General conclusions and recommendations

6.1 Exploring genetic variation

The genetic variation that exists in tick count between and within breeds should be exploited to successfully address the challenge of tick infestations. Tick-resistant indigenous breeds, such as the Nguni cattle are generally not considered in commercial beef production systems due to their poor growth performance and meat attributes. This precludes the incorporation of the needed economically important trait of tick resistance into the beef production systems. Farmers use chemical acaricides, vaccines, and environmental methods, which have not been effective in eradicating ticks. Crossbreeding is a common mating system used to exploit breed complementarity that can be used to incorporate tick resistance into beef production systems. The South African Agricultural Research Council recently initiated a research program, where the tick-tolerant Nguni was crossed with tick-susceptible Angus, which has excellent growth and meat characteristics. This project aimed to create a resource population for the discovery of genetic markers or causative mutations for tick resistance.

Since this is the first crossbreeding experiment involving these breeds, the potential of improving tick count in this crossbred population was explored by determining genetic variations for this trait. The artificial infestation was performed successfully using calico bags to assess the tick count. It is generally known that enumeration of tick counts is costly, time consuming, and labor-intensive; therefore, there is a need for the development of new phenotyping strategies for assessing tick burdens/loads. The use of calico bags serves as a new approach to assessing tick count in cattle. The disadvantage associated with this method is related to the longevity of the adhesion. In this research study, the use of calico bags to measure tick count provided a labor-manageable, cost and time-effective approach to measure the trait. Therefore, this is a promising method, and further use of this method is recommended to explore and improve its effectiveness.

Genetic parameters were initially assessed using pedigree data. The current study showed that tick count is moderately heritable, with a heritability estimate of 21%, and this was consistent with previous studies. Moreover, the heritability of tick count estimated using an animal model fitting kinship matrix showed a genetic variation of 12%. This discrepancy indicates the need for the collection of more data since the data used for the current study only included 266 animals. The alternative to improve tick count directly is to select a correlated trait that affects resistance to ticks. Skin thickness is one of the traits that affect tick

count. This is because the skin is the attachment site for tick feeding and its vascular system, which is the source of blood for the ticks, varies from one animal to another. In this regard, a moderate heritability estimate for skin thickness was obtained. This implies that a response to selection may be expected if genetic selection is applied to skin thickness. However, due to the limited number of animals, genetic correlations could not be estimated.

Equally important in beef production systems are the growth traits. Thus, genetic variations for birth, weaning, and yearling weights were evaluated in the F_2 Angus x Nguni population to determine the potential of simultaneously improving these traits and tick count. On the other hand, all the growth traits studied were moderately heritable. Thus, sufficient genetic variation exists for these traits and their improvement through genetic selection is possible. However, due to the population size, it was not possible to estimate the genetic correlations to determine if selecting growth traits can result in a correlated response in tick count. This would have determined if simultaneous selection for growth traits and tick count is possible.

6.2 Genetic influence on economic traits

Genomic technologies have made it possible to identify genes and genomic regions affecting a given trait. Genome-wide association analyses were performed for tick count, skin thickness, and growth traits. Genomic regions influencing these traits were subsequently identified, where two SNPs were identified on chromosomes 4 and 22 for tick count. For skin thickness SNPs were found on chromosomes 2, 7, 8, 9, and 12. For birth weight, the SNPs were found on chromosomes 3 and 12, while SNPs associated with weaning weight were found on chromosomes 5, 8, and 18. One SNP located on chromosome 8 was associated with yearling weight. These results show that tick count does not have any common SNPs with the weight traits analyzed, which may indicate that there are no common genes.

The absence of common genes between tick count and weight traits may suggest a low genetic correlation. This suggests that tick count should be included in the breeding objective together with the weight traits to achieve simultaneous improvement of both traits. Improvement in tick count or vice versa may thus be expected without adversely affecting the growth traits. Although skin thickness has been reported to affect resistance to ticks, results from the current study show that there are no common SNPs and genes between tick count and skin thickness. This may also indicate a low genetic correlation between these traits, suggesting that skin thickness may not be selected with an expectation of a correlated response in tick count. It should be noted, however, that the absence of common SNPs and

genes does not preclude the interaction of genes located in different chromosomes. The interaction of such genes may also result in significant genetic correlations. More reliable results may be obtained if a larger population size is used in the analysis.

The F_2 population was a suitable design for the study undertaken, however, the small population may be the reason significant SNPs were missed because the power of the study was reduced. It is generally known that it is not easy to discover significant SNPs for complex traits using a small population size. In addition to population size, a single SNP GWAS approach was used for the analysis, this approach has the potential to miss SNPs, which contributes to the heritability of the trait being studied. Another challenge is that the single-locus approach is not capable to estimate SNP effects and detect epistatic interactions within or between close genes relevant to the expression of the trait. In this context, an integrated approach of single-locus GWAS with multi-locus GWAS is recommended. Integration of these two approaches may improve the capacity of an association study to detect significant SNPs and a better understanding of the genetic architecture of the traits studied. In the current study, the integration approach was not conducted; however, it is one of the next steps for our project.

6.3 Functional annotation of genes

Genes control and or influence traits in animals; hence, their expression and ultimately the functions they perform to exert their influence are of paramount importance. These genes are involved in a variety of pathways that influence certain biochemical activities in an animal. In the current study, several genes associated with tick count were identified, which control a variety of cellular, molecular, and biological processes. Most of these genes are involved in response to stimuli, immune response, and disease resistance. Therefore, these genes can be used as potential biomarkers and targets for selection in the improvement programs for tick count. Five genes were mapped for skin thickness, which like the genes for ticks, were involved in the immunity of the animal. These genes can also be considered in the selection to improve tick count. Although there are no common genes between tick count and skin thickness, these genes perform common functions.

Genes associated with birth weight control skeletal development and growth. Similarly, genes were uncovered for weight-controlled skeletal muscle development and growth. They also affected cardiac muscle development and meat colour. For weaning weight, the genes uncovered were associated with skeletal muscle and myofibril growth and cell differentiation. These results show that growth traits are controlled by genes that perform functions that are

different from those that control tick count. Therefore, tick count must be included in breeding programs, together with growth traits.

Initially, DAVID was used for gene set enrichments as a broader tool. However, since this bioinformatics platform was purposely developed for humans, we integrated it with ShinyGo, which is a recently developed tool for animals and plants. ShinyGo is an intuitive graphical tool for gene set enrichment; therefore integrating it with DAVID provided an actionable insight into the biological mechanism underlying tick count, growth traits, and skin thickness.

6.4 Recommendations for future work

Heritability estimates obtained from pedigree analysis were characterized by large standard errors, which could be due to the sample size used in the current study. Furthermore, if the sample size is increased, an animal model may be used instead of a sire model used in this study, to yield more accurate genetic parameter estimates. A larger sample size may reveal significant SNPs, especially for tick count, which only had SNPs reaching suggestive thresholds.

This study used a 150K SNP chip, which only works with predetermined SNPs along the genome. A denser chip, such as the 777K SNP chip may be considered for future studies that will contain more SNPs than the 150K SNP chip. Furthermore, instead of working with predetermined SNPs loaded onto a chip, whole genome sequence data could be used. The advantage of whole-genome sequence data is that, in addition to the genes, it can also detect single nucleotide variants, insertions/deletions, copy number changes, and large structural variants. These are additional genetic determinants that are known to affect tick count.

The SNP chip used in this study was developed using exotic breeds; hence genetic content from local South African breeds may not be well represented in these chips. Representation of local breeds such as the Nguni in the SNP chip may be advantageous. The absence of information on the African breed in current bovine SNP genotyping arrays impedes the success of identifying SNPs influencing production and adaptive traits in African breeds. Therefore, the establishment of new African breeds' specific SNP genotyping arrays will enhance the utility of commercial chips in gene discovery studies. Furthermore, this will improve the economic traits breeding criterion. Improved breeding criteria will ensure improved animal welfare, food security, and profitability.

6.5 Shortfalls of the study

Nutrition plays an important role in influencing the traits studied. At the age of 12 months, the F_2 Angus X Nguni cattle were raised at a feedlot for a period of 3 months, however, the nutrition data was not factored into the study analyses. Moreover, the trials were conducted in batches, represented by different years; this implies that the environmental conditions were not the same. This underscored the need to factor in climate data during data analyses. The factoring of climate data into the analyses was not possible because the attempt to get hold of climate data was not successful. The inclusion of nutrition and climate data during data analyses would have improved the findings of the study. In closing, the last trial was conducted during the year 2020, from March until the end of June, which was the time during the COVID-19 pandemic lockdown, this could have affected the handling of the animals since there was a shortage of manpower, co-workers were at home because of the government restrictions.

List of Annexures

Annexure 1.1: Clustering of enriched GO for tick count.

Annotation Cluster 1	**Enrichment Score: 3.72**						
Category	**Term**	Genes	**Count**	**%**	**PValue**	**FE**	**FDR**
INTERPRO	IPR006703:AIG1	GIMAP7, GIMAP8	5	25	2.39E-10	438.25	6.46E-09
UP_SEQ_FEATURE	DOMAIN:AIG1-type G	GIMAP7, GIMAP8	5	25	3.04E-10	416.3	3.07E-08
GOTERM_MF_DIRECT	GO:0005525~GTP binding	GIMAP7, GIMAP8	5	25	4.11E-04	12.84	1.00E-02
UP_KW_LIGAND	KW-0342~GTP-binding	GIMAP7, GIMAP8	5	25	0.00	6.77	2.00E-02
INTERPRO	IPR027417:P-loop containing nucleoside triphosphate hydrolase	GIMAP7, GIMAP8	5	25	0.01	6.49	3.00E-02
GOTERM_CC_DIRECT	GO:0005829~cytosol	ZNF746, GIMAP7, GIMAP8	6	30	0.04	2.85	6.60E-01
Annotation Cluster 2	**Enrichment Score: 3.33**						
Category	**Term**	Genes	**Count**	**%**	**PValue**	**FE**	**FDR**
INTERPRO	IPR022137:Protein of unknown function DUF3669, zinc finger protein	ZNF746,ZNF777, ZNF212, ZNF398	4	20	7.56E-09	818.07	1.02E-07
SMART	SM00355:ZnF_C2H2	REPIN1,ZNF746,ZNF777, ZNF775,ZNF467,ZNF786, ZNF212, ZNF398	8	40	2.26E-07	12.49	2.71E-06
INTERPRO	IPR013087:Zinc finger C2H2-type/integrase DNA-binding domain	REPIN1, ZNF746, ZNF777, ZNF775, ZNF467, ZNF786, ZNF212, ZNF398	8	40	2.43E-07	15.41	2.18E-06
GOTERM_BP_DIRECT	GO:0006357~regulation of transcription from RNA polymerase II promoter	REPIN1, ZNF746, ZNF777, ZNF775, ZNF467, ZNF786, ZNF212, ZNF398	8	40	1.10E-05	7.74	3.85E-04
INTERPRO	IPR001909:Krueppel-associated box	ZNF746, ZNF777, ZNF862, ZNF212, ZNF398	5	25	2.49E-05	26.56	1.68E-04
UP_SEQ_FEATURE	DOMAIN:KRAB	ZNF746, ZNF777, ZNF862, ZNF212, ZNF398	5	25	2.89E-05	25.79	1.03E-07
SMART	SM00349:KRAB	ZNF746, ZNF777, ZNF862, ZNF212, ZNF398	5	25	3.73E-05	21.44	1.28E-07

KEGG_PATHWAY	bta05168:Herpes simplex virus 1 infection	ZNF746, ZNF777, ZNF786, ZNF212, ZNF398	5	25	6.23E-05	15.48	0.00E+00
UP_SEQ_FEATURE	DOMAIN:C2H2-type	ZNF777, ZNF775, ZNF467, ZNF786, ZNF212, ZNF398	6	30	9.39E-05	11.6	0.00E+00
GOTERM_MF_DIRECT	GO:0003700~transcription factor activity, sequence-specific DNA binding	ZNF746, ZNF467, ZNF212, ZNF398	4	20	0.00	16.6	1.00E-02
GOTERM_MF_DIRECT	GO:0000978~RNA polymerase II core promoter proximal region sequence-specific DNA binding	ZNF746, ZNF775, ZNF467, ZNF786, ZNF212, ZNF398	6	30	0.00	5.9	1.00E-02
UP_KW_DOMAIN	KW-0677~Repeat	REPIN1, ZNF746, LRRC61, SSPO, ZNF777, ZNF212, PDIA4, ZNF398	8	40	0.02	2.58	1.20E-01
UP_KW_DOMAIN	KW-0863~Zinc-finger	REPIN1, ZNF746, ZNF775, ZNF467, ZNF786	5	25	0.03	4.04	1.20E-01
GOTERM_MF_DIRECT	GO:0046872~metal ion binding	REPIN1, ZNF746, ZNF777, ZNF212, ZNF398	5	25	0.06	3.14	1.52E-07
UP_KW_LIGAND	KW-0862~Zinc	REPIN1, ZNF746, ZNF777, ZNF775, ZNF467, ZNF786, ZNF212, ZNF398	8	40	0.22	1.44	1.76E-07
UP_SEQ_FEATURE	REGION: Disordered	REPIN1, ZNF746, SSPO, ZNF777, ZNF775, KRBA1, ZNF467, ZNF786, ZNF212, GIMAP8, PDIA4, ZNF398	12	60	0.37	1.16	1.00E+00
GOTERM_CC_DIRECT	GO:0005634~nucleus	ZNF746, ZNF775, ZNF467, ZNF786	4	20	0.53	1.38	1.00E+00
UP_KW_LIGAND	KW-0479~Metal-binding	REPIN1, ZNF746, ZNF777, ZNF775, ZNF467, ZNF786, ZNF212, ZNF398	8	40	0.59	1.08	7.40E-01

Annexure 1. 2: List of identified genes within the 1 Mb region surrounding the suggestive SNPs for growth traits (birth, weaning, and yearling weight), and skin thickness

Trait	BTA	position	**Ensembl ID**	**Nearest gene**	**Description**
BWT	12	51067836	ENSBTAG00000048557	U6	U6 spliceosomal RNA
			ENSBTAG00000010693	LMO7	LIM domain 7
			ENSBTAG00000054578	COMMD6	COMM domain containing 6
			ENSBTAG00000008024	UCHL3	Ubiquitin C-terminal hydrolase L3
			ENSBTAG00000032878	Unknown	Uncharacterized
			ENSBTAG00000049325	bta-mir-2285ab	bta-mir-2285ab
	3	113815818	ENSBTAG00000014652	TRPM8	Transient receptor potential cation channel subfamily M member 8
			ENSBTAG00000002030	SPP2	Secreted phosphoprotein 2
			ENSBTAG00000024726	HJURP	Holliday junction recognition protein
			ENSBTAG00000017434	MROH2A	Maestro heat like repeat family member 2A
			ENSBTAG00000018756	DNAJB3	DNAJ heat shock protein family (Hsp40) member B3
			ENSBTAG0000005483	Unknown	Uncharacterized
			ENSBTAG00000051335	Unknown	Uncharacterized
			ENSBTAG00000043312	Unknown	Uncharacterized
			ENSBTAG00000053153	Unknown	Uncharacterized
	12	84095269	ENSBTAG00000018237	MY016	Myosin XVI
			ENSBTAG00000002432	ABHD13	Abhydrolase domain containing 13
			ENSBTAG00000011563	TNFS13B	TNF superfamily member 13b
			ENSBTAG00000015868	LIG4	DNA ligase 4
WWT	18	840780	ENSBTAG00000046786	UQCRFS1	Ubiquinol-cytochrome c reductase, Rieske iron-sulfur polypeptide 1
			ENSBTAG00000018244	VSTM2B	V-set and transmembrane domain containing 2B

	18	156809	ENSBTAG00000053809	OR4P4	Olfactory receptor family 4 subfamily P member 4
	5	73514352	ENSBTAG00000010533	HMGXB4	HMG-box containing 4
			ENSBTAG00000045785	TOM1	The target of myb1 membrane trafficking protein
			ENSBTAG00000015595	MCM5	Minichromosome maintenance complex component 5
			ENSBTAG00000007674	ISX	Intestine specific homeobox
			ENSBTAG00000017116	RASD2	RASD family member 2
			ENSBTAG00000005333	MB	Myoglobin
			ENSBTAG00000020125	RBFOX2	RNA binding fox-1 homolog 2
			ENSBTAG00000049066	Unknown	Uncharacterized
YWT	8	90226480	ENSBTAG00000052146	MSANTD3	Myb/SANT DNA binding domain containing 3
			ENSBTAG00000021991	TMEFF1	transmembrane protein with EGF like and two follistatin like domains 1
			ENSBTAG00000021992	CAVIN4	caveolae associated protein 4
			ENSBTAG00000015171	CDK20	cyclin dependent kinase 20
			ENSBTAG00000021235	unknown	Uncharacterized
			ENSBTAG00000025756	Unknown	Uncharacterized
			ENSBTAG00000054632	Unknown	Uncharacterized
STHICK	8	78145059	ENSBTAG00000010647	NTRK2	Neurotrophic receptor tyrosine kinase 2
	7	108006295	ENSBTAG00000045264	bta-mir-2465	bta-mir-2465
			ENSBTAG00000051829	Unknown	Uncharacterized
			ENSBTAG00000003051	FER	FER tyrosine kinase
			ENSBTAG00000021675	PJA2	Praja ring finger ubiquitin ligase 2

BWT- birth weight, WWT- weaning weight, YWT- yearly weight, STHICK- skin thickness, BTA – Bos Taurus chromosome