GENETIC ANALYSIS OF HYBRIDISATION, MUTATION LOAD AND RUNS OF HOMOZYGOSITY OF GOATS FROM KENYA AND UGANDA

THEODORA CHIKOKO

A Thesis Submitted to the Graduate School in Partial Fulfilment of the Requirements for the Master of Science Degree in Animal Breeding and Genomics of Egerton University

Remove pagination from the title page

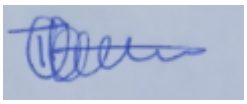
EGERTON UNIVERSITY

OCTOBER 2024

DECLARATION AND RECOMMENDATION

Declaration

This thesis is my original work and has neither wholly or in part been presented nor concurrently been presented for the award of any degree in this or any other institution.

Signature:  Date: 21st October 2024.....

Theodora Chikoko

KM111/14551/18

Recommendation

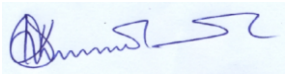
This thesis has been submitted with our approval as University supervisors.

Signature:  ..... Date: 21st October 2024……

Dr. Thomas K. Muasya, (*Dr. rer. agr.*)

Department of Animal Sciences

Egerton University

Signature:  ..... Date: 21st October 2024……

Dr. Kiplangat Ngeno, PhD.

Department of Animal Sciences

Egerton University

COPYRIGHT

DEDICATION

This thesis is dedicated to my lovely daughter Rose, the Chikoko family, and all smallholder farmers from whom I drew inspiration for this study.

ACKNOWLEDGEMENTS

First and above all, I would like to thank God Almighty for the gift of good life, wisdom, spiritual strength, and inspiration to do this research. Wish to acknowledge the CESAAM project for the financial support throughout my study period. The Directors, Department of Animal Health and Livestock Development, DAHLD-MALAWI for granting me unconditional full-time study leave. Special appreciation and my sincere gratitude to my supervisors Dr. Kiplangat Ngeno (PhD) and Dr. Thomas K. Muasya (*Dr. rer. agr.*) for the best and endless academic support, guidance, and encouragement offered throughout the research period. Thank you so much for accepting me to be your student and nurturing me throughout this academic journey. I am confident now that I am somewhere much better today than where I was. I want to appreciate Dr. Robert Onzima for his generosity in making data available to be used in this study. Dr Gilson Njunga, Dr. Joel Khobondo, Dr. Richard Habimana, and Dr. Marvin Phonera, your support and encouragement is highly appreciated. My colleagues and master students for 2019, thank you all for the unlimited support and advice. Finally, my deepest appreciation to my parents, Mr & Mrs. Chikoko, sisters -Jane, Faith, and Catherine, and my only brother Innocent for encouraging me to take up this challenge. God bless you all for taking care of my little Rosemary throughout my study period, your contribution to my study has a special place in my heart. Thank you!

ABSTRACT

Increase in human movements, improvement in reproductive technologies, and diversity of livestock management systems coupled with climate change determine the genetic structure and diversity within or between populations. However, little is known about different genetic parameters in most African goat populations, making it difficult to plan and implement genetic improvement and conservation programmes. This study focused on joint genome-wide SNP genetic analysis of goat populations from Uganda and Kenya to characterize the hybridization pattern of goats from the two nations, mutation load, Runs of Homozygosity and effective population sizes of goats from Kenya. A total of ten local and exotic goat genotypes sampled from different agro-ecological regions of Kenya (n = 94) and Uganda (n = 144) were used. Principal component, phylogenetic and admixture analysis were applied on the hybridization analysis. Results revealed that the studied populations are genetically distinct according to country of origin. This suggests that the transfer of genetic materials across the studied populations is possible for genetic improvement programs. For the mutational load (deleterious mutations), four Kenyan goat genotypes were used (Alpine, Saanen, Galla and Toggenburg). Gene annotation was done on ENSEMBLE (*Capra hircus*) using the Variant Effect Predictor (VEP) and the Biomart tool was adopted for gene ontology. Mutation load calculated through the ratio of missense to synonymous mutations was similar (0.37). Five genes were obtained from the highly deleterious mutations (SIFT score > 0.01) which include *PROS1, EHBP1, LTN1, LRRN4,* and *FNDC3A*. These results can be used as a base to predict the future rate of mutation load for future genetic improvement and conservation. Finally, Runs of Homozygosity (RoHs), effective population size, and prediction of future trends were estimated for the four goat genotypes from Kenya. Across the genotypes, 348 RoHs were detected and their distribution per chromosome was breed-specific. Higher inbreeding coefficients were observed for the Toggenburg, Saanen, and Alpine with values of 0.68, 0.35, and 0.23 respectively. Galla recorded a F_{ROHS} value of 0.09, suggesting that the genetic material for this genotype is well managed at the government station. Effective population size decreased over time across the genotypes indicating reduced genetic diversity, this will continue to decrease if recommended measures are not implemented to improve the Ne. In conclusion, goats from Uganda and Kenya are genetically distinct and can be used for cross-border genetic improvement. Furthermore, within breed selection and introduction of new exotic genetic lines must be promoted in any goat population to improve the genetic diversity.

TABLE OF CONTENTS

DECLARATION AND RECOMMENDATION i

COPYRIGHT ii

DEDICATION iii

ACKNOWLEDGEMENTS iv

ABSTRACT v

LIST OF TABLES x

LIST OF ABBREVIATIONS AND ACRONYMS/ SYMBOLS xii

CHAPTER ONE 1

INTRODUCTION 1

1.1 Background Information 1

1.2 Statement of the problem 2

1.3 Objectives 3

1.3.1. Overall Objective 3

1.3.2 Specific Objectives 3

1.4 Research Questions 3

1.5 Justification of the study 3

1.6 Limitation of the study 4

CHAPTER TWO 5

LITERATURE REVIEW 5

2.1 Goat production in East African community 5

2.1.1 Introduction to goat production in Africa 5

2.1.2 Goat breeds and breed preferences in East Africa 6

2.1.3 Breeding practices of goats in East Africa 7

2.1.4 Genetic diversity of goats in Africa 7

2.1.5 Productive and reproductive performance of goats 8

2.2 Effect of the ecological zone on the gene pool 8

2.3 Genetic improvement strategies in livestock 9

2.3.1 Within breed selection 9

2.3.2 Cross breeding strategy 10

2.3.3 Breed replacement 10

2.4 Methods of genetic characterisation 11
2.4.1 Phenotypic characterisation 11
2.4.2 Molecular characterisation 11
2.5 Runs of Homozygosity 12
2.5.1 Measurements of Inbreeding using Runs of Homozygosity 13
2.6 Effective population size 14
2.7 Mutation 14
2.8 Functional annotation 15
CHAPTER THREE 16
GENETIC ANALYSIS OF HYBRIDIZATION PATTERN IN GOAT GENOTYPES FROM KENYA AND UGANDA 16
Abstract 16
3.1 Introduction 17
3.2 Materials and methods 18
3.2.1 Study sites and Sampling 18
3.2.2 Sample collection, DNA extraction and genotyping 18
3.2.3 Quality control 19
3.2.4 Statistical analysis 19
3.3 Results 20
3.3.1 Principal Component Analysis (PCA) 20
3.3. 2 Multi-Dimensional Scaling (MDS) 21
3.3.3 Admixture analysis 22
3.3.4 Measure of genetic differentiation using the FST approach 24
3.3.5 Phylogenetic analysis 25
3.4 Discussion 25
3.4.1 Between population genetic diversity 26
3.4.2 Within populations genetic diversity 27
3.5 Conclusion 28
CHAPTER FOUR 29
MUTATIONAL LOAD AND DELETERIOUS MUTATIONS IN GOAT GENOME FROM KENYA 29
Abstract 29
4.1 Introduction 30

4.2. Materials and methods 31
4.2.1 Study site and sampling 31
4.2.2 Data analysis 31
4.3 Results 31
4.4 Discussion 35
4.5 Conclusion 37
CHAPTER ~~5~~FIVE 39
RUNS OF HOMOZYGOSITY AND EFFECTIVE POPULATION SIZE OF DIFFERENT GOAT GENOTYPES IN KENYA 39
Abstract 39
5.1 Introduction 40
5.2 Materials and Methods 41
5.2.1 Study site and sampling 41
5.2.2 Data analysis 41
5.3 Results 43
5.4 Discussion 47
5.4.1 Runs of homozygosity 47
5.4.2 Effective population size and future trends 49
5.5 Conclusion 50
CHAPTER SIX 51
GENERAL DISCUSSION, CONCLUSION AND RECOMMENDATIONS 51
6.1 Aim of the study 51
6.2 Study methodology 52
6.3 Genetic characterization of hybridization and gene introgression between the goat genotypes from Uganda and Kenya 52
6.4 Identification of mutation load and deleterious mutations in goat genotypes 53
6.5 Estimation of runs of homozygosity and effective population size of goats from Kenya 54
6.6 Conclusions 56
6.7 Recommendations 56
REFERENCES 57
APPENDICES 73
Appendix A: Percentage of chromosomes covered by Runs of homozygosities 73
Appendix B: Inbreeding coefficients per individual 74

Appendix C: Genomic regions with high frequency of Runs of homozygosities 76
Appendix D: Estimated ancestral effective population size over past generations 77
Appendix E: Genes identified from the genomic regions islands 78
Appendix F: Publication 79
Appendix G: Tropentag Conference presentation abstract 80
Appendix H: AFMA Conference presentation abstract 82
Appendix I: Research permit 83

LIST OF TABLES

Table 1. Distribution of goats in East African Community from 2018 - 2022 5
Table 2. The FST values for the ten genotypes of Kenya and Uganda obtained from admixture analysis 24
Table 3. Summary of SNP annotation for goat genotypes in Kenya 32
Table 4. GO terms enrichment for non-synonymous deleterious variants in goat populations based on SIFT prediction 33
Table 5. Descriptive statistics of runs of the homozygosities per genotype 43
Table 6. Runs of the homozygosity length categories, number of individuals with RoHs, number and mean sum of RoHs length (Mb) across the genotypes 45
Table 7. Inbreeding coefficients, Mean, and F_{RoHs} per genotype 46

LIST OF FIGURES

Figure 1. Principal components (PCA1 and PCA2) showing genetic relationships among 238 goats from Kenya (ALP, GAL, SAA and TOT) and Uganda (BOE, KAR, KIG, MUB, SEA and SEB). 21
Figure 2. Multidimensional scaling plot of 238 goats of ten genotypes from Kenya (ALP, GAL, SAA and TOT) and Uganda (BOE, KAR, KIG, MUB, SEA and SEB). 22
Figure 3. Population structure of ten goat genotypes for K = 2 to K = 10 indicating ancestral population per individual. Boe (Boer); KAR (Karamojong); KIG (Kigezi); MUB (Mubende); SEA (Small East African); SEB (Sebei); SAA (Saanen); GAL (Galla); ALP (Alpine); TOT (Toggenburg). 23
Figure 4. Cross-validation plot indicating the optimal K- value 24
Figure 5. Neighbour-joining (NJ) tree for the Ugandan (MUB, KIG, SEA, KAR, BOE and SEB) and Kenyan (GAL, TOT, SAA and ALP) goat populations studied 25
Figure 6. Coding consequences for local and exotic goats from Kenya 33
Figure 7. Runs of the homozygosity per chromosome per genotype 44
Figure 8. Distribution of mean sum of RoHs coverage per length category averaged per breed across four genotypes 45
Figure 9. The effective population size over generations for Alpine, Galla, Saanen, and Toggenburg goat genotypes 47

LIST OF ABBREVIATIONS AND ACRONYMS/ SYMBOLS

AFLP	Amplified Fragment Length Polymorphism
CBBP	Community Based Breeding Programme
CV	Cross-Validation error
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EHBP1	EH Domain Binding Protein 1
FAO	Food and Agriculture Organisation
FNDC3A	Fibronectin Type III Domain Containing 3A
F_{PED}	Pedigree inbreeding coefficient
F_{RoHs}	Runs of Homozygosity inbreeding coefficient
GenAgo	Generations Ago
GO	Gene Ontology
IBD	Identical by Descent
KSCAP	Kenya Climate Smart Agricultural Productivity Project
LD	Linkage Disequilibrium
LRRN4	Leucine Rich Repeat Neuronal 4
LTN1	Listerin E3 Ubiquitin Protein Ligase 1
Mb	Megabytes
MDS	Multi-Dimensional Scale
Ne	Effective population size
PCA	Principle Component Analysis
PROS1	Protein S
RAPD	Random Amplified Polymorphic DNA
RoHs	Runs of Homozygosity
SIFT	Sorting Intolerant From Tolerant
SNP	Single Nucleotide Polymorphism
TOT	Toggenburg
VEP	Variant Effect Predictor

CHAPTER ONE
INTRODUCTION

1.1 Background Information

East African community goat population is estimated at 98,933,106 (FAOSTAT, 2022). Goat production plays a significant socio-economic role in most developing African countries by ensuring food security through the availability of meat, milk, and sales of both meat and meat products, among other benefits (Chenyambuga & Lekule, 2014; Monau *et al.,* 2020). Goats are found in different ecological regions under diverse production and management systems mostly raised under extensive and semi-extensive management systems (Muingai *et al.,* 2017). Their feeding behaviour of grazing and browsing makes them more adaptable to various environments and management conditions; hence, they effectively utilize a wide range of feed resources compared to other ruminant species (Pragna *et al.,* 2018). Adaptation of goats in many agro-ecological regions leads to unique animals with special alleles adaptable to that particular environment due to natural selection (Ngeno *et al.,* 2015; Visser *et al.,* 2016).

Generally, goats have a wide genetic resource that allows them to survive in different agro-ecological regions hence utilizing wide range of feed resources compared to other livestock species (Pragna *et al.,* 2018). However, the productivity of meat and milk in goats is limited which can be improved through good management practices and following the recommended methods of improving desired genetic material (Abraham *et al.,* 2019; Haile *et al.,* 2020). Breeding strategies applied in most African goat populations include crossbreeding, within-breed selection, or breed replacement. Crossbreeding and within-breed selection systems are preferred in most African countries due to fast genetic gain (Peacock *et al.,* 2011). Genetic materials for different exotic breeds like Boer, Toggenburg, Anglo Nubian, German Alpines, and Saanen are found among most indigenous goat breeds in Africa due to crossbreeding (Ahuya *et al.,* 2005; Manirakiza *et al.,* 2020). However, crossbreeding programs in most nations are not implemented strategically to ensure genetic improvement and sustainability of genetic diversity since breeding of goats is not controlled (Chenyambuga & Lukule, 2014; Kahi & Wasike, 2019). The consequence is a random combination of alleles leading to the dilution of local genetic materials and the availability of unidentified goat breeds due to a combination of different private alleles (Yakubu *et al.,* 2010).

Furthermore, animal migration influenced by human migration and trade, facilitates continuous gene flow between countries, in the process affecting global genetic diversity. There is

currently enough evidence of gene structuring and some admixture levels of goat populations within East African Community (Rahmatalla *et al.,* 2017). However, the assessment of these genetic resources in Kenya and Uganda focused on within-country goat populations, ignoring the possibility of gene flow between countries (Onzima *et al.,* 2018a; Waineina *et al.,* 2021). According to the research done on cattle and goats across nations, there are more benefits in evaluating different animal genetic parameters at the international level to increase genetic progress and future breeding strategies (Opoola *et al.,* 2020; Visser *et al.,* 2016). The limited molecular characterization of goat genetic resources in Africa, uncontrolled goat breeding practices among smallholder farmers, and the advancement of technological tools for genetic characterization create the need for molecular genetic evaluation. This research therefore focused on the genetic analysis of hybridization between Kenya and Uganda, mutation load, runs of homozygosity of goat genotypes and the effective population sizes of goat genotypes from Kenya using Single Nucleotide Polymorphisms (SNPs) data. The information generated through this study is useful for better conservation and utilization of goat genetic resources in Kenya and East Africa at large.

1.2 Statement of the problem

There is limited production and productivity of goats in most populations failing to meet the evergrowing human population demands of goat meat and meat products. Different goat production programs have been implemented across African nations to improve the productivity. One of the programs is crossbreeding of goats using both local and exotic goat genotypes being used in Kenya and Uganda to boost productivity. However, the introgressive hybridization might have some evolutionary effects on animals' traits such as fitness, behaviour, reproduction, and immunology. Yet, the hybridization pattern of goats in the two nations is unknown. In addition, goats are found in different geographical localities where they are exposed to different climatic conditions, diets, diseases, and breeding practices. These lead to unique alleles that are adaptable to those particular environments due to natural selection but their effects on the current low productivity of goats is not known. Finally, the essential information of different genetic parameters that is required for implementation of goat genetic improvement programs such as RoHs is limited in most goat populations in Kenya. Furthermore, the effects of ecological zones on the goats' genetic make-up, especially on mutational load, and the extent of genetic erosion between the goat genotypes in Kenya are unknown. These hinders the implementation of effective

genetic improvement and conservation breeding programs to boost productivity in many livestock populations.

1.3 Objectives

1.3.1. Overall Objective

The overall objective of this study was to contribute toward increased goat productivity and sustainability of goat genetic improvement programs in Kenya and Uganda through joint genetic analysis using genome-wide SNP data.

1.3.2 Specific Objectives

i. To genetically characterize hybridization between different goat genotypes from Kenya and Uganda
ii. To identify the mutational load and deleterious mutations in goats from Kenya
iii. To estimate runs of homozygosity and effective population size and predict their future trends in goats from Kenya.

1.4 Research Questions

i. What are the characteristics of hybridization between different goat genotypes from Kenya and Uganda?
ii. What are the mutational load and deleterious mutations of goats from Kenya?
iii. What are the runs of homozygosity and effective population size, and the predicted future trends of goats from Kenya?

1.5 Justification of the study

Information about the genetic structure and admixture levels of goat genotypes from East Africa has enabled identification of a genetic pool for unique genes, which are adaptable to all environments and are of consumer and farmer interest. The findings further help to facilitate the development of genetic improvement programs targeting the producer to ensure high productivity to meet the continually growing human demand for meat and meat products. The acquired knowledge helps in the conservation and utilization of unique genes to ensure longer-term survival in the region. Finally, different stakeholders have benefitted with the results in policy development and setting up genetic improvement and conservation programs. This helps to ensure sustainable utilization and maintenance of genetic diversity for goats in the East African community.

1.6 Limitation of the study

This study was limited to a few genotypes from a few East African countries (Kenya and Uganda) due to resource constraints and unavailability of SNP data, therefore the results might not be generalized for all goat genotypes in the East African community.

CHAPTER TWO

LITERATURE REVIEW

2.1 Goat production in East African community

2.1.1 Introduction to goat production in Africa

It is reported that domesticated goats (*Capra-hircus*) have some genes of wild Bezoar goats (*Capra- aegarus*) and was first discovered in the Near East of the fertile crescent and South-East Asia around 10,000 – 11,000 years ago (Pereira *et al.,* 2009). Human migration and commercial networks have facilitated the worldwide spread of goats due to their adaptability and ease of transportation. East African community is estimated to have over 98 million goats (FAOSTAT, 2022) distributed across the seven member countries which include Tanzania, Rwanda, Kenya, Uganda, South Sudan, Congo and Burundi. The highest number of goats in the community are found in Kenya and the lowest in Congo (Table 1).

Table 1. Distribution of goats in East African Community from 2018 - 2022

Nation	2018	2019	2020	2021	2022
Burundi	3,249,827	3,227,903	3,365,897	3,281,209	3,291,670
D. R. Congo	4, 105,026	4,111,782	4,118,849	4, 125,764	4, 125,935
Kenya	26,710,775	34,998,238	36,021,177	32,570,314	34,529,910
Rwanda	2,731,795	2,279,812	1,745,806	1,705,054	1,513,140
S. Sudan	16,990,626	17,330,400	17,590,326	14,412,722	16,444,483
Uganda	16,491,153	15,920,638	16,421,328	16,667,994	16,881,480
Tanzania	19,798,444	20,223,004	20,930,155	21,538,321	22,146,488
EAC Totals	**90,077,646**	**98,091,777**	**100,193,538**	**94,301,378**	**98,933,106**

These figures vary both within and between nations over the years due to variations of computations which is a combination of official, imputations or estimated figures (FAOSTAT, 2022). Table 1, shows that more goats are found in Kenya seconded by Tanzania which may be attributed to the involvement of different stakeholders that aims at promoting production and productivity of goats in the two nations using different approaches (Ahuya *et al.,* 2005; Chenyambuga & Lekule, 2014). The lowest number of goats in Rwanda and Burundi are expected to increase due to increase in human population creating high demand for goat meat (Manirakiza

et al., 2020) and also the increase in human population contributes to decline in space for farming hence facilitate a shift from large ruminants to small ruminants. On the other hand, there has been a decrease in goat population in the region from 2020 to 2021 attributed to high decreases in the goat populations from Kenya and Sudan. The management systems under which the goats are reared vary across the nations from semi-intensive to extensive systems, with little or no attention from farmers (Chenyambuga & Lekule, 2014). In some cases, goats are tethered during the farming season with little or no feed supplementation and are allowed to scavenge in fields after harvesting. Despite the exposure of goats to the diverse management system, goats have a good adaptation mechanism to survive and reproduce in their environment (Gaughan *et al.*, 2019; Manirakiza *et al.*, 2020).

2.1.2 Goat breeds and breed preferences in East Africa

According to (Polak *et al.*, 2022), out of the 576 goat breeds found throughout the world, Africa is home to 96 breeds. The most commonly known indigenous breeds found in East Africa include the Small East Africa, Mubende, Kingezi, Galla, and their crosses while some breeds are known by local names based on their community (Muigai *et al.*, 2017; Onzima *et al.*, 2018b). The local goats are characterized by small body size, low carcass, and growth rate hence the exotic breeds with superior genetic materials (like boer goats) are used to improve the level of production in the local goat population through crossbreeding (Onzima *et al.*, 2018b). However, most smallholder farmers prefer indigenous breeds or their crosses to exotic breeds due to the poor survival rate of offspring and the high management costs of exotic breeds (Chenyambuga & Lekule, 2014). The indigenous goats have genetic materials that favour them to survive well under diverse management systems (Mdladla *et al.*, 2017). Therefore, they can do well with little or no production costs in terms of housing, feeding, and frequent veterinary services that are required in the management of exotic breeds.

In Africa, farmers keep goats for different purposes. Bett *et al.* (2009) found out that dairy goat farmers in Kenya keep goats mainly for milk production and selling of breeding stock to fellow farmers. This objective makes them to secure and breed for goat that have good traits of milk yield, fertility, size of the body, resistance to diseases and growth rate (Bett *et al.*, 2009). Similary, in Uganda Onzima *et al.* (2018b) discovered that most farmers keep different types of goat breeds mainly for sale therefore breeding is mostly based on adaptation traits, production

traits and ease of marketing. This shows that the purpose of keeping goats by most farmers determines the type of breed a farmer prefers to keep.

2.1.3 Breeding practices of goats in East Africa

Most goat populations in Africa are raised under extensive and semi-extensive production systems with little or no human involvement (Mdladla *et al.,* 2017). Under extensive management systems, doe and buck flocks are mixed randomly, creating the possibility of random mating where mating ratio and selective mating is not observed. This type of mating system alters the genome's shape in both short and long time due to a random combination of alleles which lead to dilution of important or possibly risking loss of important goat genetic resource (Yakubu *et al.,* 2010). Under natural random mating, the inbreeding rate becomes high, coupled with disease spread and high genetic erosion (Chenyambuga & Lekule, 2014). In such management systems, the implementation of genetic improvement programs becomes difficult. Where selection is applied, farmers usually follow the positive assortative mating method where goats are allowed to mate based on phenotypic characteristics such as large body size in bucks, twinning ability, or milk yield potential (Zergaw *et al.,* 2016). In addition, some potential farmers do castrate their male kids during weanning as one way of controlling inbreeding

2.1.4 Genetic diversity of goats in Africa

Genetic variation within and between livestock populations is paramount to ensure the continuation of genetic improvement. Determining the level of genetic diversity in a population using molecular markers provides accurate results for the selection of breeding stock. Evaluation of genetic diversity in Sub-Saharan goat populations has been done using microsatellite markers and SNPs. According to Chenyambuga *et al.* (2002), indigenous goat breeds in Africa are divided into five major categories based on the wide geographical location, climatic differences, topography, and natural pressure that affects their genetic make-up. These goat types include long lop-eared goats of North-east Africa, the Small short-eared goats (East & part of southern Africa), the Dwarf short-eared goats (coast west and central Africa), the Southern lop-eared goats (southern Africa), the Sahelian and Intermediate goats found in the Savannah belt. Within these categories, there are some indigenous goats bearing local names according to their locations such as the Nubian, Taggar, Nilotic, and Desert Sudanese goats (Rahmatalla *et al.,* 2017). Genetic diversity in East African countries of Tanzania, South Sudan, Uganda, and Kenya shows different goat genotypes with limited genetic variability within individuals, populations, and between

populations (Chenyambuga *et al.,* 2002; Onzima *et al.,* 2018a; Rahmatalla *et al.,* 2017; 2018; Waineina *et al.,* 2021). Improved genetic diversity in East African countries' goat populations provides a potential for genetic improvement through breed selection.

2.1.5 Productive and reproductive performance of goats

Inorder to discover the deficiency in the entire goat flock, it is important to monitor the productive perfomance by evaluating key perfomance indicators and reproductive parameters. Some of the productive traits include milk yield, weaning weight, birth weight, daily weight gain and weaning age, while the reproductive parameters can include age at puberty, age at first kidding, kidding interval and litter size (Bilal *et al.,* 2018; Raheem *et al.,* 2024). These production and reproduction indicators are interconnected. Due to differences in breeding, management systems and ecological region of goats there are also differences in the general perfomances of goats. The performance of goats depends on various factors which include the genetic material of an animal, environment such as nutrition, the prevalence of disease and parasites, e.t.c and their interactions (Baloch, 2014; Khandoker *et al.,* 2018). Normally, goats are known to be low producers compared to other livestock species like cattle. However, under good management of feeding, breeding, and disease control goats are able to express their genetic potential and produce more compared to other goats under little or no management. For instance, it is reported that under semi-intensive management systems, goats reach the puberty stage early, have their first kid at a younger age, and have lower kidding intervals compared to those raised under the extensive system (Abraham *et al.,* 2019; Ndeke *et al.,* 2015). However, it is reported that the management system does not affect some reproductive traits like litter size as observed in the Begait goat of Ethiopia (Abraham *et al.,* 2019). Similarly, production performances such as growth rate and milk yield increase with improvements in production management such as feed availability. Animals that have access to quality feed throughout attain good health and have a good growth rate than other animals with limited feed.

2.2 Effect of the ecological zone on the gene pool

Goats are generally adapted to a variety of geographical locations. However, their genome of the same varies according to the population’s demographic history (Peripolli *et al.,* 2017). Visser *et al.* (2016) verified that different selection strategies and topography influence the genetic make-up of animals that are separated from each other and those that are in close contact as observed in the Angora goat population of Argentina, France, and South Africa. Goats in Kenya and Uganda

are found in different agro-ecological zones and are exposed to various conditions in terms of feeding, breeding, climate, pests, and diseases. Variations of these conditions create unique animals with special alleles adaptable to that particular region due to natural selection (Ngeno *et al.,* 2015; Visser *et al.,* 2016).

Geographical barriers like mountains hinder gene flow between different animal populations, creating non-overlapping distinct clusters for a given population (Michailidou *et al.,* 2019). This can lead to inbreeding in small populations unless other ways of gene transfer between highly isolated populations are implemented such as physical mating of new breeding stock or the use of reproductive technologies in the population. Where there are no geographical barriers, continuous gene flow between populations facilitated by the management system and people's movement helps animal populations to be genetically diverse (Visser *et al.,* 2016).

2.3 Genetic improvement strategies in livestock

The productivity of any livestock population is affected by different factors which may include but not limited to animal health, nutrition as well as breeding and genetics (Lamy *et al.,* 2012). Genetic improvement is more essential in a population where the genetic potential is limited. This is achieved through selection of individuals with improved genetic potential to safeguard effective and sustainable livestock production. A number of strategies have been recommended to ensure overall improvement of a population and these include; within breed selection, cross breeding and breed replacement (Peacock *et al.,* 2011).

2.3.1 Within breed selection

Genetic improvement through within-breed selection is used by many farmers to ensure the improvement and conservation of desirable genetic resources. The availability of genetic variation within the breed population facilitates the effective selection of high-producing animals to be used as parent stock within the breed. For effective implementation and sustainability of within-breed selection focusing on smallholder farmers, a bottom-up approach with full involvement of farmers and local institutions during planning and implementation is paramount (Peacock *et al.,* 2011). Farmers need to be given support in terms of feeding and health care for their animals. The approach of within-breed selection is more similar to the Community-Based Breeding program (CBBP) approach where smallholder farmers who have a common interest in livestock production work together to improve the genetic materials within their geographical

location (Mueller *et al.,* 2015). However, this approach applies both within and between breeds. The parent stock (bucks) can be reared at government test stations or by the community's trusted members. Sometimes, animals with desirable superior traits are identified among farmer's flocks to produce breeding males for the community. In most developing countries, this program is reported to be profitable leading to an improvement in the livelihood of goat farmers (Haile *et al.*, 2020; Karnuah *et al.,* 2018).

2.3.2 Cross breeding strategy

Genetic improvement programs for goats through crossbreeding are being implemented by both government and non-governmental organizations in Africa using either local or exotic genetic superior materials. This strategy is aimed at poverty reduction, improvement in nutritional health, and ensuring increased meat and milk productivity of indigenous goats while maintaining the population's genetic diversity (Muigai *et al.,* 2017). British Toggenburg dairy goat has been used for crossbreeding indigenous goats in Kenya's eastern highlands through the Meru dairy goat and animal health care project (Peacock & Hastings, 2011). In addition, the Galla goat in Kenya is used to improve the body weight of the Small East African breed while the Boer goats are commonly used to improve indigenous growth rate and disease resistance in most African countries (Onzima *et al.,* 2018b; Manirakiza *et al.,* 2020). A sustainable and effective crossbreeding program is supposed to be well planned, monitored, and controlled to avoid inbreeding and dilution of well-adapted genetic resources in the population.

2.3.3 Breed replacement

In most African goat populations, replacement occurs when the existing population has limited production levels. This breeding strategy involves the introduction of commercial breeding male or female animals of a different breed to the flock to improve its genetic potential. For instance, Boer bucks in many African nations have been used to replace local bucks to improve local meat productivity where the replacement source can either be from the own flock, outside the flock, or both (Laouadi *et al.,* 2018). However, for the replacement to be effective, an assessment must be done on the flock's current performance, environmental restrictions, type of genetics required, and replacement stock source since the existing genotypes might be enough to improve the flock with good management. Furthermore, to avoid the loss of local and unique genes in the population, the use of high-producing animals must be controlled and strategic to ensure sustainable genetic diversity.

2.4 Methods of genetic characterisation

Genetic characterization of any livestock species is essential in planning both short- and long-term strategies for best utilization, conservation, and improvement (Msanga *et al.*, 2012). This is done either by following the phenotypic or molecular method. Currently different genetic characterization results have been reported on various livestock species populations including goats (Acaitum, 2018, Waineina *et al.*, 2021). Even though the methods can be done separately, it is recommended to complement the phenotypic with molecular characterization for effective results.

2.4.1 Phenotypic characterisation

This type of characterization is based on phenotypic measurements of both qualitative and quantitative parameters whose information gathered is useful for planning purposes at different levels (local, national, or international) (FAO, 2011; Ramukhithi *et al.*, 2019). Well-designed surveys are used to collect information on different phenotypic parameters such as physical characteristics or performance levels, which are essential for characterization. This is done by following the recommended FAO guidelines for conducting phenotypic characterization to gather information on the existing differences in morphometric features of a population (FAO, 2011). Phenotypic characterization of goat populations has been reported in different African nations of Tanzania (Nguluma *et al.*, 2016), Uganda (Acaitum, 2018), and South Africa (Ramukhithi *et al.*, 2019). In all the cases, results are used as a reference in strategic planning to maintain or improve genetic diversity, avoid genetic erosion and ensure genetic improvement in indigenous goat species.

2.4.2 Molecular characterisation

Molecular characterization complements phenotypic evaluation when carrying out genetic diversity studies. This method shows the genetic basis of phenotypes, inheritance patterns from one generation to another, within breed genetic structure, level of variability, and the relationship between breeds (Scherf *et al.*, 2015). To conduct a molecular evaluation for a given population, researchers must follow FAO guidelines to have the best biological sample, well-represented to give out consistent data to be easily included in international scale analysis.

Molecular evaluation has become more relevant due to the development of whole-genome sequences for different livestock species, technologies for measuring polymorphism at loci spread across the genome, and technologies for measuring gene transcription and large expression scale

(FAO, 2011). Many biotechnological tools are currently accessible for assaying DNA and some of the molecular markers available for molecular analysis include Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), Single Nucleotides Polymorphism (SNP), and Microsatellite (SSR). However, the most commonly used tool in analysing genetic relationships within and between populations is the SNPs (FAO, 2011; Weinman *et al.,* 2015).

Single Nucleotide Polymorphism (SNP) refers to a single base change in a DNA sequence (Vignal *et al.,* 2002). The DNA sequence is made up of four base nucleotides A, T, C, and G, which exist as a double-strand and are always complementary to each other; A&T; G&C. The SNPs are identified from the raw sequence data produced by the next generation sequence platforms whose identification procedure follows systematic steps using different bioinformatics tools (Ngeno, 2018). The steps as reported by Ngeno. (2018) involve base calling, quality control trimming, read mapping or alignment, alignment post-processing, base quality recalibration, SNP calling, and filtering of the SNP to improve SNP calls' quality. A SNP can be in the form of transition or transversion, widely distributed in the genome, co-dominant, very specific, and can be inherited (Koopaee & Koshkoiyeh, 2014). They are useful in detecting internal sequence variations of an individual, which become applicable in different study areas like genome-wide association studies.

2.5 Runs of Homozygosity

The mating of related individuals in a population produces progeny with identical copies of genetic material (Curik *et al.*, 2014). Runs of homozygosity (RoHs) are defined as contiguous homozygous stretches in an individual genome due to the transmission of identical haplotypes from parents to offspring (Ceballos *et al.,* 2018). Detection of RoHs is essential in improving the existing mating system to reduce inbreeding levels. Accurate identification of RoHs depends on the genotyping quality, the minimum size of runs of homozygosity, a permitted number of heterozygotes, and the type of chip used. According to (Zhang *et al.* (2015), RoHs shorter than 5.0Mb is accurately detected when using chips with densities greater than 50000 SNPs.

Accumulation of RoHs in a genome provides information about the population's genetic events such as the demographic evolution of a population over time and its relatedness (Sölkner *et al.,* 2010). Different studies have detected RoHs in various livestock species populations describing different parameters within or among the species populations (Bosse *et al.,* 2012;

Nandolo *et al.,* 2019). The size, frequency, and distribution of RoHs in livestock depend on different factors working together or in isolation like the number of generations in which inbreeding has occurred, effective population size, and recombinant rate among other factors (Curik *et al.,* 2014).

The development of high-density SNP arrays to scan RoHs in the genome provides an effective method of identifying Identical by Descent (IBD) haplotypes. Different researchers apply different methodologies to predict RoHs in a genome since there is no known standard method. For that reason, it is currently difficult to compare RoHs results from different studies. The first study on RoHs in livestock was done in cattle to identify RoHs and their relationship with demographic history and inbreeding (Sölkner *et al.,* 2010). Follow-up studies of RoHs were done on different species of goats and pigs, among others (Bertolini *et al.,* 2018; Purfield, *et al.,* 2017). Currently, more research on RoHs is being called for domesticated species like goats, sheep, and horses due to limited information.

2.5.1 Measurements of Inbreeding using Runs of Homozygosity

According to Nandolo *et al.* (2019), inbreeding levels of African goats are diverse with many goat populations having low levels of inbreeding than expected. The development of genomic technology has ensured the availability of molecular information creating the possibility of estimating inbreeding levels. Runs of Homozygosity (RoHs) is used to estimate inbreeding levels within individuals, in a population, and between populations rather than the use of pedigree information (F_{PED}), (Peripolli *et al.,* 2017). The F_{RoHs} is the best tool for estimating inbreeding due to its greater accuracy and its ability to disclose the inbreeding age based on RoHs length (Peripolli *et al.,* 2017; Rebelato & Caetano, 2018). Comparison studies to estimate inbreeding using F_{PED} and F_{RoHs} have been reported in human and cattle populations with accurate results obtained from RoHs estimates (Marras *et al.,* 2015; McQuillan *et al.,* 2008). The F_{PED} is associated with errors due to failure in capturing the actual proportion of the genome that is identical by descent (Curik *et al.,* 2014). When calculating inbreeding coefficients using F_{PED}, absent individuals in the population are considered unrelated, which means they do not consider the stochastic nature of inheritance resulting from a finite number of chromosomes and the small number of recombination events during meiosis (Curik *et al.,* 2014).

2.6 Effective population size

Effective population (Ne) size refers to the size of an idealized population that undergoes the same genetic drift rate and inbreeding as the actual population under study (Falconer & Mackay, 1996). The Ne is an important genetic parameter whose output information is useful in genomic selection and Genome-Wide Association Studies in various livestock species. According to Barbato *et al.* (2015), different factors affect the population's effective population sizes such as changes in population size, differences in sex ratio, and differences in reproductive success (number of offspring). The Ne can be estimated from either pedigree information, demographic parameters, or genetic markers from a randomly sampled population (Corbin *et al.,* 2012).

Marker-based method of estimating effective population size is done based on the relationship between Linkage disequilibrium (LD), recombination rate, and population size (Wang *et al.,* 2016). Linkage disequilibrium has been used in different livestock species include cattle and goats to estimate effective population size (Fabri *et al.,* 2020; Flury *et al.,* 2010). This is the non-random association of alleles at different loci in a population and the extent of LD is due to evolutionary history and the effective population size (Deng *et al.,* 2019). According to Corbin *et al.* (2012), the rate of decrease in LD as a function of marker distance measured as a squared correlation coefficient (r^2) is used to estimate effective population size and how it changes with time. Small effective population size is associated with low genetic diversity; therefore, the population becomes vulnerable to a low adaptation rate to evolution and accumulation of deleterious mutations through genetic drift.

2.7 Mutation

Mutations are changes in Deoxyribonucleic Acid (DNA) sequence that causes genetic diversity between organisms. Mutation effects can be beneficial, harmful, or neutral depending on the context or location. In every generation, new mutations will arise which can affect the animal's fitness and phenotype. Agrawal and Whitlock, (2012) defined mutation load as the loss of an animal's fitness due to deleterious alleles maintained by mutation-selection balance. According to Charlesworth (2012), adaptive evolution is due to the mutations that increase animals' fitness since selection is based on such variants. Improvement in genomic technology facilitates the identification, quantification, and segregation of deleterious variants in a population. Effective estimation of mutation load in different species is however limited by finding mutation-free reference genotypes (Agrawal & Whitlock, 2012).

Some mutations can persist in a population for a long time before being eliminated through selection. Regardless of the mutation levels in a population, the presence of deleterious mutations can cause a reduction in population size and possible species extinction if not properly managed (Agrawal & Whitlock, 2012). The presence of long homozygous regions in inbred animals helps to assess deleterious mutations in a given population in addition to the detection of functional variants during Genome-Wide Association studies for complex traits (Szpiech *et al.*, 2013). High coverage of Runs of homozygosity has a high fraction of deleterious mutations that occur in long runs of homozygosity. According to Szpiech *et al.* (2013) and Zhang *et al.* (2015), there is a strong direct correlation between the frequency of deleterious homozygous mutations and genomic RoHs regions. The accumulation of recessive harmful mutations in the genome due to inbreeding causes a reduction in progeny fitness called inbreeding depression.

2.8 Functional annotation

Functional annotation is the process where biological information is attached to a given DNA sequence of a genome (Abril *et al.,* 2019). It is divided into molecular function, biological function, and cellular components in non-overlapping ontologies (Milano, 2019). The animal performance or behaviour is due to the genome it carries hence sequencing alone is not complete until the sequences are interpreted. Therefore, sequencing technologies must be followed with an update of genome annotation to make it complete. Different genome steps are used in the annotation that shares the same features and researchers have grouped them into two general phases: computational and annotation (Abril *et al.,* 2019). The computational stage involves sequence alignment where the generation of an evidence-driven gene is predicted while an annotation is synthesized into gene annotation (Yandell & Ence, 2012). The existing annotation protocols focus only on protein-coding genes except for ENSEMBL which has the potential to annotate non-coding RNAs.

For accurate annotation, genome assembly must be complete by meeting the minimum required standards described using different parameters like; N50, the average gap size of a scaffold, and the average number of gaps per scaffold (Humann *et al.,* 2019). However, at least 90% of completeness is required for annotation. According to Yandell & Ence (2012), the accuracy of the gene annotation is difficult to quantify since most species sequenced today do not have reference data set. Therefore, the only solution is to quantify the proportion of annotations that encodes a protein with known domains using tools used for such analysis.

CHAPTER THREE

GENETIC ANALYSIS OF HYBRIDIZATION PATTERN IN GOAT GENOTYPES FROM KENYA AND UGANDA

Abstract

East African community is home to 98,933,106 goats kept in different geographical regions and are exposed to diverse climatic, production, and management conditions. This has shaped the goat genome due to selection. This study investigated the hybridization pattern between distantly isolated goat populations from Kenya and Uganda using Single Nucleotide Polymorphism (SNP) data, genotyped using a 50K goat SNP chip. In Uganda, data for 144 goats representing six genotypes sampled from five agro-ecological zones were retrieved from the Zenodo genome archive. In Kenya, 94 goats representing four genotypes sampled from three regions were used. DNA was extracted and genotyping was done. Quality control procedures were performed in PLINK v 1.9. Data from Kenya (48,303 SNPs) and Uganda (46,105 SNPs) was then merged in Tassel software resulting in 94,408 SNPs available for joint downward analysis. Principle component and phylogenetic analyses were used to visualize relationships between the studied populations. All genotypes were separated into two main groups according to country of origin. Four well-defined clusters, two from each country were observed with some visible outliers. From the two clusters in each country, genetically admixed genotypes were also observed in one cluster. This might suggest the mating of individuals between the genotypes due to the proximity between them. The degree of genetic differentiation measured using Fst ranged from 0.191 to 0.324, indicating that all the genotypes within a country are to some extent isolated from each other. Only Boer genotype from Uganda was highly isolated from all genotypes in this study. The diversity shown between Kenya and Uganda goat genotypes might be due to different ancestry or isolation by distance. These results are useful in the implementation of future genetic conservation, utilization, and improvement programs in Kenya, Uganda and the rest of African community.

3.1 Introduction

Worldwide, goats are found in different geographical regions where they are exposed to various climatic, production, and management conditions mostly kept under an extensive system (Mdladla *et al.,* 2018). Over eons, goats have evolved and become adapted to their specific environment with unique alleles due to natural selection (Michailidou *et al.,* 2019; Ngeno *et al.,* 2015; Visser *et al.,* 2016). In Kenya and Uganda, goat population is estimated at 34,529,910 and 16,881,480 respectively (FAOSTAT, 2022). According to Muingai *et al.* (2017), goats are grouped into three major categories in Africa which include lop-eared, short-eared twisted horn, and short-eared small-horned and are commonly identified by local names.

Due to limited productivity of goats in Kenya and Uganda, different stakeholders have been working on genetic improvement programs for indigenous goats using exotic or local breeds from within or distantly related populations. For instance, Boer and Galla breeds are commonly used in meat production improvement while Toggenburg and Saanen are used in the improvement of milk production (Ahuya *et al.,* 2005; Onzima *et al.,* 2014). However, the grazing system in most goat populations is mainly a communal that unavoidably leads to uncontrolled natural mating among the populations (Otieno *et al.,* 2015; Semakula *et al.,* 2010). A communal grazing system of management facilitates hybridization between different livestock populations.

Hybridisation refers to the mating of distantly related individuals which are different in more than one trait that is heritable (Harrison, 1993). This mating will produce F_1 hybrid that have different genotype and consequently change in the phenotype compared to both parents (Liu, 2010). Despite that hybridisation can lead to genetic extinction in smaller populations, it is beneficial in that the offsprings becomes more strong, reproductive, high immunity and increase in longevity than the parents (Yadav *et al.,* 2019). According to Chan *et al.* (2019), the known risks of hybridisation are overstated hence there is under-utilisation of hybridization because from the few studies where hybridisation was applied there have been recommendable positive outcomes.

Genetic hybridization in African goat populations have been reported using Single Nucleotide Polymorphism (Onzima *et al.,* 2018a; Waineina *et al.,* 2021) and microsatellites (Nguluma *et al.,* 2018). These studies, however, focused on distantly isolated goat populations from within their nations. However, there might be a possibility of gene flow between nations catalysed by human migration, trade, or advancement in reproductive technologies. Due to the

isolation by distance and possible differences in environmental conditions between Kenya and Uganda, a joint genetic analysis of goat populations was conducted. This study aimed at investigating the hybridization and gene introgression pattern between Kenyan and Ugandan goat populations using Single Nucleotide Polymorphism data genotyped using a 50K SNP chip.

3.2 Materials and methods

3.2.1 Study sites and Sampling

The study was done using Single Nucleotide Polymorphism (SNP) datasets from Kenya and Uganda. In Uganda, SNP data for 144 goats from six goat genotypes sampled from five agro-ecological zones were retrieved from the Zenodo genome archive (Onzima *et al.*, 2018). The zones sampled in Uganda were Mubende district (mid-altitude farmlands and central wooded savannah), Kabale and Kisoro districts (South-western highlands), Arua district (short savannah grasslands), Moroto district (north-eastern semi-arid region), Sironko (Eastern highlands), Ssembabule and Kabarole districts (mid-altitude zone).

In Kenya, Blood samples were collected from a total of 94 goats from four goat genotypes obtained from 53 farms and one government breeding station. Being part of the Kenyan Climate Smart Agriculture Project (KCSAP), the goats were purposively selected in different ecological zones of Kenya, namely; Nyeri (Mukurweini Sub-County), Meru (Central Imenti Sub-County) and Homa Bay (Homa Bay town) located in the Central (wet-dry), Eastern (wet) and Western regions (wet area) respectively. The selected areas were some of the entry points of exotic breeds in the country. The goat genotypes that were investigated included; Saanen (n = 24), Alpine (n = 29) and Toggenburg (n = 31) sampled from members of goat farmer associations across the selected Counties. Galla (n = 12) was sampled from Naivasha, Sheep and Goat government station. The number of goats varied between the breeds and within the sampled households which led to variations in sample size across the genotypes. Blood samples were collected at each selected farm. A member with two does only one doe was used and where there were more, the relationship of the does was confirmed by the farmer to avoid selecting full and half siblings. For Galla goats, pedigree information was used to ensure sampling of unrelated goats.

3.2.2 Sample collection, DNA extraction and genotyping

Blood samples were collected at each selected household by a qualified veterinary officer. The animals were constrained during blood collection and all FAO protocols for sampling of blood for DNA were observed. The blood samples were collected into Ethylenediaminetetraacetic acid

(EDTA) tubes from the Jugular vein and stored at -20^{O}C for two months before genomic extraction. Blood sample duplicates were also collected and kept separately.

DNA extraction was done using the Qiagen DNeasy Blood and Tissue Kits. Purified DNA quality and quantity were validated using the Qubit dsDNA BR (Broad-Range) Assay Kit on the Qubit 2.0 and Nanodrop Spectrophotometer (Nanodrop ND-1000). DNA extraction and quality control checks were done at Kenta Agricultural and livestock Research Organisation (KALRO) biotechnology laboratory while Genotyping was outsourced (Illumina, Inc.San Diego, CA92122 USA) and were conducted using the Illumina goat SNP50 Bead chip developed by International Goat Genome Consortium (IGGC).

3.2.3 Quality control

Quality control procedures were performed in PLINK v 1.9 (Chang *et al.,* 2015) using the following parameters; SNPs less than 95% call rate, less than 0.05 Minor Allele Frequency (MAF < 0.05), Hardy-Weinberg Equilibrium <0.001, and SNPs with more than 10% missing genotypes were filtered. The study protocol was approved by the Egerton University Research Ethics committee.

3.2.4 Statistical analysis

To describe hybridization and gene introgression patterns between Kenya and Uganda goat populations, the two SNP datasets were merged in Tassel software (Bradbury *et al.,* 2007). The population structure of goat genotypes were investigated using principal components analysis (PCA) and multi-dimension scaling (MDS). These approaches use the pairwise relationship between individuals to visualize the genome-wide population structures. Principal Component Analysis (PCA) were done in Tassel software by uploading the merged SNP dataset in the software and set the commands to run PCA. This gave output in an excel sheet which were visualized in R-package ggplot2 (Walter *et al.,* 2015). The PCA excel output were uploaded on the Rstudio then insert the PCA-R commands to give out graphs. Genome-wide identical-by-state distances were calculated in PLINK v 1.9 using the --cluster command together with the --mds-plot option on the uploaded merged data set. The MDS-plot for the first two components identified were visualized in R-package ggplot2 as describe above (Walter *et al.,* 2015).

The merged dataset was also then pruned for Linkage Disequilibrium to reduce data redundancy in PLINK v 1.9 using simple pairwise threshold model command; (--indep-pairwise 50 5 0.2) SNP window size: 50, 5 SNPs shifted per step and r2 threshold of 0.2 as used by (Visser

et al., 2016). Pruned data were uploaded in the admixture software 1.3.0. (Alexander *et al.*, 2015) in KENET v 0.19 to assess the population structure and admixture. The analysis was run for K values 2 to 10 under unsupervised mode since there was no information for the individual's origin. Fivefold cross-validation error (CV) values and 500 iterations were considered for each K value to reach convergence. The optimal number of clusters present within the dataset were determined by the lowest CV error of the K-value. A stacked bar plot was drawn in R-package ipADMIXTURE (Amornbunchornvej *et al.*, 2020). The Fst values obtained from the admixture analysis were used to establish the quantitative measure of genetic differentiation between the genotypes in the study.

Finally, Tassel software was used to construct a distance matrix from the pruned data and then create a neighbour-joining tree. This was done by uploading the merged Kenya and Uganda dataset on the software then set commands for analysis of relatedness and creation of an Archaeopteryx tree.

3.3 Results

3.3.1 Principal Component Analysis (PCA)

After merging the two data sets, 94,408 SNPs were available for joint downward analysis with 48,303 and 46,105 SNPs from Kenya and Uganda respectively. Principle component analysis indicated that 42.39% and 18.64% genetic variation was explained by PC1 and PC2 respectively (Figure 1). Four well-defined clusters were observed with some visible outliers. Two clusters were formed in each country and genetically mixed genotypes were identified in only one cluster. The two clusters in Kenya were comprised of Saanen (SAA), Alpine (ALP), and Toggenburg (TOT) genotypes and an independent cluster of Galla genotype only. While in Uganda, the Boer genotype had an independent cluster, and the second cluster was comprised of Sebei (SEB), Kingezi (KIG), Mubende (MUB), Small East African (SEA), and Karamojong (KAR). Few individuals were observed as outliers from Boer (BOE), ALP, SAA, KIG, and MUB genotypes. All Ugandan indigenous genotypes were separated from the remaining genotypes in the study on PC1 while PC2 separated the BOE genotype from all the genotypes in the study.

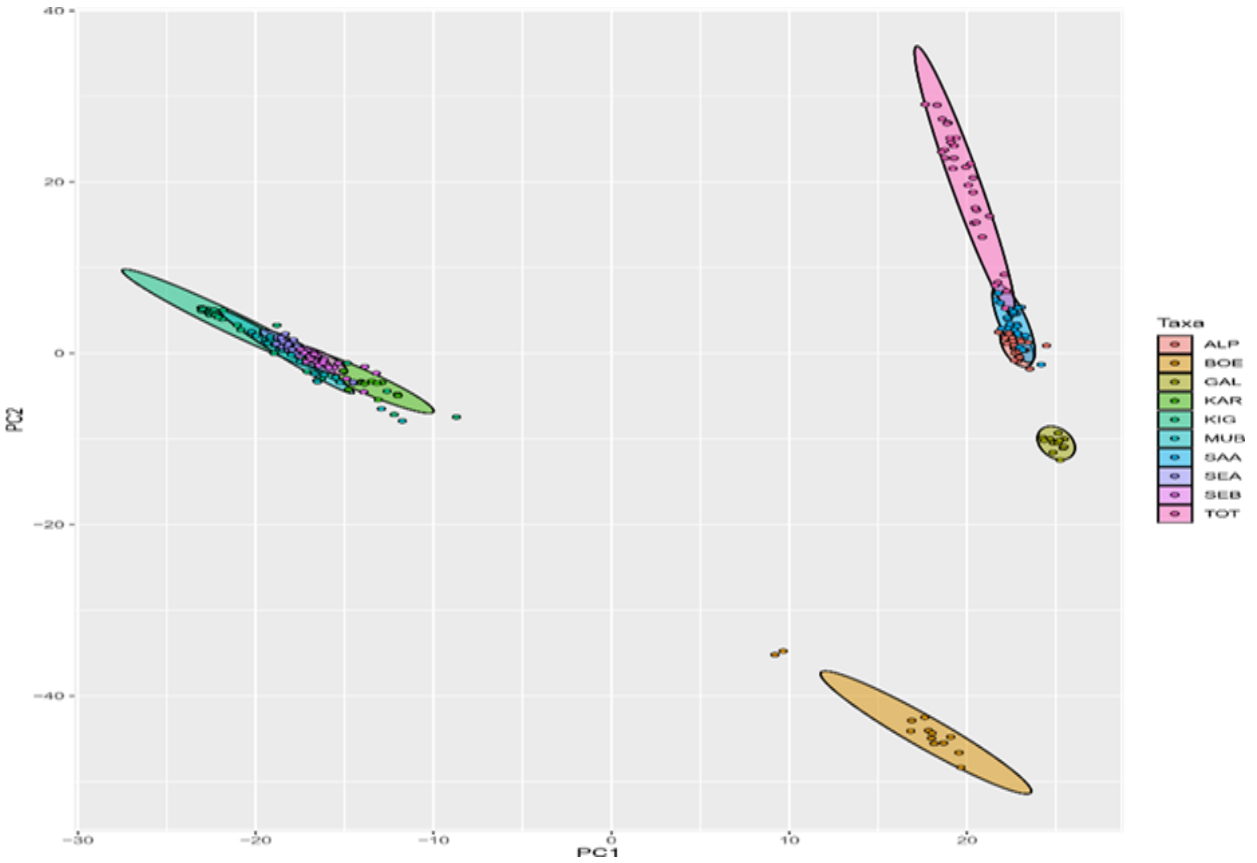


Figure 1. Principal components (PCA1 and PCA2) showing genetic relationships among 238 goats from Kenya (ALP, GAL, SAA and TOT) and Uganda (BOE, KAR, KIG, MUB, SEA and SEB).

3.3. 2 Multi-Dimensional Scaling (MDS)

The genetic relationship between the genotypes revealed in an MDS-plot separated all the genotypes according to country of origin (Figure 2). MDS-plot revealed three distinct clusters, one cluster composed of all the genotypes (GAL, SAA, ALP, and TOT) used in the study from Kenya. Two clusters were observed in the Ugandan goat population with the Boer genotype forming an independent cluster while the other cluster had mixed genotypes of SEB, KIG, MUB, SEA, and KAR. From the analysed sample, two individuals were observed as outliers from KIG and TOT.

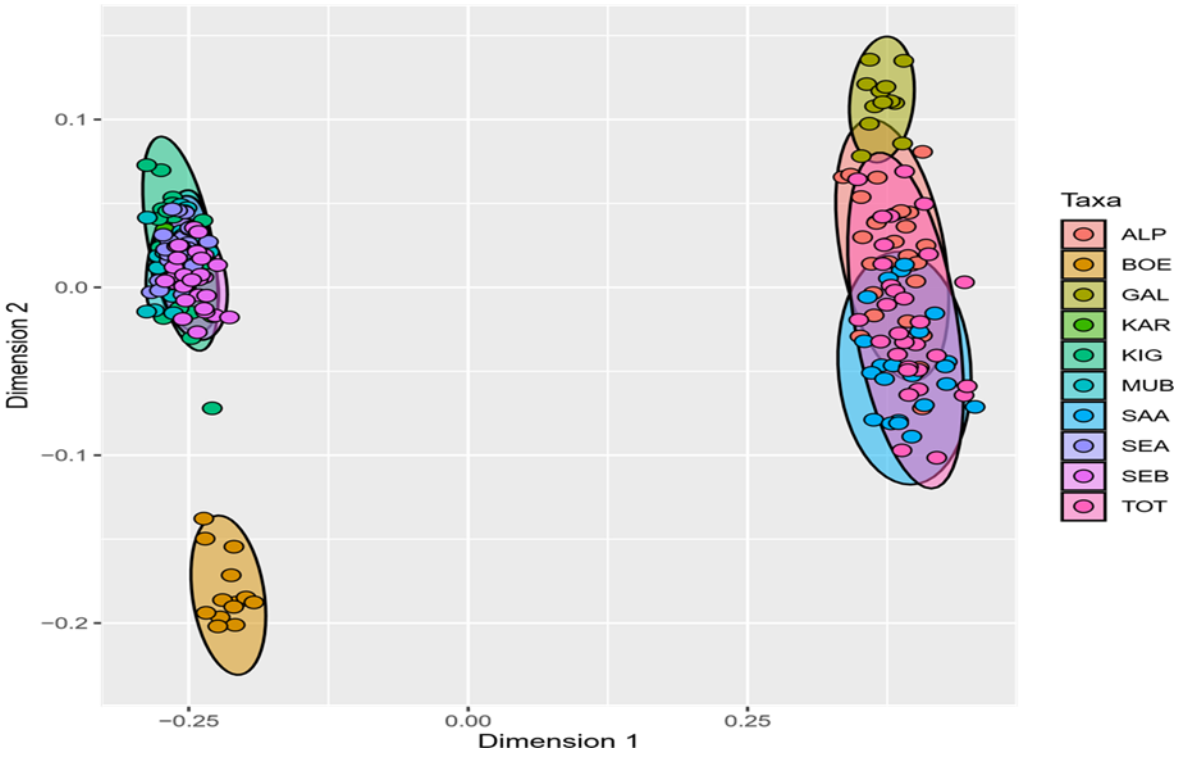


Figure 2. Multidimensional scaling plot of 238 goats of ten genotypes from Kenya (ALP, GAL, SAA and TOT) and Uganda (BOE, KAR, KIG, MUB, SEA and SEB).

3.3.3 Admixture analysis

An unsupervised admixture analysis was done to determine the degree of genetic mixing between the goat genotypes. Admixture analysis for K = 2 to K = 10 with 500 bootstrap replicates, revealed K = 4 as an optimal number of clusters present in the data set. At K = 4, differences were observed between the pure and admixed genotypes (Figure 3). Therefore, K = 4 is the number of ancestry populations having the lowest Cross Validation error of 0.62784 (Figure 4). Only Boer maintained pure genetic makeup, which is consistent with PCA, MDS, and Phylogenetic analysis results. Karamojong, KIG, MUB, SEA, and SEB are all mixed up with some Boer ancestries. Alpine and Toggenburg are almost pure but have some traces of genes from Boer. Furthermore, GAL genotypes indicated gene introgression from BOE, ALP, and TOT.

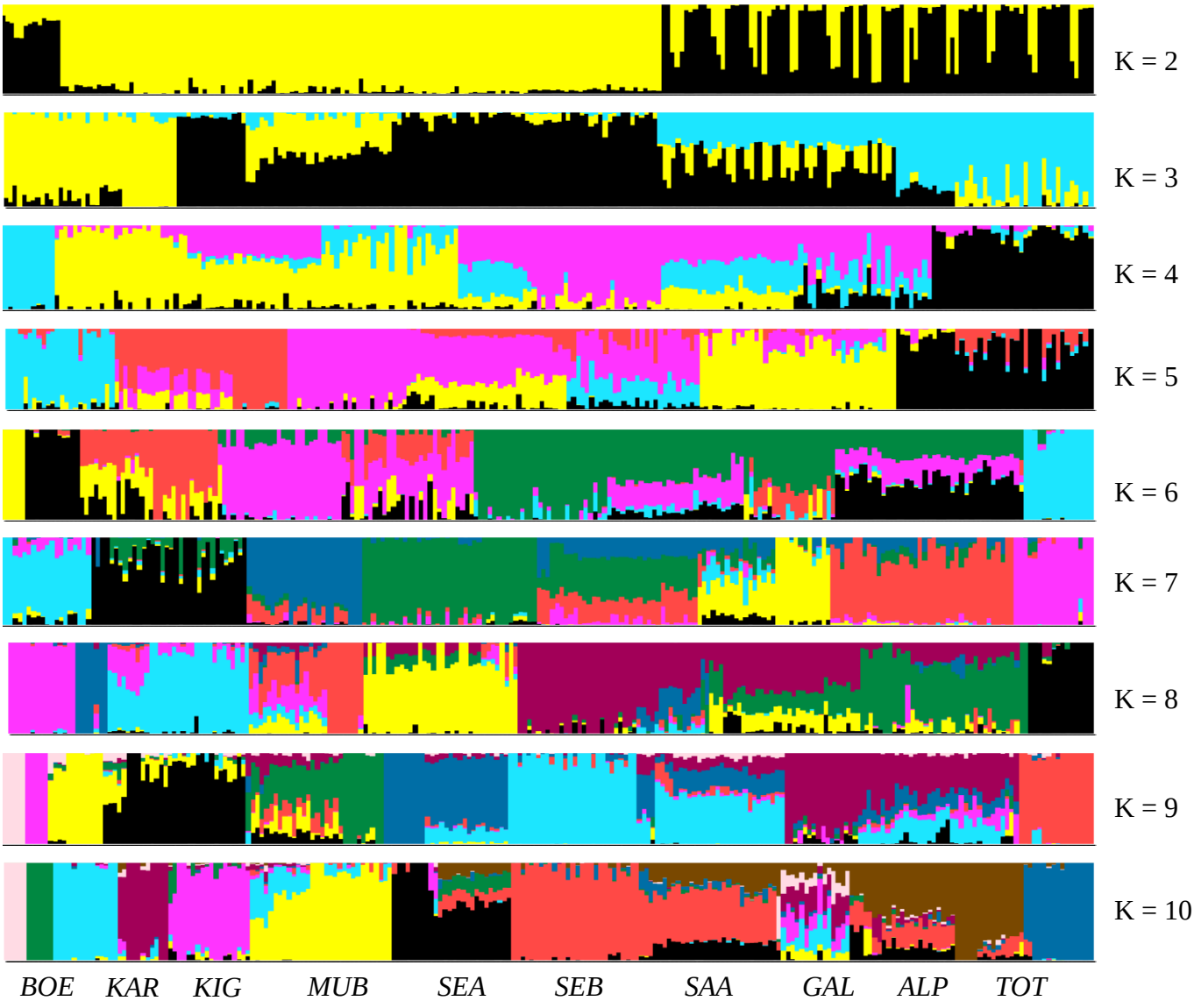


Figure 3. Population structure of ten goat genotypes for K = 2 to K = 10 indicating ancestral population per individual. Boe (Boer); KAR (Karamojong); KIG (Kigezi); MUB (Mubende); SEA (Small East African); SEB (Sebei); SAA (Saanen); GAL (Galla); ALP (Alpine); TOT (Toggenburg).

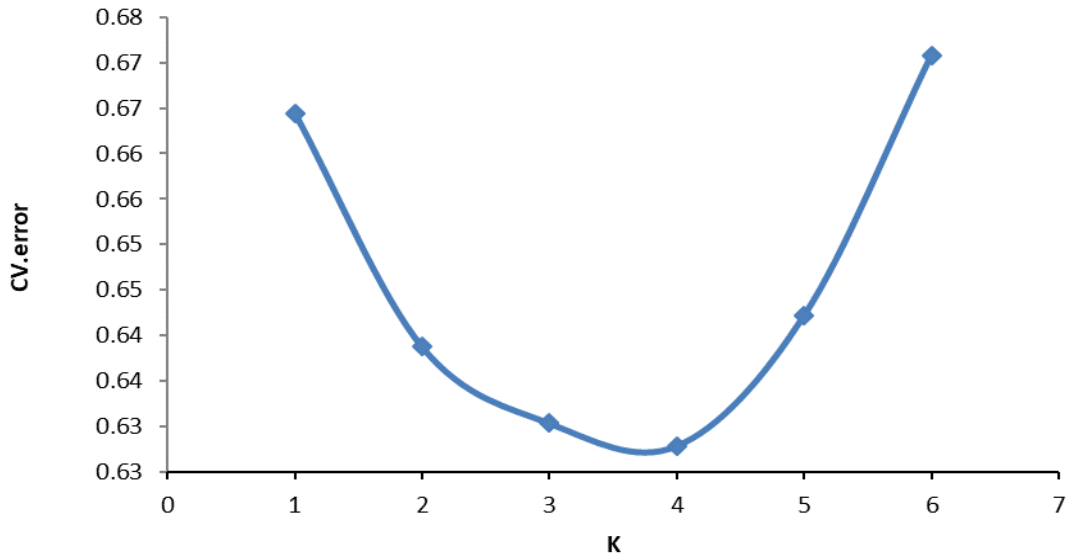


Figure 4. Cross-validation plot indicating the optimal K- value

3.3.4 Measure of genetic differentiation using the FST approach

From the Fst values estimated, all the genotypes in the study are to some extent isolated from each other indicating genetic differentiation. The values ranged from 0.191 to 0.324 between

Table 2. The FST values for the ten genotypes of Kenya and Uganda obtained from admixture analysis

	BOE	KAR	KIG	MUB	SEA	SEB	SAA	GAL	TOT
BOE									
KAR	0.222								
KIG	0.246	0.246							
MUB	0.214	0.217	0.255						
SEA	0.23	0.191	0.259	0.248					
SEB	0.251	0.278	0.298	0.272	0.284				
SAA	0.273	0.269	0.273	0.284	0.287	0.322			
GAL	0.229	0.221	0.243	0.233	0.244	0.284	0.275		
TOT	0.27	0.268	0.291	0.281	0.286	0.324	0.31	0.269	
ALP	0.209	0.206	0.238	0.217	0.222	0.261	0.253	0.208	0.241

the genotypes (Table 3). High genetic differentiation was observed between TOT and SEB (FST = 0.324) and between SAA and SEB (FST = 0.322). On the other hand, little genetic differentiation was observed between SEA and KAR genotypes (FST = 0.191).

3.3.5 Phylogenetic analysis

Genetic relationships between the 10 genotypes under study from a neighbour-joining tree are shown in Figure 5, where each genotype has been assigned a colour. Based on the genetic distances, almost all individuals were clustered together according to their genotypes. The Boer genotype was the most separated genotype from all other genotypes in the study. The indigenous genotypes of Uganda formed a distinct cluster. Galla, Toggenburg, and Alpine shared a common clade, however, the distance between these genotypes was larger compared to the distance between Ugandan indigenous genotypes. Some individuals from Alpine have been clustered together with Saanen, which might indicate gene introgression.

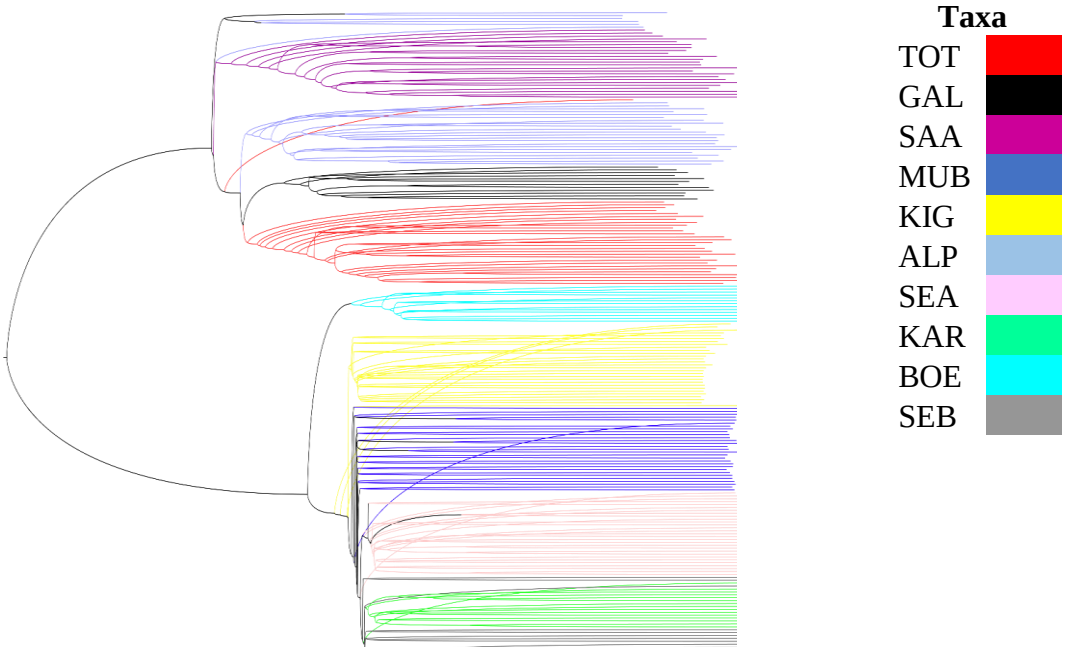


Figure 5. Neighbour-joining (NJ) tree for the Ugandan (MUB, KIG, SEA, KAR, BOE and SEB) and Kenyan (GAL, TOT, SAA and ALP) goat populations studied

3.4 Discussion

From the three methods (PCA, Admixture, and Phylogenetic analysis) used to investigate the hybridization and gene introgression patterns in goat genotypes from Kenya and Uganda,

results are in agreement that distantly isolated livestock populations have unique alleles adaptable to their isolated environments (Visser *et al.,* 2016). All livestock populations are exposed to different forces that affect the shape of their genome hence contributing to differences in phenotypes (Michailidou *et al.,* 2019; Peripolli *et al.,* 2017)

3.4.1 Between population genetic diversity

Principal Component Analysis and NJ-tree separated the genotypes into four different genetic pools. Individuals from each country were grouped into two different clusters for each country. These findings are similar to previous studies investigating the genetic variance of goat genotypes across nations that reported high genetic diversity between the distantly isolated population (Monau *et al.,* 2020; Muema *et al.,* 2009; Tarekegn *et al.,* 2019; Visser *et al.,* 2016). The extent of genetic diversity between populations depends on several factors such as differences in ancestry parents, differences in evolutionary forces of genetic drift, mutations, selective breeding, adaptation, geographical isolation, and migration or importation (Colli *et al.,* 2018; Groeneveld *et al.,* 2010). Some of the mentioned factors like geographical isolation with natural barriers and a low level of animal migration catalyzed by human activities contributed more to the observed genetic distinctiveness of goat genotypes from Kenya and Uganda. At the international level, livestock genetic materials are also exposed to movement restrictions hence affecting the rate of gene flow between the nations.

According to Muema *et al.* (2009), high levels of homozygosity are observed in un-isolated populations compared to isolated subpopulations hence the proximity of any animal population determines the extent of the genetic relationship. The genetic diversity observed between Kenya and Uganda is comparable to the joint genetic analysis done in the genetic analysis of world sheep breeds (Kijas *et al.,* 2012), domesticated pigs and wild pigs (Sprem *et al.,* 2014), and dairy cattle across Sub Saharan African countries (Opoola *et al.,* 2020).

Local genotypes like Galla and Mubende from Kenya and Uganda respectively are reported to have superior genetic materials preferred by farmers (Ahuya *et al.,* 2003; Ojango *et al.,* 2016; Onzima *et al.,* 2018), yet they are not utilized in genetic improvement programs across the nations. Similar to other exotic genotypes, traits of economic importance in local goats need to be genetically improved to be used as genetic material across nations. According to Ahuya *et al.* (2003), different research studies and indigenous knowledge confirmed that the Galla genotype possesses unique traits that can be used as a source of improved genetic material to increase milk

and meat productivity, provided good management practices are followed by farmers. This is similar to the Mubende genotype which has a large live body weight and is commonly used in breeding by farmers in Uganda.

The distinct clusters for Galla and Boer genotypes could be because they were sampled from controlled government breeding stations. These genotypes are conserved for genetic improvement programs in their nations since they have superior genes for meat and milk production (Ahuya *et al.,* 2003; Onzima *et al.,* 2014). Boer was introduced in Uganda from South Africa while Galla is a local Kenyan genotype. The high separation of the Boer genotype from other Ugandan genotypes might be due to strict control of the breeding bucks. Even though farmers are encouraged to utilize these genotypes to genetically improve their populations, the level of gene flow from Boer and Galla showed in this study is limited. This might be attributed to lack of knowledge, access to the breeding stock by the farmers, or farmers' preference. Most smallholder farmers in Africa who are knowledgeable about improved genotypes do not prefer improved breeds due to either poor survival rate of the progeny, high production costs, or due to inconsistency in the supply of replacement breeding stock (Berhanu *et al.,* 2012; Chenyambuga *et al.,* 2014).

3.4.2 Within populations genetic diversity

Genetic differentiation within the studied genotypes was moderate, consistent with studies done by Onzima *et al.* (2018) and Wainena *et al.* (2021). The genotypes within populations were not clearly separated from each other indicating evidence of gene flow between them. This is because of the close proximity between the genotypes coupled with the management history of the goats that facilitates uncontrolled breeding in communal grazing areas (Chenyambuga & Lekule, 2014; Semakula *et al.,* 2010). According to Muema *et al.* (2009), genetic relationships of goat populations in Sub-Saharan Africa using microsatellite markers were found to depend on the location proximity compared to morphological classification. Results from this study are similar to other recent studies of genetic structure analysis using microsatellite and SNP markers which displayed low to average genetic variations (Nguluma *et al.,* 2018; Rahmatalla *et al.,* 2017; Tarekegn *et al.,* 2019).

Despite the genotype clustering separately, admixture analysis revealed the optimal number of clusters to be four (K = 4), confirming the occurrence of crossbreeding between the studied genotypes as shown in the stacked bar plot (Figure 3). This is in agreement with mixed

clusters observed in PCA and short branches between genotypes in NJ-tree plots. At K = 4, the Boer genotype maintained pure ancestry, since it is an imported exotic breed that is being maintained in a controlled environment. The indigenous genotypes in Uganda are admixed among themselves with low genetic differentiation. This indicates possible gene flow between the genotypes due to close proximity and communal grazing which facilitates random mating. On the other hand, Galla has some genes from ALP and TOT following the introduction of these two exotic breeds in Kenya to improve the dairy productivity in local goats under the FARM-AFRICA community-based goat improvement program (Peacock & Hastings *et al.,* 2011). Alpine and Toggenburg exotic genotypes maintained some gene purity with few traces of ancestries from Boer indicating possible long-time crossbreeding between the genotypes.

The Fst analysis confirmed the genotypes understudy to be genetically distinct with values ranging from 0.191 to 0.324. According to Brito *et al.* (2017), remarkable genetic differentiation between populations should have Fst values greater than 0.15. The highest genetic difference was observed between genotypes of different nations with SEB of Uganda and TOT from Kenya. This indicates restrictions of gene flow between the studied genotypes attributed to isolation distance. These results were expected due to the evidence of genetic relationships among the most studied isolated populations.

3.5 Conclusion

The joint genetic analysis of hybridization and gene introgression between Kenya and Uganda goat genotypes has revealed distinct genetic pools in East Africa. With the in-depth genetic evaluation of the identified gene pools, all nations benefit from the obtained genetic information, which is essential in genetic improvement programs. These programmes can be successfully implemented through the identification and importation of economically important traits from the identified genetic pools that are of farmer and consumer interest.

CHAPTER FOUR

MUTATIONAL LOAD AND DELETERIOUS MUTATIONS IN GOAT GENOME FROM KENYA

Abstract

Mutations that occur in any animal's genome in every generation have an impact on their fitness. Different domestication processes such as selection and effective population sizes are among other factors responsible for the accumulation of mutations in an animal. This study investigated the mutation load and accumulation of deleterious mutations in Kenyan goat populations using Single Nucleotide Polymorphism (SNP) data obtained from local (Galla = 12) and exotic (Saanen = 24, Alpine = 28, and Toggenburg = 30) goat genotypes. The SNP annotation was done on the ENSEMBLE goat (Capra hircus) using the Variant Effect Predictor (VEP) and Sorting Intolerant From Tolerant (SIFT) was done on the annotated file using a score of > 0.01 to describe highly deleterious mutations. Biomart tool in the ENSEMBL goat Capra hircus was used for Gene Ontology (GO) for the genes identified from highly deleterious mutations. The analysis has disclosed no differences in mutation load among the studied genotypes. Overall, the synonymous mutations were in abundance compared to missense mutations, totaling 693 and 258 representing 1.01% and 0.37%, respectively. Mutation load calculated through the ratio of missense to synonymous mutations was therefore similar (0.37) suggesting exposure to similar domestication processes that are creating similar effects on the animal's genome. Further analysis of the missense variants revealed 126 deleterious mutations and 132 tolerated mutations. A total of 111 genes were identified from the deleterious mutations and only 5 were among the highly deleterious mutations (SIFT score > 0.01) which include *PROS1, EHBP1, LTN1, LRRN4,* and *FNDC3A*. Gene Ontology describes these genes as responsible for different functions associated with animal reproduction, diseases, and memory. This study's results have confirmed the levels of mutation load and the presence of deleterious mutations in Kenyan goats. The results can be used as a base to predict the future rate of mutation load hence ensuring the implementation of conservation programs to avoid an accumulation of deleterious mutations.

4.1 Introduction

Mutation is one of the processes of evolution that involve a change in the nucleotide sequence of the genome in an organism. The change can be positive or negative in a population hence the need to know the type of mutation so as to understand the phenotypic perfomance of a population. Negative mutations are harmful or deleterious hence not required in any animal genome due to their effect on an animal's productivity (Pausch *et al.,* 2015). Generally, individuals with deleterious mutations are more vulnerable to diseases or other phenotypic disorders. Occurrences of deleterious mutations in a population is a function of many factors including selection level, domestication, geographical history, and effective population size which are exposed to different plant and animal populations (Bosse *et al.,* 2019). Domesticated goats are exposed to many processes that facilitate the occurrences of deleterious mutations in their populations. For instance, continuous artificial selection for improved traits, reduction in effective population size, and high inbreeding levels in domesticated animals were proved to create a conducive environment for the presence of deleterious mutations due to increased homozygous alleles in a genome (Bosse *et al.,* 2019; Makino *et al.,* 2018; Marsden *et al.,* 2016). In most animal populations, phenotypic variations are observed while their mutational load are not well understood.

A number of studies have been conducted on different livestock species to analyse deleterious variants including goats and cattle (Grossen *et al.,* 2020; Yakubu *et al.,* 2018). Deleterious mutation in *CWC15* and *SLC37A2* genes are reported to be possible causes of low reproductive efficiency in Jersey cattle and fetus abortion in Vorderwald cattle (Reinartz & Distl 2016; Sostengard *et al.,* 2013). Also, a retinal genetic defect in European cattle breeds was discovered to be caused by PR1gene mutation (Michot *et al.,* 2016). Accumulation of deleterious mutations in goat populations were reported in different studies including nine goat breeds of Asian countries (Nomura *et al.,* 2013), the Alpine ibex population (Grossen *et al.,* 2020), and worldwide domesticated goats (Zheng *et al.*, 2020).

Information on the availability of deleterious mutations in a population is not enough. An effort must be made to understand the functions of those identified mutations to facilitate effective selection during breeding. This study investigated the variations in mutation load and deleterious mutations between Kenyan local and exotic goat genotypes and identified affected genes with their associated functions. The findings help in understanding the effects of different factors that expose

Kenyan goat populations to mutations and facilitate measures on how to maintain them to minimize the overload of deleterious mutations.

4.2. Materials and methods

4.2.1 Study site and sampling

The study was done using Single Nucleotide Polymorphism (SNP) data of goats from Kenya. The SNP data were obtained following the materials and methods described in section 3.2.1 from paragraph number 2, section 3.2.2 and section 3.2.3 of this study.

4.2.2 Data analysis

Functional annotation was done on the ENSEMBLE goat (*Capra hircus*) using the Variant Effect Predictor (VEP) based on a method described by Howe *et al.* (2021). The method involves uploading the goat SNP data file on the Ensemble VEP web interface with identified specie *Capra hircus*. Identify the job to be done i.e mutation and then run the selected file in the software. The output includes different statistics of annotation such as SIFT, Consequences, location, genes and transcript e.t.c. A genome annotation file was then downloaded from the Ensemble-VEP containing all the functional consequences information. The number of missense and synonymous variants were calculated from the downloaded file in excel and they were used in calculating the ratio of missense and synonymous to get the mutation load for the population (Bosse *et al.*, 2019).

Sorting Intolerant From Tolerant (SIFT) algorithm was used to identify the mutation effect on protein function and understand the effect of changes in amino acids. The SIFT score ranged from 0 to 1 where a SIFT of 0 to 0.05 is regarded to be deleterious while a SIFT score of greater than 0.05 is tolerable (Kumar *et al.,* 2009; Sim *et al.,* 2012). The SIFT score of less than 0.01 were regarded in this study as the most deleterious hence the variant genes found below the set value were extracted and analysed for their functions using the function settings on Biomart tool in the ENSEMBL goat Capra hircus.

4.3 Results

Annotation of SNPs in ENSEMBLE-VEP revealed the presence of different variants located within and outside the genes. The variants located outside the genes (intergenic and up/downstream variants) accounted for 48.57% while the remaining percentages were variants located within the genes (Table 3). Classification of variants to be missense (mutations that affect amino acid sequence in a protein) or synonymous (mutations with no effect on the amino-acid

produced) showed that the synonymous variants were in abundance compared to missense variants totalling 693 and 258 representing 1.01% and 0.37%, respectively (Figure 6). Further classification of the 258 missense variants showed that 126 variants were deleterious while 132 are tolerated. Only 0.02 % of the variants were found to be eligible for loss or gain of the stop codon. Generally, the ratio of missense to synonymous variants was found to be 0.37 across the populations.

A total of 111 genes were identified under missense variants. Using a SIFT score of less than 0.01, five genes were identified from 21 highly deleterious variants (Table 3).

Table 3. Summary of SNP annotation for goat genotypes in Kenya

Annotation Category	**Numbers**	**Percentage (%)**
Intergenic	25792	37.42
Intronic	30701	44.54
UTR3'/UTR5'	560	0.81
Upstream/downstream	7683	11.15
Exonic		
-Nonsynonymous	258	0.37
-Nonsynonymous deleterious	126	0.18
-Nonsynonymous deleterious (>0.01)	21	0.03
-Nonsynonymous tolerated	132	0.19
-Stop gain/loss	13	0.02
Synonymous	693	1.01
Splicing	91	0.13
Intron non-coding regions	2996	4.35
Protein non-coding regions	128	0.19
Stop retained variants	12	0.02

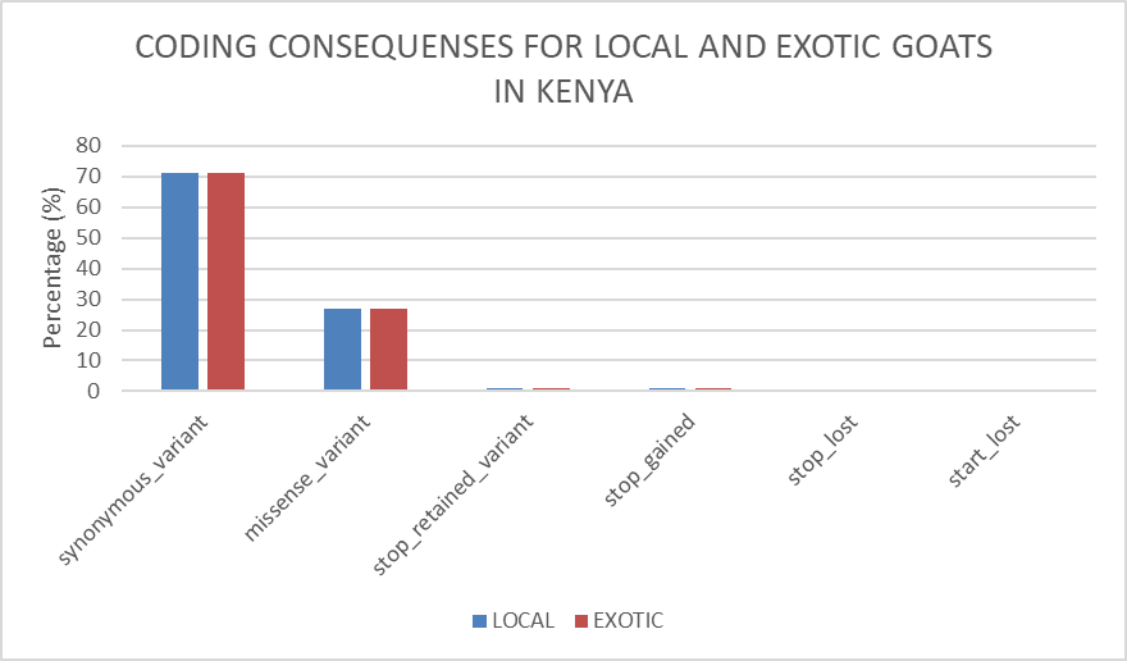


Figure 6. Coding consequences for local and exotic goats from Kenya

The identified genes include *PROS1*, *EHBP1*, *LTN1*, *LRRN4*, and *FNDC3A* which are associated with different functionalities according to Gene Ontology analysis (Table 4). Some of the gene functions analysed include; negative regulation of blood coagulation (*PROS1*), mediation of spermatid-sertili adhesion during spermatogenesis (*FNDC3A*), regulation of vesicular trafficking, and linking endosomes to the actin cytoskeleton (*EHBP1*), animal memory of cells (*LRRN4*) and regulation of translation, elongation and proteasome-mediated ubiquitin-dependent protein catabolic process (*LTN1*).

Table 4. GO terms enrichment for non-synonymous deleterious variants in goat populations based on SIFT prediction

Gene description	GO term name	GO term accession	GO domain	COUNT
Protein S (*PROS1*)	Negative regulation of blood coagulation	GO:0030195	Biological-process	2
	extracellular region	GO:0005576	Cellular-component	2
	calcium ion binding	GO:0005509	Molecular-function	2
	Fibrinolysis	GO:0042730	Biological-process	1
	extracellular space	GO:0005615	Cellular-component	2
EH domain-binding protein 1 (*EHBP1*)	Cytosol	GO:0005829	Cellular-component	1
	plasma membrane	GO:0005886	Cellular-component	1
Fibronectin type III domain-containing 3A (*FNDC3A*)	Membrane	GO:0016020	Cellular-component	4
	An integral component of membrane	GO:0016021	Cellular-component	4
	protein binding	GO:0005515	Molecular-function	4
	Golgi apparatus	GO:0005794	Cellular-component	1
Listerin E3 ubiquitin protein ligase 1 (*LTN 1*)	ubiquitin-protein transferase activity	GO:0004842	Molecular-function	1
	protein autoubiquitination	GO:0051865	Biological-process	1
	RQC complex	GO:1990112	Cellular-component	3
	ribosome-associated ubiquitin-dependent protein catabolic process	GO:1990116	Biological-process	3
	zinc ion binding	GO:0008270	Molecular-function	3
	Cytosol	GO:0005829	Cellular-component	3

	metal ion binding	GO:0046872	Molecular-function	2
	transferase activity	GO:0016740	Molecular-function	2
	protein ubiquitination	GO:0016567	Biological-process	2
	ribosomal large subunit binding	GO:0043023	Molecular-function	2
	proteasome-mediated ubiquitin-dependent protein catabolic process	GO:0043161	Biological-process	2
	rescue of stalled ribosome	GO:0072344	Biological-process	2
leucine rich repeat neuronal 4 (*LRRN4*)	Membrane	GO:0016020	Cellular-component	1
	An integral component of membrane	GO:0016021	Cellular-component	1
	long-term memory	GO:0007616	Biological-process	1
	visual learning	GO:0008542	Biological-process	1
	An integral component of plasma membrane	GO:0005887	Cellular-component	1
	protein binding	GO:0005515	Molecular-function	1

4.4 Discussion

The levels of mutation load obtained in this study through the ratio of missense to synonymous mutations were similar (0.37) across the studied genotypes (Figure 6). This shows that mutations are present in all Kenyan goat populations regardless of the genotype. The similarities can be attributed to the exposure of these animals to almost similar selection criteria towards the same breeding goal creating similar effects on the genome (Ngeno *et al.,* 2015; Visser *et al.,* 2016). Surprisingly, more studies on genomes report only the deleterious mutations observed in a population and not the levels of mutation load making it difficult to compare mutation load across populations (Henn *et al.,* 2015). Conclusions on levels of mutation load are mostly based

on the comparison between populations within or across the study since there are no recommended estimates to base on.

This study followed the principle of Bosse *et al.* (2019) who described the mutation load of 0.025 and 0.03 to be higher in the European pig genome than the 0.02 load observed in the Asian pig genome. In the same study, a mutation load of 0.018 in commercial chickens was described as higher compared to 0.014 observed in the village chicken. Also, Makino *et al.* (2018), described the mutation load of 0.83 and 0.46 for domesticated rabbits and chickens respectively as higher than their wild counterparts with lower values. It is reported that mutation load is common in domesticated animals compared to their wild counterparts (Marsden *et al.,* 2016). Concerning the results of other species mentioned above, it can be concluded that the mutation load of goat genotypes from Kenya is lower than the reported domesticated dogs, rabbits, and chickens but higher than the mutation load of the pig's genome. This can be attributed to the breeding interventions that are being implemented in the sampled goat associations in Kenya. The breeding bucks were being introduced in the farmer associations to improve the existing genetic materials. However, the overuse of these bucks in the population might have led to inbreeding resulting to increase in mutation load and accumulation of deleterious mutation due to reduction in effective population sizes.

Recent mutation studies are focusing on deleterious mutations due to their effects on population fitness. The deleterious mutations observed in this study accounted for 0.18 %, and only 0.03 % were considered to be more deleterious. However, this is lower than the 0.17 % reported in seven Capra species (Grosson *et al.,* 2020). It is reported that deleterious alleles occur at a very low frequency as observed in domesticated goats (Zheng *et al.,* 2020). This can lead to the conclusion that different forces that contribute to deleterious mutations have a limited impact on the accumulation of deleterious mutations in the studied goats.

Out of 111 genes identified from the deleterious mutations, five genes were found within the highly deleterious mutations that are associated with diseases, reproduction, vision, and memory (Table 5). This result helps in understanding the genetic basis of phenotypic variations among individual goats hence ensuring accurate selection during breeding. According to Asayama *et al.* (2020), the *PROS1* gene help in protecting mammals from different diseases such as bronchial asthma in humans. Currently, the effect of the *PROS1* gene in goats has not been reported

however it might also contribute to cardiac disorders hence further research needs to be done to verify.

Naturally, goats have good term memory regardless of sex, age, or breed whose mechanism is unclear (Briefer *et al.,* 2014). It is reported that goats can remember any visual shape within 42 days and a doe can recognize its kids' voice 12 months after weaning (Briefer *et al.,* 2014; Langbein *et al.,* 2008). The *LRRN4* gene found in this study might be the contributing factor to memory, natural feeding, and social interaction behaviour in some goats during management. Goats with difficulties in visual and memory can be due to mutations in the *LRRN4* gene in their genome among other contributing factors. Furthermore, the *LTN1* gene is involved in the biological process of rescuing stalled ribosomes. Stalled ribosomes can occur during protein synthesis due to several factors like depletion of amino acids leading to incomplete protein synthesis. This causes the accumulation of proteotoxic components that lead to cellular stress and neurodegenerative diseases in animals. Scrapie is one of the reported fatal neurodegenerative diseases associated with various clinical signs and it is difficult to diagnose in goats and sheep (Baylis *et al.,* 2004).

Male infertility can be due to several factors including genetic disorders. According to Obholz *et al.* (2006), the *FNDC3A* gene which is found in testis tissues helps in spermatogenesis by mediating the adhesion between spermatids and sertoli cells. The mutation of the *FNDC3A* gene causes male sterility by failure to release mature spermatids. Some male goats in a population might not be reproductive due to unknown causes, this study has therefore provided insight into the possible cause of infertility which requires further studies to validate. Finally, the *EHBP1* gene is essential for several functions in mammals, which include vesicular trafficking, early development as well as cancer (Rai *et al.,* 2020). However, the biochemical processes associated with this function are still unknown.

4.5 Conclusion

The phenotypic differences among individuals can be a factor of many genes working together not the function of a single gene. This study has proven that goat genotypes in Kenya have mutations with presence of some deleterious mutations in their genome. The identified genes from highly deleterious mutations might have negative effects on performance if the phenotype is undesirable. Therefore, there is a need to avoid factors that influence presence of deleterious mutations in a population such as inbreeding. In addition, the identification of mutations, genes,

and their associated functions in goat populations should be prioritised for strategic implementation of breeding programs to avoid increasing gene mutations.

CHAPTER 5

RUNS OF HOMOZYGOSITY AND EFFECTIVE POPULATION SIZE OF DIFFERENT GOAT GENOTYPES IN KENYA

Abstract

Limited genetic information in most goat populations hinders the implementation of better breeding strategies for genetic conservation and improvement. Runs of Homozygosity (RoHs), RoHs distribution, frequency and length, inbreeding coefficients, and effective population size (Ne) were estimated from a total of 48808 Single Nucleotide Polymorphisms (SNP). The SNP data of four goat genotypes from Kenya were used; Galla (n = 12), Alpine (n = 28), Saanen (n = 24) and Toggenburg (n = 30). Across the genotypes, 348 RoHs were detected with the highest number (180) observed in Toggenburg and the lowest (22) in Galla. The distribution of RoHs per chromosome was breed-specific without a clear pattern across the genotypes. Furthermore, 32 genomic regions with a high frequency of RoHs were detected. Sixteen genes associated with various phenotypic functions were identified from these RoHs islands. High inbreeding coefficients (> 0.1) were observed in all the exotic genotypes suggesting continuous use of few breeding bucks. The trend of effective population size decreased over time across the genotypes indicating limited genetic diversity. This study outcome has provided the required genetic reference for sustainable implementation of different breeding programs which will help in increasing the genetic diversity of goat genotypes in Kenya. Such might include opening up the existing nucleuses, use different lines of goat genotypes, improved technologies, and implementation of controlled breeding programs.

5.1 Introduction

Goat production in Kenya like in many other African nations is important in improving rural livelihood through the provision of meat, milk, and income among other benefits (Monau *et al.,* 2020a). Kenya has a diverse gene pool of both exotic and local genotypes essential for genetic improvement programs (Kivila *et al.,* 2018; Wainena *et al.,* 2021). The shape of the animal genome in any structure depends on many factors such as geographical location, production, and breeding systems that have the potential to increase or decrease genetic diversity (Bosse *et al.,* 2012; Szmatola *et al.,* 2019). When genetic diversity is low in a population, close relatives are mated leading to inbreeding, reduced genetic variation, and affecting animals' fitness. Inbreeding levels can be measured at both individual and population levels. Due to improvements in genomic technologies, Runs of Homozygosity (RoHs) is known to be the most effective way of measuring inbreeding in a population through the estimation of inbreeding coefficients (Peripolli *et al.,* 2017; Rebelato & Caetano, 2018).

Ceballos *et al.* (2018), defines Runs of homozygosity as the continuous homogenous regions of the genome in an individual that occurs due to the inheritance of identical alleles from their parents. According to Szmatola *et al.* (2019), unlimited artificial selection for beneficial alleles in a population can also increase homozygosity in genomic regions. In a genome, RoHs can be either long or short and they are not randomly distributed in the animal genome (Zhang *et al.,* 2015). Furthermore, recent inbreeding is presented by long RoHs while short RoHs indicate ancient individual relationships. It is reported that differences between livestock genotypes and other genomic information can be drawn from the patterns of RoHs in specific genomic regions of individuals (Islam *et al.,* 2019; Upadhyay *et al.,* 2017).

Effective population size (Ne) is defined as the size of an idealized population that undergoes the same rate of genetic drift and inbreeding as the actual population under study (Falconer & Mackay, 1996). The Ne helps to describe the genetic diversity level of a population and it is estimated by measuring pairwise Linkage Disequilibrium (LD) as a squared correlation coefficient (r^2). Linkage Disequilibrium refers to the non-random association of alleles which depends on evolutionary history and the effective population size (Deng *et al.,* 2019). Changes in Ne over time in a population help to measure population genetic diversity and implement conservation of important animal genetic resources.

Inbreeding levels in most goat populations in Kenya measured by RoHs are not known despite uncontrolled breeding being the common practice. Furthermore, the information on the Ne for different goat genotypes is limited hence difficult to facilitate the implementation of effective breeding programs. Using Single Nucleotide Polymorphism (SNP) data for four goat genotypes (Galla, Alpine, Saanen and Toggenburg), this study focused on genomic characterization of RoHs distribution, inbreeding coefficients, and Ne among exotic and local goat genotypes kept in Kenya. This analysis provided important information to ensure the maintenance of genetic distinctiveness, better conservation strategies, and identification of adaptive goat genes in Kenya.

5.2 Materials and Methods

5.2.1 Study site and sampling

The study was done following the materials and methods described in sections 3.2.1 from second paragraph, section 3.2.2 and section 3.2.3 of this thesis.

5.2.2 Data analysis

Distribution of runs of homozygosity

Runs of Homozygosity were determined using runs of the homozygosity tool in PLINK v1.9 (Chang *et al.*, 2015). This uses a window of a specific length or number of homozygous SNPs to scan each SNP marker position to detect homozygous segments (Howrigan *et al.,* 2011). Homozygosity in this study was defined based on the following parameters; having a minimum number of 15 consecutive homozygous SNPs, a minimum physical length of 1 Mb, 1 maximum missing genotype and 1 heterozygous call was allowed within the RoHs for genotyping errors (Islam *et al.,* 2019; Kumar *et al.,* 2018). For each chromosome, the number of RoHs and the percentage of chromosomes covered by RoHs were calculated following the formula by Al-Mamun *et al.* (2015) as shown below;

$$\text{Mean RoHs length} = \frac{\text{Sum of all RoHs on a chromosome}}{\text{Total No. of individuals that has RoHs on that particular chromosome.}}$$

The percentage of chromosomes covered by RoHs were calculated through;

$$\% \text{ of Chr. With RoHs} = \frac{\text{Mean RoHs length}}{\text{Respective chromosome length}} \times 100$$

The total number, frequency, and length distribution of RoHs (in Megabases) were identified in each genotype. To evaluate the distribution of RoHs length in each genotype, the identified lengths were categorized into four classes; 2–4 Mb, 4–8 Mb, 8–16 Mb, and above 16 Mb as used in related studies (Mastrangelo *et al.,* 2021; Onzima *et al.,* 2018a). The sum of all RoHs segments for each length category per genotype divided by the total number of individuals in that breed produced the frequency and mean RoHs length for each category.

Estimation of inbreeding coefficient

Inbreeding coefficients were estimated per individual genome and genotype. Runs of homozygosity inbreeding coefficients (F_{RoHs}) were determined using the formular below as used by Kumar *et al.* (2018) and Islam *et al.* (2019).

$$F_{RoHs} = \frac{\text{Total RoHs length } (L_{RoHs})}{\text{Autosomal genome length } (L_{AUTO})}$$

Genomic regions with high runs of homozygosity frequency

To identify genomic regions with high RoHs frequencies, the number of times each SNP occurred in the RoHs were calculated throughout the population to get the percentage of SNP occurrence. From the percentages obtained in each genotype, the top 10% of genomic regions with a high frequency of RoHs were selected for downward analysis. This 10% percentage was randomly selected from within the range which other scholars used. Different scholars used different values which ranges from 1% to 45% (Islam *et al.,* 2019; Mastrangelo *et al.,* 2018; Purfield *et al.,* 2017) among others. The identified genomic regions were extracted using vcftools software and uploaded in the ENSEMBL goat Capra hircus using the Variant Effect Predictor (VEP) for gene annotation.

Effective population size

The software SNeP v1.1 (Barbato *et al.,* 2015) was used to estimate the effective population size among the genotypes based on linkage disequilibrium. This is done by running a SNP - - ped and - - map files on the command prompt using different commands. The software follows the formula described by Sved (1971):

$$E(r2) = \frac{1}{1 + 4\,Nec}$$

Where;

Ne is the effective population size,

c is the genetic distance between SNPs in Morgans

$E\ (r^2)$ is the expected correlation between allele frequencies of two loci

The estimated effective population sizes (Ne) were plotted against the past 1000 generations to determine its trend.

5.3 Results

Distribution of the runs of homozygosity

A total of 348 RoHs were detected in the four goat genotypes with a mean of 4.703 per genome. The highest number of RoHs was observed in exotic genotype (Toggenburg) while local genotype (Galla) recorded the lowest number (Table 6). Twenty individuals out of 94 individuals from across the genotypes (GAL=7, ALP=8, SAA=2, and TOT=3) did not have RoHs.

Table 5. Descriptive statistics of runs of the homozygosities per genotype

Genotype	No. of RoHs	Mean No. of RoHs	No. of individuals without RoHs	RoHs length (Mb)	Mean RoHs length	CHR. With longest RoHs	RoHs length (Mb)	No. of SNP
Alpine	54	2.7	8	554.92	27.75	4	68.04	1411
Galla	22	4.4	7	211.17	42.23	2	34.91	712
Saanen	92	4.2	2	846.67	38.49	6	113.33	2338
Toggenburg	180	6.7	3	1631.53	60.43	4	244.6	5023

Toggenburg had the longest average sum of RoHs length with a mean of 60.43 followed by the GAL genotype with a mean sum length of 42.23 Mb. Alpine showed the lowest average RoHs length of 27.75 Mb compared to all other genotypes.

Runs of homozygosity detected per chromosome varied according to the genotype in all the 28 chromosomes (Figure 7). Chromosome four of TOT recorded the highest number of RoHs with 5023 SNPs and a corresponding chromosomal length of 244.6 Mb. The total number of RoHs across the chromosome was lowest in GAL genotype and highest in TOT genotype (Figure 7). The percentage of chromosomes covered by RoHs was calculated across the genotypes for all the chromosomes with the highest percentage observed in Galla chromosome 16 (Appendix 1).

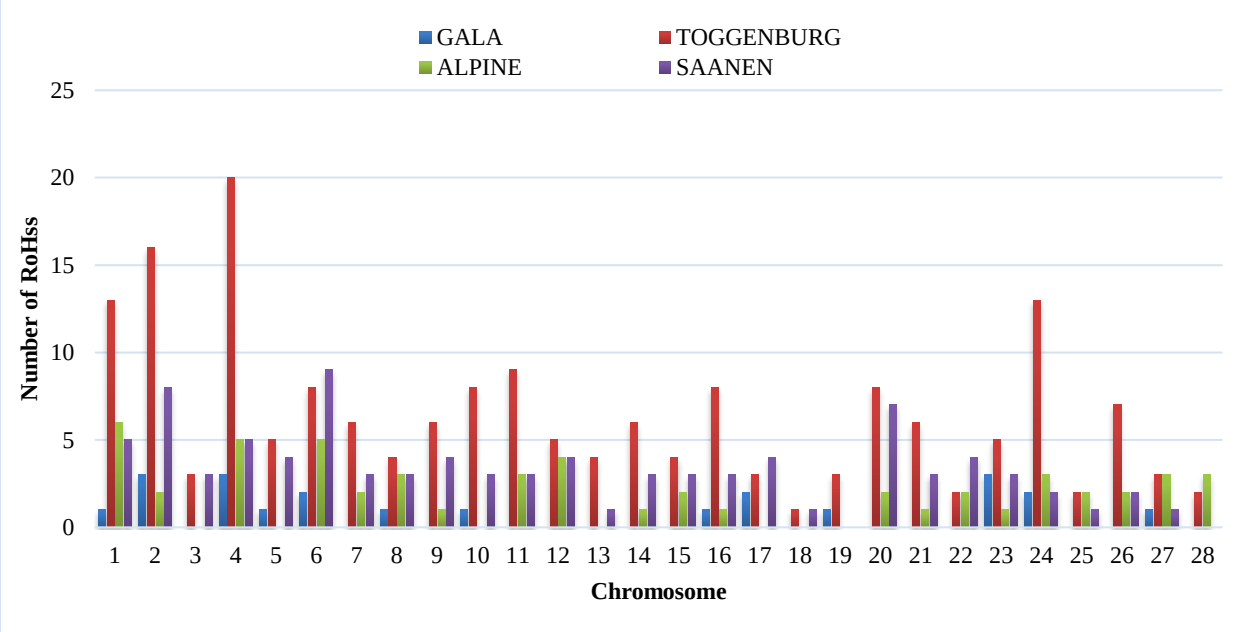


Figure 7. Runs of the homozygosity per chromosome per genotype

To identify the RoHs patterns, the identified RoHs were grouped into four categories. Shorter RoHs were more than the longer ones in this analysis. The highest numbers of RoHs were observed in the 4-8 Mb length category across all genotypes with the greatest number of RoHs in TOT and the lowest in the GAL genotype (Table 7). The number of RoHs per genotype decreased with an increase in the length category, while the average RoHs categories increased as the RoHs category increased in all the genotypes. The highest mean length was observed in the category of >16 Mb (Table 7 & Figure 8). For instance, the average RoHs length for TOT at 4-8 Mb, 6-16 Mb, and >16Mb was 5.99, 10.45, and 21.25, respectively. In general, the ALP genotype had the highest

mean length in all the categories followed by GAL and SAA becomes the least among all genotypes.

Table 6. Runs of the homozygosity length categories, number of individuals with RoHs, number and mean sum of RoHs length (Mb) across the genotypes

RoHs length category	Galla			Toggenburg			Alpine			Saanen		
	RoHs No.	No. of Indv.	Mean Length	RoHs No.	No. of Indv	Mean Length	RoHs No.	No. of Indv	Mean Length	RoHs No.	No. of Indv	Mean Length
2-4Mb	0	0	0	0	0	0	0	0	0	0	0	0
4-8Mb	12	5	6.13	102	27	5.99	30	17	6.33	50	22	5.91
8-16Mb	8	5	12.02	59	22	10.45	16	12	12.02	32	14	10.68
>16Mb	2	2	20.76	19	13	21.25	8	4	24.84	10	8	20.95

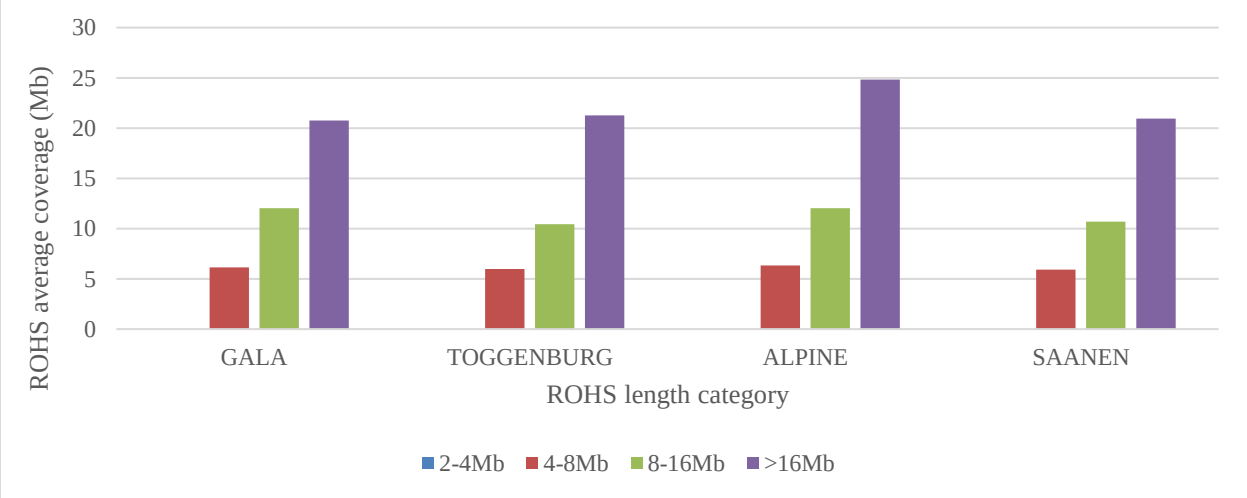


Figure 8. Distribution of mean sum of RoHs coverage per length category averaged per breed across four genotypes

Inbreeding coefficients

Toggenburg genotype is the most inbred with an inbreeding coefficient of 0.68 followed by SAA (0.35), ALP (0.23), and then GAL (0.09). Mean F_{RoHs} across the genotypes ranged from

0.01 for ALP to 0.03 for TOT (Table 8). As per individuals, F_{RoHs} ranged from 0.00 to 0.07 with the highest F_{RoHs} of 0.07 observed in TOT (Appendix 2).

Table 7. Inbreeding coefficients, Mean, and F_{RoHs} per genotype

Genotype	Galla	Toggenburg	Alpine	Saanen
Inbreeding coefficient per genotype	0.09	0.68	0.23	0.35
Total length (RoHs)	211.17	1631.53	554.92	846.67
Mean Length (RoHs)	42.234	60.427	27.746	38.485

Genomic regions with the high frequency of RoHs

From the analysis, 32 genomic regions with the high frequency of RoHs were identified (Appendix 3). Sixteen genes of missense and synonymous nature and with different functionalities were identified from these RoHs islands.

Effective population size (Ne) and its trend

The estimates of ancestral effective population size (Ne) over past generations obtained in this analysis are presented in Appendix 4. As the number of generations increased, effective population size across the genotypes decreases at different rates. The Ne for Alpine tends to decrease rapidly compared to all other genotypes in the study (Figure 9). At the most recent generation (13^{th}), the Ne for Alpine, Galla, Saanen, and Toggenburg was 109, 49, 81, and 93, respectively indicating a narrow genetic pool for all the genotypes except Alpine. The Ne for the furthest distant generation (983 generation ago) was 3709, 2428, 7515, and 2548 for ALP, GAL, SAA, and TOT respectively.

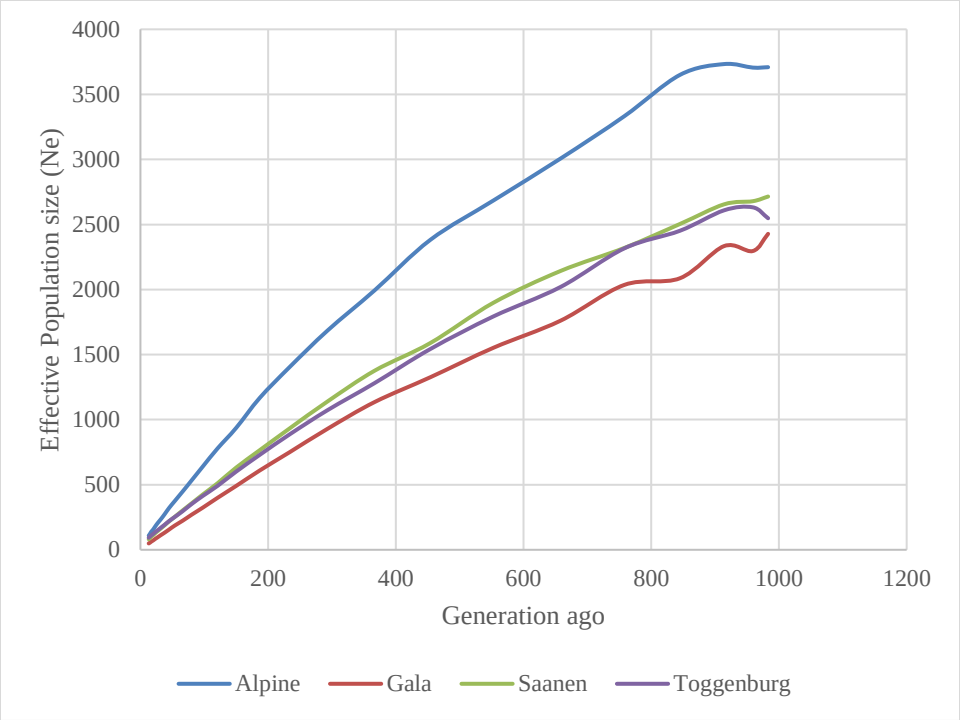


Figure 9. The effective population size over generations for Alpine, Galla, Saanen, and Toggenburg goat genotypes

5.4 Discussion

Different studies have been conducted to analyse various genetic parameters of goat genotypes in Kenya. However, the characterization of RoHs and Ne in goat genotypes remained unknown. Waineina *et al.* (2021) reported the genetic diversity and population structure among indigenous (Galla) and exotic (Alpine, Saanen, and Toggenburg) goat breeds in Kenya.

5.4.1 Runs of homozygosity

Results indicated that all the genotypes under study have RoHs in their genomes whose presence varies in terms of the number, length, and distribution. This is similar to the distribution of RoHs observed in different goat populations and other livestock like cattle (Cardoso *et al.*, 2018; Michailidou *et al.*, 2019; Szmatola *et al.*, 2019). The formation of RoHs in any population is a factor of different demographic events and recombination rates, hence the observed variation in the RoHs for the Kenyan goat genotypes (Bosse *et al.*, 2012). The large numbers of RoHs per individual observed in exotic genotypes compared to the local genotypes indicate increased homozygosities in exotic genotypes. However, the average RoHs length was higher in Galla compared to Alpine and Saanen, which recorded low numbers of RoHs. The observed results in

this study are comparable to related studies done on goat genotypes from different agro-ecological regions (Cardoso *et al.,* 2018; Michailidou *et al.,* 2019).

From the assigned length categories, the short RoHs (4-8Mb) were more frequent across the genotypes. However, the average sum of RoHs coverage across the genotypes was higher for longer RoHs than the shorter RoHs (Figure 8). The highest average sum of RoHs (24.84 Mb) was observed in Alpine (Figure 8). This suggests high levels of recent inbreeding in the Alpine genotype that can be attributed to the artificial selection of best breeding bucks, the presence of few replacement stocks, or limited population of imported breeding stock. This result is in tandem with the findings in Asian pigs (Bosse *et al.,* 2012) and Italian goats (Mastrangelo *et al.,* 2021c). On the contrary, the average RoHs were more common in the short RoHs category compared to long RoHs consistent with RoHs studies of goats and sheep (Purfield *et al.,* 2017; Onzima *et al.,* 2018c). Generally, the majority of the mean sum of RoHs coverage is reported at the length of >16Mb (Figure 8) suggesting recent inbreeding across the genotypes. The long RoHs genomic sections in exotic genotypes were also observed in similar related studies of goats and sheep (Kim *et al.,* 2016; Onzima *et al.,* 2018c). This information is important for planning better breeding programs since most deleterious variants are known to be carried in long RoHs (Szpiech *et al.,* 2013).

The distribution of RoHs per chromosome was different across the genotypes with each genotype displaying a non-specific pattern. In similar studies of RoHs done in sheep and goats, the distribution of RoHs per chromosome were found to be breed-specific (Islam *et al.,* 2019; Mastrangelo *et al.,* 2017a). Chromosome four of Toggenburg has recorded the highest number of RoHs, indicating the possible continuous transfer of ancestral genes specific to that chromosome, unlike other chromosomes that were observed without RoHs.

Observed F_{RoHs} per individual were generally low, ranging from 0.00 to 0.07 indicating limited inbreeding levels within individuals. For instance, F_{RoHs} for Alpine were reported below 0.05 in this study, similar to the findings of Marete *et al.* (2011). However, F_{RoHs} values per genotype were high for the exotic Toggenburg (0.65) and moderate for Alpine and Saanen compared to local genotype Galla (0.09). According to Luingi-Sierra *et al.* (2019), a population with low inbreeding levels must have inbreeding coefficient levels of less than 0.1. In this study, only local Galla recorded an F_{RoHs} value of 0.09 (Table 8) which corresponds with the observed low numbers of RoHs suggesting that the genetic material for this genotype is at least well

managed at the government station in Kenya but measures must be implemented to maintain recommended inbreeding levels. An analysis of Spanish Churra and Italian sheep breeds had similar results as observed in Kenyan goat genotypes were obtained (Chitneedi *et al.,* 2017; Mastrangelo *et al.,* 2018).

The moderate to high inbreeding coefficients in exotic genotypes were in agreement with values observed in goats from different geographical locations of the world in a study by Bertolini *et al. (*2018) and it can be attributed to the extensive use of exotic bucks by smallholder farmers since the small population of the genotypes was imported into Kenya to improve local goat productivity. Special considerations should therefore be made to have different lines of exotic goat genotypes or the use of Artificial Insemination (AI) for effective genetic improvement programs. From the genomic regions associated with the high frequency of RoHs, this study focused on the presence of synonymous and missense genes. Missense genes (13) were more compared to synonymous genes (3), (Appendix 5). These identified genes are associated with genetic disorders, diseases, reproduction, and general body immunity according to gene ontology analysis.

5.4.2 Effective population size and future trends

The Ne for all the genotypes in the very distant past (983 generations ago) was high with Ne values of above 2000 across the genotypes. This indicates high genetic varialility across the studied genotypes. In the 20^{th} generation ago, the Ne for all exotic genotypes was above the threshold of Ne = 100 with 156, 113, and 122 for Alpine, Saanen, and TOT. These results are comparable with Ne observed in the same breeds in the 20^{th} generation ago by Brito *et al.* (2015). Over time until the recent present, a decrease trend in the Ne was observed (Figure 9) indicating limited genetic variability and possibility of inbreeding between individuals. This decreasing trend was also observed in local Swiss sheep (Burren *et al.,* 2014), Australian and Canadian Boer goats (Brito *et al.,* 2015), Buffalo populations (Deng *et al.*, 2019), and local South African goats (Monau *et al.*, 2020a).

To ensure the long-term viability of any livestock population, the effective population size must reach a threshold of Ne = 100 (Meuwissen *et al.,* 2009). However, in the 13^{th} generation ago (recent), Ne for all the genotypes except Alpine did not meet the required Ne threshold (Ne = 100) indicating limited genetic diversity. Similar results were also obtained at the 13^{th} generation ago in two goat populations of China (Islam *et al.,* 2019). In order to have an increase in Ne trend in the future generations there is a need to implement measures such as the exchange of breeding

bucks or the use of artificial insemination can be implemented in Galla, Saanen, and Toggenburg to ensure the required levels of diversity are sustained. It is important to ensure that populations of local genotypes have high genetic variations at all times since they are a source of many genetic materials adaptable to the local environment (Monau *et al.,* 2020a).

5.5 Conclusion

In conclusion, RoHs have been detected with inbreeding coefficients of greater than 0.09 in all Kenyan goat genotypes under this study. This indicates that there is considerable levels of homozygosity in the studied populations which can affects the future productivity of goats if different genetic interventions are not implemented. Similarly, the Ne decreases with time across the genotypes except for Alpine indicating declining in genetic diversity. Therefore, different lines for these exotic genotypes must be prioritized and introduced to ensure a diverse genetic pool for improved productivity as well as effective improvement and conservation programs.

CHAPTER SIX

GENERAL DISCUSSION, CONCLUSION AND RECOMMENDATIONS

6.1 Aim of the study

Local genetic resources for small ruminants have been shown to provide a significant impact on livelihood in any production area where better feeding and health care are provided (Kaumbata *et al.,* 2021). Goat breeding is aimed at the improvement of animals using a well-developed breeding programme. A successful animal breeding program must start with definition of the breeding objectives which will guide the farmers on measures to implement to achieve the set of objectives which will be evaluated in the subsequent generations (Nandolo *et al.,* 2016). In the process of formulating breeding objectives, priority is given to the economic traits. Goat populations are exposed to different factors that affect their genetic make-up, creating special genetic structures and diversity within or among populations as presented in chapter 2.1 of this thesis. In animal breeding, 50 % of DNA is passed on from each parent to the offspring. Therefore, the collection of information and genetic evaluation are some of the breeding activities essential during the implementation of any breeding programme. Information about genetic relationships is essential for selection and breeding during the implementation of the long-term breeding strategies in the livestock sector for future productivity and adaptation to environmental changes. Therefore, understanding the genetic basis of goat performance and the level of diversity across nations is beneficial but limited in most goat populations. The development of different genomic technologies has however enabled easy and effective evaluation of various genetic parameters within and among populations using different markers such as Single Nucleotide Polymorphism (SNP).

This thesis aimed at contributing toward increased goat productivity and sustainability of goat genetic improvement programs in Kenya and Uganda communities through joint genomic analysis using genome-wide SNP data. The research questions that were addressed included: i) what are the characteristics of hybridization between different goat genotypes in Kenya and Uganda? ii) what are the mutation load and deleterious mutations of goats from Kenya? iii) what are the Runs of Homozygosity and effective population size, and the predicted future trends of goats from Kenya? These questions were derived from the following specific objectives: i) to genetically characterize hybridization between different goat genotypes from Kenya and Uganda, ii) to identify the mutation load and deleterious mutations in goats from Kenya, iii) to estimate

runs of homozygosity and effective population size and predict their future trends of goats from Kenya.

6.2 Study methodology

This study used genome-wide SNP data from 144 goats from six agro-ecological zones in Uganda (Onzima *et al.,* 2018a) and 94 goats from different counties in Kenya representing four goat genotypes as presented on chapter 3.2.1, 3.2.2 and 3.2.3 of the thesis. To genetically characterize hybridization and gene introgression between the goat genotypes from Uganda and Kenya, Principal Component and Phylogenetic analysis were done in Tassel software v 5. This helped to predict the genetic relationships between the studied genotypes. For comparison, Plink v 1.9 software was adopted to measure the sub-structuring of the studied genotypes using identical by-state distances (IBD) and multi-dimension scaling (MDS). To reduce data redundancy, SNP data were pruned in Plink v 1.9 to allow effective evaluation of genetic admixture between the individuals. Admixture software v 1.3.0 was used to determine the level of genetic admixture.

Survivability of any population is not only dependent on the population size but also on other genetic parameters within a population such as mendelian sampling variances, breeding values, and effects of mutation accumulation (Grossen *et al.*, 2020; Wellmann & Bennewitz, 2019). The second objective, therefore, focused on the identification of mutation load and deleterious mutations in goat genotypes from Kenya by annotating the SNPs using the Variant Effect Predictor (VEP) based on a method described by Howe *et al.* (2021). This study used the ratio of missense and synonymous SNPs to get the mutation load for the population and deleterious variants were identified by Sorting Intolerant from Tolerant (SIFT) with SIFT score of less than 0.01. Runs of homozygosity (RoHs) were identified in Plink v 1.9 following a specific window length to scan each SNP marker position to detect homozygous segments (Howrigan *et al.*, 2011). Finally, SNeP v1.1 (Barbato *et al.,* 2015) was used to estimate the effective population size among the genotypes based on Linkage Disequilibrium (LD).

6.3 Genetic characterization of hybridization and gene introgression between the goat genotypes from Uganda and Kenya

Chapter three of this study shows that there is no genetic introgression and hybridization between the Ugandan and Kenyan goat genotypes meaning they are genetically distinct from each other. This concurs with the fact that isolation distance and variations in breeding strategies between populations affects the genetic make-up of animals (Visser *et al.,*2016). The study output

can be applied across nations in the development of policies to govern the development and implementation of national breeding and conservation programs of animal genetic resources across the two or more nations. The results have justified that genetic improvement and conservation of goats across nations is very possible. Different nations can genetically benefit from each other through the exchange of the best genetic materials as parental breeding stock. According to Faraque *et al.* (2010), the selection of breeding stock for genetic improvement depends on genetic variations within or between populations. Individuals with traits of economic importance that are of farmers' interest such as disease resistance and better survivability, can be identified and selected as parent stock across different nations. This approach must also prioritize the local genetic material due to their unique adaptability in diverse conditions compared to the exotic genotypes.

Furthermore, the information generated from this study will be utilized in a conservation strategy of local genes required to survive and produce in future changing environments. For instance, the indigenous goats in Kenya are not genetically distinct (Githui *et al.,* 2016; Kivila *et al.,* 2018). On the other hand, Ugandan indigenous goat breeds are genetically diverse with narrow sub-structuring (Onzima *et al.,* 2018a). Therefore, using local genotypes from across nations will help in improving local genetic materials adaptable to future climate change and ensure genetic variability.

Exotic breeds have been used for a long time to improve local goats posing a threat of dilution of important local genes. Currently, controlled breeding is lacking in most African goat populations. Therefore, adopting the Community Based Goat Breeding Programs (CBBPs) which are farmer-oriented is ideal for using the local improved bucks identified in the distinct clusters. To ensure the sustainability of using local genotypes in the CBBPs, the program must be supported by permanent established institutions and organizations (Haile *et al.,* 2019). This program has proved to be a success in the selected studied goat populations. Improvement in goat productivity and genetic diversity among small ruminant populations were reported in Ethiopia, Malawi, Uganda, Kenya, and Liberia (Ahuya *et al.,* 2003; Haile *et al.,* 2019; Karnuah *et al.,* 2018; Kaumbata *et al.,* 2021).

6.4 Identification of mutation load and deleterious mutations in goat genotypes

The distinct genetic pools observed in chapter three were attributed to natural and artificial selection in addition to differences in geographical locations. This led to the analysis of the second objective as described in chapter four of this thesis. In-depth genetic evaluation of any genetic pool

across the world is essential in animal breeding programmes. This led to the identification of mutation load and deleterious variants in the exotic and local genotypes of goat populations in Kenya. The presence of deleterious mutations in a population is due to high coverage of homozygosity runs in the long runs of homozygosity among other factors. It has been shown in chapter four of this thesis that there is an accumulation of deleterious mutations in goat genotypes from Kenya which can cut across many goat populations within and across nations due to uncontrolled breeding systems. The selection history and evolution of a population determine the impact of deleterious alleles in a population which is shown as cumulative, not in an individual. Therefore, the deleterious alleles are cumulatively passed on to the offspring creating a large impact in livestock populations such as fertility problems as observed in pigs and cattle (Derks *et al.,* 2017; VanRaden *et al.,* 2011). Since deleterious alleles are eliminated from the population through selection, there is a need to effectively remove those that are detrimental.

In most livestock populations, the effects of many deleterious mutations on phenotypes are not clear. This Thesis (Chapter Four) reported the few functions of identified deleterious mutations with much focus on the coding regions. It has to be noted that all mutations have associated functions hence the focus in this study on only highly deleterious mutations indicates a loss of other information in the study. The *PROS1, EHBP1, LTN1, LRRN4,* and *FNDC3A* genes were found among the most deleterious variants that are beneficial to the animal's survival hence the need to be maintained in the population. To the author's knowledge, the functions of these genes were only reported in other mammals such as humans and mice (Burstyn-Cohen *et al.,* 2009; Joazeiro, 2019; Zhang *et al.,* 2012). None of the identified genes in this study has been evaluated and documented for any goat population. This calls for the focus on the functions of deleterious genes in future studies.

6.5 Estimation of runs of homozygosity and effective population size of goats from Kenya

Runs of homozygosity's are a home for homozygous deleterious mutations and it is the best indicator for determining inbreeding levels hence Chapter Five. Some of the contributing factors to the accumulation of deleterious mutations in a genome is due to low effective population size and inbreeding. In Chapter Five of this Thesis, it has been shown that RoHs are present across genotypes with high levels observed in exotic genotypes compared to local genotypes. This applies to many goat populations in Africa due to almost similatities in unsustainable breeding systems. For instance, the highest RoHs number and length were also observed in improved Awassi sheep

breeds compared to the local Ethiopian sheep breeds, (Getachew *et al.,* 2020). For a long time, exotic breeds have been used to increase meat and milk productivity through selection while in local breeds little or no selection is done for breeding across many nations despite possessing important traits. This breeding system increases the alleles that are selected for in an individual genome, creating genomic changes that contribute to the genetic differences between genotypes (Maynard & Haigh, 2007).

Homozygosities cause various effects on animal production and productivity. For instance, in Murciano-Granadina goats, RoHs are reported to reduce production and productivity by affecting udder health and milk yield (Luigi – Sierra *et al.,* 2022). Furthermore, Makanjuola *et al.* (2021), reported the unfavourable production and fertility traits in Holstein cattle due to a single and unique RoHs genotype found in the genomic region of interest. Homozygosity is related to inbreeding which occurs due to poor genetic structure and variability. Hence, the negative effects of homozygosity can be avoided by controlling the breeding rate per generation and ensuring the maintenance of recommended breeding ratio (Muasya *et al.,* 2013; Mwangi *et al.,* 2020). This, therefore, calls for controlled breeding practices and the use of different lines of exotic genotypes to avoid the negative impacts of RoHs in a population in all African goat populations. Similarly, the local Galla genotype from small-holder farmers requires good breeding management to ensure improved productivity while avoiding inbreeding. The results in this study are for a local breed that is under controlled management hence low levels of RoHs and inbreeding. However, due to uncontrolled breeding among farmers' goats, it can be possible that inbreeding levels are high therefore the need to be verified in future research work. The analysis of RoHs indicated some of the SNPs being linked to different genes, both missense and synonymous (Chapter Five) which are associated with animal health and reproduction. However, the role of the identified genes in goats is not very clear, hence the need for further research.

Finally, there is a decrease in trend on effective population size across the studied genotypes across the genotypes which concurs with results presented in chapter 4 and 5 of this thesis. The decrease in Ne contributes to inbreeding in any animal population. To ensure the long-term survivability of any livestock population the recommended Ne is supposed to be 100 (Meuwissen *et al.,* 2009; Mwangi *et al.,* 2020). There is a need to improve the breeding stock across the genotypes through use of new blood lines to facilitate increase in population size. Much

emphasis should be on the local Galla genotype which is more threatened with the lowest effective population sizes (Ne = 49).

6.6 Conclusions

i. Uganda and Kenya goat genotypes used in this study are genetically distinct and clustered into two main gene pools according to their country of origin
ii. In Kenyan goat genotypes, there is accumulation of mutations with the presence of some deleterious mutations hence exposing the population to some disorders.
iii. The presence of Runs of homozygosities in the entire individual's genome has been verified and their effective population size is decreasing with increase in the generations.

6.7 Recommendations

From the results of this study, the following recommendations are made:

i. Develop a regional breeding programme to exploit more on the goat population genetic diversity inorder to identify genes of economic importance that are of farmer's interest
ii. Encourage use of improved breeding systems among goat farmers inorder to reduce the mutation load and deleterious mutations in Kenyan goat genotypes.
iii. Incorporate measures of managing the genetic diversity to reduce presence of RoHs and maintain the recommended levels of effective population size
iv. Further studies using high density markers are necessary to evaluate the level of genetic autozygosity and signatures of selection of goat breeds.

REFERENCES

Abraham, H., Gizaw, S., & Urge, M. (2019). Growth and reproductive performances of Begait goat under semi-intensive and extensive management in western Tigray, north Ethiopia. *Livestock Research for Rural Development, 31(32).* http://www.lrrd.org/lrrd31/3/hagsa31032.htm

Abril, J. F., & Castellano Hereza, S. (2019). Genome annotation. In M. Yandell & D. Ence. (Eds *Encyclopedia of Bioinformatics and Computational Biology. (pp. 195-209).* Elsevier: Amsterdam, Netherlands.

Acaitum, M. (2018). *Phenotypic characterisation of indigenous goats and perception of farmers on indigenous versus cross-bred goats in North-Eastern Uganda* [Makerere University].

Agrawal, A. F., & Whitlock, M. C. (2012). Mutation load: the fitness of individuals in populations where deleterious alleles are abundant. *Annual Review of Ecology, Evolution, and Systematics, 43*, 115-135.

Ahuya, C. O., Okeyo, A. M., Mosi, R. O., & Murithi, F. M. (2003). *Growth, survival and milk production performance of Toggenburgs and their crosses to East African and Galla goat breeds in the Eastern slopes of Mount*, Proceedings of LPP International Workshop on Small Ruminants Research and Dissemination. Embu, Kenya, 3-7th March 2003

Ahuya, C. O., Okeyo, A. M., & Peacock, C. (2005). Developmental challenges and opportunities in the goat industry: the Kenyan experience. *Small Ruminant Research, 60*(1-2), 197-206.

Alexander, D. H., Shringarpure, S. S., Novembre, J., & Lange, K. (2015). Admixture 1.3 software manual. *Los Angeles: UCLA Human Genetics Software Distribution*, *Genome Research, 19*, 1655-1664

Al-Mamun, H. A., Clark, S. A., Kwan, P., & Gondro, C. (2015). Genome-wide linkage disequilibrium and genetic diversity in five populations of Australian domestic sheep. *Genetics Selection Evolution*, *47*(1), 1-14.

Amornbunchornvej, C., Wangkumhang, P., & Tongsima, S. (2020). ipADMIXTURE: R package for inferring sub-population clusters based on genetic admixture. *BioRxiv, 10*, 1-12 doi:https://doi.org/10.1101/2020.03.21.001206

Asayama, K., Kobayashi, T., D'Alessandro-Gabazza, C. N., Toda, M., Yasuma, T., Fujimoto, H., & Gabazza, E. C. (2020). Protein S protects against allergic bronchial asthma by modulating Th1/Th2 balance. *Allergy, 75*(9), 2267-2278.

Baloch, S. N. (2014). *Study on performance analysis of Kamohri goat flock under semi-intensive management at Government Bhagnari Cattle farm Usta Muhammad Baluchistan*. M. Sci. Master of Science Thesis, Sindh Agriculture University Tandojam.

Barbato, M., Orozco-terWengel, P., Tapio, M., & Bruford, M. W. (2015). SNeP: a tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Frontiers in Genetics*, *6*, 109.

Baylis, M., & Goldmann, W. (2004). The genetics of scrapie in sheep and goats. *Current Molecular Medicine*, *4*(4), 385-396.

Berhanu, T., Thiengtham, J., Tudsri, S., Abebe, G., Tera, A., & Prasanpanich, S. (2012). Purposes of keeping goats, breed preferences and selection criteria in pastoral and agro-pastoral districts of South Omo Zone. *Livestock Research for Rural Development*, *24*(12), 2012.

Bertolini, F., Cardoso, T. F., Marras, G., Nicolazzi, E. L., Rothschild, M. F., & Amills, M. (2018). Genome-wide patterns of homozygosity provide clues about the population history and adaptation of goats. *Genetics Selection Evolution*, *50*(1), 1-12.

Bett, R. C., Kosgey, I. S., Kahi, A. K., & Peters, K. J. (2009). Analysis of production objectives and breeding practices of dairy goats in Kenya. Tropical Animal Health and Production, 41, 307-320.

Bilal, G., Moaeen-ud-Din, M., Khan, M., & Reecy, J. (2018). Productive and reproductive performance of goat breeds of Sindh. Journal of Animal Science, 96, 258-259.

Bosse, M., Megens, H. J., Madsen, O., Paudel, Y., Frantz, L. A., Schook, L. B. & Groenen, M. A. (2012). Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. *PLoS Genet*, *8*(11), e1003100.

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

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23(19)*, 2633-2635.

Briefer, E. F., Haque, S., Baciadonna, L., & McElligott, A. G. (2014). Goats excel at learning and remembering a highly novel cognitive task. *Frontiers in Zoology*, *11*(1), 1-12.

Brito, L. F., Jafarikia, M., Grossi, D. A., Kijas, J. W., Porto-Neto, L. R., Ventura, R. V., Salgorzaei, M., & Schenkel, F. S. (2015). Characterization of linkage disequilibrium, consistency of gametic phase and admixture in Australian and Canadian goats. *BMC Genetics*, *16*(1), 1-15.

Brito, L. F., Kijas, J. W., Ventura, R. V., Sargolzaei, M., Porto-Neto, L. R., Cánovas, A., Schenkel, F. S., Feng, Z., Jafarikia, M., & Schenkel, F.S. (2017). Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers. *BMC Genomics*, *18*(1), 1-20.

Burren, A., Signer-Hasler, H., Neuditschko, M., Tetens, J., Kijas, J., Drögemüller, C., & Flury, C. (2014). Fine-scale population structure analysis of seven local Swiss sheep breeds using genome-wide SNP data. *Animal Genetic Resources/Ressources génétiques animales/Recursos genéticos animales*, *55*, 67-76.

Burstyn-Cohen, T., Heeb, M. J., & Lemke, G. (2009). Lack of protein S in mice causes embryonic lethal coagulopathy and vascular dysgenesis. *The Journal of Clinical Investigation*, *119*(10), 2942-2953.

Cardoso, T.F., Amills, M., Bertolini, F., Rothschild, M., Marras, G., Boink, G., & Lenstra, J. A. (2018). Patterns of homozygosity in insular and continental goat breeds. *Genetics Selection Evolution*, *50*(1), 1-11

Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics*, *19*(4), 220.

Chan, W. Y., Hoffmann, A. A., & van Oppen, M. J. (2019). Hybridization as a conservation management tool. *Conservation Letters*, *12*(5), e12652.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, *4*(1), 1-16.

Charlesworth, B. (2012). The effects of deleterious mutations on evolution at linked sites. *Genetics*, *190*(1), 5-22.

Chenyambuga, S. W., Watts, P. C., Hirbo, J., Kemp, S. J., Hanotte, O., Kifaro, G. C. & Petersen, P. H. (2002). Analysis of Genetic Diversity and Relationships of Tanzania local Goat

Populations using Microsatellite DNA Markers. *Tanzania Journal of Agricultural Sciences*, 5(1).

Chenyambuga, S. W., & Lekule, F. P. (2014). Breed preference and breeding practices for goats in agro-pastoral communities of semi-arid and sub-humid areas in Tanzania. *Livestock Research for Rural Development, 26*, 1-14.

Chitneedi, P. K., Arranz, J. J., Suarez-Vega, A., García-Gámez, E., & Gutiérrez-Gil, B. (2017). Estimations of linkage disequilibrium, effective population size and RoHs-based inbreeding coefficients in Spanish Churra sheep using imputed high-density SNP genotypes. *Animal Genetics*, *48*(4), 436-446

Colli, L., Milanesi, M., Talenti, A., Bertolini, F., Chen, M., Crisà, A., Daly, K.G., Del Corvo, M., Guldbrandtsen, B., Lenstra, J.A. & Rosen, B.D. (2018). Genome-wide SNP profiling of worldwide goat populations reveals strong partitioning of diversity and highlights post-domestication migration routes. *Genetics Selection and Evolution*, *50*(1), 1-20.

Corbin, L. J., Liu, A. Y. H., Bishop, S. C., & Woolliams, J. A. (2012). Estimation of historical effective population size using linkage disequilibria with marker data. *Journal of Animal Breeding and Genetics*, *129*(4), 257-270.

Curik, I., Ferenčaković, M., & Sölkner, J. (2014). Inbreeding and runs of homozygosity: a possible solution to an old problem. *Livestock Science, 166*, 26-34.

Deng, T., Liang, A., Liu, J., Hua, G., Ye, T., Liu, S. & Yang, L. (2019). Genome-wide SNP data revealed the extent of linkage disequilibrium, persistence of phase and effective population size in purebred and cross-bred buffalo populations. *Frontiers in Genetics*, *9*, 688.

Derks, M. F., Megens, H. J., Bosse, M., Lopes, M. S., Harlizius, B., & Groenen, M. A. (2017). A systematic survey to identify lethal recessive variation in highly managed pig populations. *BMC Genomics*, *18*(1), 1-12.

Fabbri, M. C., Dadousis, C., & Bozzi, R. (2020). Estimation of linkage disequilibrium and effective population size in three Italian autochthonous beef breeds. *Animals*, *10*(6), 1034.

Falconer, D. S., & Mackay, T. F. C. (1996). Introduction to quantitative genetics. Longman, Harlow, England. *Introduction to Quantitative Genetics. 4th ed. Longman, Harlow, England.* Correctly cite books and book chapters

FAO. (2011). Guideline for Molecular Genetic Characterisation of Animal Genetic resources.

FAOSTAT. (2022). https://www.fao.org/faostat/en/#data/QCL 28/02/2024; 15:10pm

Faruque, S., Chowdhury, S. A., Siddiquee, N. U., & Afroz, M. A. (2010). Performance and genetic parameters of economically important traits of Black Bengal goat. *Journal of the Bangladesh Agricultural University*, *8*(1), 67-78.

Flury, C., Tapio, M., Sonstegard, T., Drögemüller, C., Leeb, T., Simianer, H., ... & Rieder, S. (2010). Effective population size of an indigenous Swiss cattle breed estimated from linkage disequilibrium. *Journal of Animal Breeding and Genetics*, *127*(5), 339-347.

Gaughan, J. B., Sejian, V., Mader, T. L., & Dunshea, F. R. (2019). Adaptation strategies: ruminants. *Animal Frontiers*, *9*(1), 47-53.

Getachew, T. A., Haile, A., Mészáros, G., Rischkowsky, B., Huson, R. H. J., Gizaw, S., Wurzinger, M., Mwai, A.O., & Sölkner, J. (2020). Genetic diversity, population structure and runs of homozygosity in Ethiopian short fat-tailed and Awassi sheep breeds using genome-wide 50k SNP markers. *Livestock Science, 1(232),* 1-40: 103899.

Githui, E. K., Kibegwa, F. M., Kamau, J. M., Mutura, S. K., Okwany, Z. A., Ngigi, D. M., & Mwangi, E. W. (2016). Genetic relationships of indigenous goats reared by pastoralists in Kenya based on mitochondria D-loop sequence. *Animal Genetic Resources/Resources génétiques animales/Recursos genéticos animales*, *59*, 73-80.

Groeneveld, L. F., Lenstra, J. A., Eding, H., Toro, M. A., Scherf, B., Pilling, D., ... & Globaldiv Consortium. (2010). Genetic diversity in farm animals–a review. *Animal genetics*, *41*, 6-31.

Grossen, C., Guillaume, F., Keller, L. F., & Croll, D. (2020). Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. *Nature Communications*, *11*(1), 1-12.

Haile, A., Getachew, T., Mirkena, T., Duguma, G., Gizaw, S., Wurzinger, M. & Rischkowsky, B. (2020). Community-based sheep breeding programs generated substantial genetic gains and socio-economic benefits. *Animal*, *14*(7), 1362-1370.

Harrison, R. G. (1993). Hybrids and hybrid zones: historical perspective. Hybrid Zones and the Evolutionary. In Harrison, R. G. (ed.), *Hybrid Zones and the Evolutionary Process* (New York, NY, 1993; online edn, Oxford Academic

Henn, B. M., Botigué, L. R., Bustamante, C. D., Clark, A. G., & Gravel, S. (2015). Estimating the mutation load in human genomes. *Nature Reviews Genetics*, *16*(6), 333-343.

Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., & Flicek, P. (2021). Ensembl 2021. *Nucleic Acids Research*, *49*(D1), D884-D891.

Howrigan, D. P., Simonson, M. A., & Keller, M. C. (2011). Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. *BMC Genomics*, *12*(1), 1-15.

Humann, J. L., Lee, T., Ficklin, S., & Main, D. (2019). Structural and functional annotation of eukaryotic genomes with GenSAS. In *Gene Prediction* (pp. 29-51). Humana, New York, NY.

Islam, R., Li, Y., Liu, X., Berihulay, H., Abied, A., Gebreselassie, G., & Ma, Y. (2019). Genome-wide runs of homozygosity, effective population size, and detection of positive selection signatures in six Chinese goat breeds.*Genes*, *10*(11), 938. doi: 10.3390/genes10110938.

Joazeiro, C. A. (2019). Mechanisms and functions of ribosome-associated protein quality control. *Nature Reviews Molecular Cell Biology*, *20*(6), 368-383.

Kahi, A. K., & Wasike, C. B. (2019). Dairy goat production in sub-Saharan Africa: current status, constraints and prospects for research and development. *Asian-Australasian Journal of Animal Sciences*, *32*(8), 1266-1274. doi: 10.5713/ajas.19.0377.

Karnuah, A. B., Dunga, G., & Rewe, T. (2018). Community-based breeding program for improving goat production in Liberia. *MOJ Current Research and Reviews*, *1*(5), 216-221.

Kaumbata, W., Nakimbugwe, H., Nandolo, W., Banda, L. J., Mészáros, G., Gondwe, T., Woodward-Greene, M.J., Rosen, B.D., Van Tassell, C.P., Solkner, J. & Wurzinger, M. (2021). Experiences from the Implementation of Community-Based Goat Breeding Programs in Malawi and Uganda: A Potential Approach for Conservation and Improvement of Indigenous Small Ruminants in Smallholder Farms. *Sustainability*, *13*(3), 1494.

Khandoker, M. A. M. Y., Afini, N., & Azwan, A. (2018). Productive and reproductive performance of Saanen goat at AZ-Zahra farm of Sandakan in Malaysia. *Bangladesh Journal of Animal Science*, *47*(1), 1-12.

Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Porto Neto, L. R., San Cristobal, M., Servin, B., McCulloch, R., Whan, V., Gietzen, K., & Paiva, S. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology*, *10*(2), e1001258.

Kim, E. S., Elbeltagy, A. R., Aboul-Naga, A. M., Rischkowsky, B., Sayre, B., Mwacharo, J. M., & Rothschild, M. F. (2016). Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity*, *116*(3), 255-264.

Kivila, J. N. (2018). *Genetic characterization of the indigenous Maasai goats (capra hircus) in Narok and Kajiado counties of Kenya using mitochondrial DNA control region*. [PhD thesis, University of Nairobi, Kenya].

Koopaee, H. K., & Koshkoiyeh, A. E. (2014). SNPs Genotyping technologies and their applications in farm animals breeding programs. *Brazilian Archives of Biology and Technology*, *57*(1), 87-95.

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, *4*(7), 1073-1081

Kumar, N. A., Shashank, C. G., & Banakar, P. S. (2018). Runs of Homozygosity and Genomic Inbreeding Concepts in Dairy Animal Breeding. *International Journal of Current Microbiology and Applied Sciences*, *7*(8), 688-693.

Lamy, E., van Harten, S., Sales-Baptista, E., Guerra, M. M. M., & de Almeida, A. M. (2012). Factors influencing livestock productivity. *Environmental Stress and Amelioration in Livestock Production*, 19-51.

Langbein, J., Siebert, K., & Nuernberg, G. (2008). Concurrent recall of serially learned visual discrimination problems in dwarf goats (Capra hircus). *Behavioural Processes*, *79*(3), 156-164.

Laouadi, M., Tennah, S., Kafidi, N., Antoine-Moussiaux, N., & Moula, N. (2018). A basic characterisation of small-holders' goat production systems in Laghouat area, Algeria. *Pastoralism*, *8*(1), 1-8.

Liu, S. (2010). Distant hybridization leads to different ploidy fishes. *Science China Life Sciences*, *53*, 416-425.

Luigi-Sierra, M. G., Cardoso, T. F., Martínez, A., Pons, A., Bermejo, L. A., Jordana, J., Delgado, J.V., Ugarte, E., Arranz, J.J. & Casellas J. (2019). Low genome-wide homozygosity in 11 Spanish ovine breeds. *Animal Genetics*, *50*(5), 501-511.

Luigi-Sierra, M. G., Fernández, A., Martínez, A., Guan, D., Delgado, J. V., Álvarez, J. F., Landi, V., Such, F.X., Jordana, J., Saura, M., & Amills, M. (2022). Genomic patterns of

homozygosity and inbreeding depression in Murciano-Granadina goats. *Journal of Animal Science and Biotechnology*, *13*(1), 1-14.

Makanjuola, B. O., Maltecca, C., Miglior, F., Marras, G., Abdalla, E. A., Schenkel, F. S., & Baes, C. F. (2021). Identification of unique ROH regions with unfavorable effects on production and fertility traits in Canadian Holsteins. *Genetics Selection Evolution*, *53*(1), 68. doi: 10.1186/s12711-021-00660-z.

Makino, T., Rubin, C. J., Carneiro, M., Axelsson, E., Andersson, L., & Webster, M. T. (2018). Elevated proportions of deleterious genetic variation in domestic animals and plants. *Genome Biology and Evolution*, *10*(1), 276-290.

Manirakiza, J., Hatungumukama, G., Besbes, B., & Detilleux, J. (2020). Characteristics of smallholders' goat production systems and effect of Boer cross-breeding on body measurements of goats in Burundi. *Pastoralism*, *10*(1), 2.

Marete, A. G., Jung'a, J. O., & Mosi, R. O. (2011). An analysis of inbreeding levels and factors affecting growth and reproductive traits in the Kenya Alpine dairy goat. *Livestock Research for Rural Development*, *23*(12), 259. http://www.lrrd.org/lrrd23/12/mare23259.htm

Marras, G., Gaspa, G., Sorbolini, S., Dimauro, C., Ajmone-Marsan, P., Valentini, A. & Macciotta, N. P. (2015). Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. *Animal Genetics*, *46*(2), 110-121.

Marsden, C. D., Ortega-Del Vecchyo, D., O'Brien, D. P., Taylor, J. F., Ramirez, O., Vilà, C., & Lohmueller, K. E. (2016). Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. *Proceedings of the National Academy of Sciences*, *113*(1), 152-157.

Mastrangelo, S., Tolone, M., Sardina, M. T., Sottile, G., Sutera, A. M., Di Gerlando, R., & Portolano, B. (2017). Genome-wide scan for runs of homozygosity identifies potential candidate genes associated with local adaptation in Valle del Belice sheep. *Genetics Selection Evolution*, *49*(1), 1-10.

Mastrangelo, S., Ciani, E., Sardina, M. T., Sottile, G., Pilla, F., Portolano, B., & Bi. Ov. Ita Consortium. (2018). Runs of homozygosity reveal genome-wide autozygosity in Italian sheep breeds. *Animal Genetics*, *49*(1), 71-81.

Mastrangelo, S., Di Gerlando, R., Sardina, M. T., Sutera, A. M., Moscarelli, A., Tolone, M., Cortellari, M., Marletta, D., Crepaldi, P., & Portolano, B. (2021). Genome-Wide Patterns

of Homozygosity Reveal the Conservation Status in Five Italian Goat Populations. *Animals*, *11*(6), 1510.

Maynard, J., & Haigh, J. (2007). The hitch-hiking effect of a favourable gene. *Genetics Research*, *89*(5-6), 391-403.

McQuillan, R., Leutenegger, A. L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A. and MacLeod A.N. (2008). Runs of homozygosity in European populations. *The American Journal of Human Genetics*, *83*(3), 359-372.

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

Mdladla, K., Dzomba, E. F., & Muchadeyi, F. C. (2018). Landscape genomics and pathway analysis to understand the genetic adaptation of South African indigenous goat populations. *Heredity*, *120*(4), 369-378.

Meuwissen, T. (2009). Genetic management of small populations: A review. *Acta Agriculturae Scand Section A*, *59*(2), 71-79.

Michailidou, S., Tsangaris, G. T., Tzora, A., Skoufos, I., Banos, G., Argiriou, A., & Arsenos, G. (2019). Analysis of genome-wide DNA arrays reveals the genomic population structure and diversity in autochthonous Greek goat breeds. *PloS One*, *14*(12), e0226179. doi: 10.1371/journal.pone.0226179

Michot, P., Chahory, S., Marete, A., Grohs, C., Dagios, D., Donzel, E., & Capitan, A. (2016). A reverse genetic approach identifies an ancestral frameshift mutation in RP1 causing recessive progressive retinal degeneration in European cattle breeds. *Genetics Selection Evolution*, *48*(1), 1-15.

Milano, M. (2019). Gene prioritization tools. In *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* (pp. 907-914).

Monau, P. I., Visser, C., Muchadeyi, F. C., Okpeku, M., Nsoso, S. J., & Van Marle-Köster, E. (2020a). Population structure of indigenous southern African goats based on the Illumina Goat50K SNP panel. *Tropical Animal Health and Production*, *52*(4), 1795-1802.

Monau, P., Raphaka, K., Zvinorova-Chimboza, P., & Gondwe, T. (2020b). Sustainable utilisation of indigenous goats in Southern Africa. *Diversity*, *12*(1), 20. https://doi.org/10.3390/d12010020

Msanga, Y. N., Mwakilembe, P. L., & Sendalo, D. (2012). The indigenous cattle of the southern highlands of Tanzania: Distinct phenotypic features, performance, and uses. *Livestock Research for Rural Development, 24*, 1-14.

Muasya, T. K., Peters, K. J., & Kahi, A. K. (2013). Breeding structure and genetic variability of the Holstein Friesian dairy cattle population in Kenya. *Animal Genetic Resources/Resources génétiques animales/Recursos genéticos animales*, *52*, 127-137.

Mueller, J. P., Rischkowsky, B., Haile, A., Philipsson, J., Mwai, O., Besbes, B. & Wurzinger, M. (2015). Community-based livestock breeding programmes: essentials and examples. *Journal of Animal Breeding and Genetics*, *132*(2), 155-168.

Muema, E. K., Wakhungu, J. W., Hanotte, O., & Han, J. (2009). Genetic diversity and relationship of indigenous goats of Sub-saharan Africa using microsatellite DNA markers, *Livestock Research for Rural Development*, 21(28). http://www.lrrd.org/lrrd21/2/muem21028.htm

Muigai, A. W., Okeyo, A. M., & Ojango, J. M. (2017). Goat production in Eastern Africa: Practices, breed characteristics, and opportunities for their sustainability. In *Sustainable Goat Production in Adverse Environments: Volume I* (pp. 31-57). Springer, Cham.

Mwangi, S. I., Muasya, T. K., Ilatsia, E. D., & Kahi, A. K. (2020). Effect of controlling future rate of inbreeding on expected genetic gain and genetic variability in small livestock populations. *Animal Production Science*, *60*(14), 1681-1686.

Nandolo, W., Wurzinger, M., Mészáros, G., Van Tassell, C., Gondwe, T., Mulindwa, H., & Sölkner, J. (2016). Identification of breeding objectives in community-based goat breeding programmes in Malawi. *Acta Agriculturae Slovenica*, *5*(104).

Nandolo, W., Mészáros, G., Banda, L. J., Gondwe, T. N., Lamuno, D., Mulindwa, H. A. & Sölkner, J. (2019). Timing and extent of inbreeding in African goats. *Frontiers in Genetics*, *10*, 537. doi: 10.3389/fgene.2019.00537.

Ndeke, A. N., Mutembei, H. M., Tsuma, V. T., & Mutiga, E. R. (2015). Reproductive Performance of the Galla and Toggenburg Goats and their Crosses in Mwingi Sub-county of Kenya. *Journal of Agricultural Science and Food Technology, 1(6)*, 78-83.

Ngeno, K., Van Der Waaij, E. H., Megens, H. J., Kahi, A. K., Van Arendonk, J. A. M., & Crooijmans, R. P. M. A. (2015). Genetic diversity of different indigenous chicken ecotypes using highly polymorphic MHC-linked and non-MHC microsatellite markers. *Animal Genetic Resources/Resources*, *56*, 1-7.

Ngeno, K. (2018). Variant Calling pipeline for Next Generation Sequence Data–A review. *Journal of Animal Science and Veterinary Medicine*, *3*(4), 90-93.

Nguluma, A. S., Msalya, G., & Chenyambuga, S. W. (2016). Phenotypic variation among four populations of Small East African goats of Tanzania. *Livestock Research for Rural Developmnt*, *28*, 136. http://www.lrrd.org/lrrd28/8/ngul28136.html

Nguluma, A. S., Huang, Y., Zhao, Y., Chen, L., Msalya, G., Lyimo, C & Chenyambuga, S. W. (2018). Assessment of genetic variation among four populations of Small East African goats using microsatellite markers. *South African Journal of Animal Science*, *48*(1), 117-127.

Nomura, K., Yonezawa, T., Mano, S., Kawakami, S., Shedlock, A. M., Hasegawa, M., & Amano, T. (2013). Domestication process of the goat revealed by an analysis of the nearly complete mitochondrial protein-encoding genes. *PLoS One*, *8*(8), e67775. doi: 10.1371/journal.pone.0067775.

Obholz, K. L., Akopyan, A., Waymire, K. G., & MacGregor, G. R. (2006). FNDC3A is required for adhesion between spermatids and Sertoli cells. *Developmental Biology*, *298*(2), 498-513.

Ojango, J. M., Audho, J., Oyieng, E., Recha, J., Okeyo, A. M., Kinyangi, J., & Muigai, A. W. (2016). System characteristics and management practices for small ruminant production in "Climate Smart Villages" of Kenya. *Animal Genetic Resources/Resources Génétiques Animales/Recursos Genéticos Animales*, *58*, 101-110.

Onzima, R., Aheisibwe, A. R., Katali, B. K., Kanis, E., & van Arendonk, J. A. M. (2014). *Economic analysis of cross breeding programs for indigenous goat breeds in Uganda.* In 10th World Congress on Genetics Applied to Livestock Production (WCGALP), Vancouver, BC, Canada.

Onzima, R. B., Upadhyay, M. R., Mukiibi, R., Kanis, E., Groenen, M. A. M., & Crooijmans, R. P. M. A. (2018a). Genome-wide population structure and admixture analysis reveals weak differentiation among Ugandan goat breeds. *Animal Genetics*, *49*(1), 59-70.

Onzima, R. B., Gizaw, S., Kugonza, D. R., van Arendonk, J. A., & Kanis, E. (2018b). Production system and participatory identification of breeding objective traits for indigenous goat breeds of Uganda. *Small Ruminant Research, 163,* 51-59.

Onzima, R. B., Upadhyay, M. R., Doekes, H. P., Brito, L., Bosse, M., Kanis, E. & Crooijmans, R. P. (2018c). Genome-wide characterisation of selection signatures and runs of homozygosity in Ugandan goat breeds. *Frontiers in Genetics*, *9*, 318. doi: 10.3389/fgene.2018.00318.

Opoola, O., Banos, G., Ojango, J. M., Mrode, R., Simm, G., Banga, C. B. & Chagunda, M. G. G. (2020). Joint genetic analysis for dairy cattle performance across countries in sub-Saharan Africa. *South African Journal of Animal Science*, *50*(4), 507-520.

Otieno, O. G., Junga, J. O., Badamana, M., & Amimo, J. O. (2015). Indigenous knowledge used in breeding and management of Capra hircus populations in Kajiado and Makueni Counties, Kenya. *Open Journal of Genetics, 5,* 111-135. doi: 10.4236/ojgen.2015.53009.

Pausch, H., Schwarzenbacher, H., Burgstaller, J., Flisikowski, K., Wurmser, C., Jansen, S., Jung, S., Schnieke, A., Wittek, T., & Fries, R. (2015). Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle. *BMC Genomics 16* (1), 1-13.

Peacock, C., Ahuya, C. O., Ojango, J. M., & Okeyo, A. M. (2011). Practical crossbreeding for improved livelihoods in developing countries: the FARM Africa goat project. *Livestock Science*, *136*(1), 38-44.

Peacock, C., & Hastings, T. (2011). Meru dairy goat and animal healthcare project. *International Journal of Agricultural Sustainability*, *9*(1), 203-211.

Pereira, F., Queiros, S., Gusmao, L., Nijman, I. J., Cuppen, E., Lenstra, J. A., Consortium, E., Davis, S.J., Nejmeddine, F., & Amorim, A. (2009). Tracing the history of goat pastoralism: new clues from mitochondrial and Y chromosome DNA in North Africa. *Molecular Biology and Evolution*, *26*(12), 2765-2773.

Peripolli, E., Munari, D. P., Silva, M. V. G. B., Lima, A. L. F., Irgang, R., & Baldi, F. (2017). Runs of homozygosity: current knowledge and applications in livestock. *Animal Genetics*, *48*(3), 255-271.

Polak, G., Sosin, E., & Martyniuk, E. (2022). FAO Commission on Genetic Resources for Food and Agriculture: what it does and how it supports the livestock sector. *Animal Science and Genetics*, *18*(4).

Pragna, P., Chauhan, S. S., Sejian, V., Leury, B. J., & Dunshea, F. R. (2018). Climate change and goat production: Enteric methane emission and its mitigation. *Animals*, *8*(12), 235.

Purfield, D. C., McParland, S., Wall, E., & Berry, D. P. (2017). The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. *PLoS One*, *12*(5), 1-23.

Raheem, K. A., Basiru, A., Raji, L. O., & Odetokun, I. A. (2024). Productive performance of goat. In *Trends in Clinical Diseases, Production and Management of Goats*, (163-177). Academic Press.

Rahmatalla, S. A., Arends, D., Reissmann, M., Ahmed, A. S., Wimmers, K., Reyer, H., & Brockmann, G. A. (2017). Whole-genome population genetics analysis of Sudanese goats identifies regions harbouring genes associated with major traits. *BMC Genetics*, *18*(1), 1-10.

Rai, A., Bleimling, N., Vetter, I. R., & Goody, R. S. (2020). The mechanism of activation of the actin binding protein EHBP1 by Rab8 family members. *Nature Communications*, *11*(1), 1-16.

Ramukhithi, F. V., Lehloenya, K. C., Nephawe, K. A., Chokoe, T. C., Seshoka, M. M., Jonker, T., & Nedambale, T. L. (2019). Phenotypic Characterisation of South African Unimproved Indigenous and Tankwa Goats, *American Journal of Animal and Veterinary Sciences*, 14(4), 207-220.

Rebelato, A. B., & Caetano, A. R. (2018). Runs of homozygosity for autozygosity estimation and genomic analysis in production animals. *Pesquisa Agropecuária Brasileira*, *53*(9), 975-984.

Reinartz, S., & Distl, O. (2016). Validation of deleterious mutations in Vorderwald cattle. *PLoS One*, *11*(7), e0160013.

Scherf, B. D., & Pilling, D. (2015). The second report on the state of the world's animal genetic resources for food and agriculture, In *FAO Commission on Genetic Resources for Food and Agriculture Assessments.* Rome.FAO.

Semakula, J., Mutetikka, D., Kugonza, D. R., & Mpairwe, D. (2010). Comparison of breeding systems by smallholder goat keepers in the humid, sub-humid and semi-arid ecological zones of Uganda. *Agricultural Journal*, *5*(2), 89-97.

Sim, N. L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Research*, *40*(W1), W452-W457.

Sölkner, J., Ferencakovic, M., Gredler, B., & Curik, I. (2010). Genomic metrics of individual autozygosity, applied to a cattle population. In *Book of Abstracts of the 61st Annual Meeting of the European Association of Animal Production* (p. 306).

Sonstegard, T. S., Cole, J. B., VanRaden, P. M., Van Tassell, C. P., Null, D. J., Schroeder, S. G., ... & McClure, M. C. (2013). Identification of a nonsense mutation in CWC15 associated with decreased reproductive efficiency in Jersey cattle. *PloS One*, *8*(1), e54872.

Šprem, N., Salajpal, K., Safner, T., Đikić, D., Jurić, J., Curik, I., Đikić, M. & Cubric-Curik, V. (2014). Genetic analysis of hybridization between domesticated endangered pig breeds and wild Boer. *Livestock Science*, *162*, 1-4.

Sved, J. A. (1971). Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theoretical Population Biology*, *2*(2), 125-141.

Szmatoła, T., Gurgul, A., Jasielczuk, I., Ząbek, T., Ropka-Molik, K., Litwińczuk, Z., & Bugno-Poniewierska, M. (2019). A comprehensive analysis of runs of homozygosity of eleven cattle breeds representing different production types. *Animals*, *9*(12), 1024. doi: 10.3390/ani9121024.

Szpiech, Z. A., Xu, J., Pemberton, T. J., Peng, W., Zöllner, S., Rosenberg, N. A., & Li, J. Z. (2013). Long runs of homozygosity are enriched for deleterious variation. *The American Journal of Human Genetics*, *93*(1), 90-102.

Tarekegn, G. M., Wouobeng, P., Jaures, K. S., Mrode, R., Edea, Z., Liu, B., Zhang, W., Mwai. O.A., Dessie, T., Tesfaye, K., & Strandberg, E. (2019). Genome-wide diversity and demographic dynamics of Cameroon goats and their divergence from East African, North African, and Asian conspecifics. *PloS One*, *14*(4), e0214843.

Upadhyay, M. R., Chen, W., Lenstra, J. A., Goderie, C. R. J., MacHugh, D. E., Park, S. D. E., Magee, D.A., Matassino, D., Ciani, F., Megens, H.J., & Van Arendonk, J.A.M. (2017).

Genetic origin, admixture and population history of aurochs (Bos primigenius) and primitive European cattle. *Heredity*, *118*(2), 169-176.

VanRaden, P. M., Olson, K. M., Null, D. J., & Hutchison, J. L. (2011). Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *Journal of Dairy Science*, *94*(12), 6153-6161.

Vignal, A., Milan, D., SanCristobal, M., & Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, *34*(3), 275-305.

Visser, C., Lashmar, S. F., Van Marle-Köster, E., Poli, M. A., & Allain, D. (2016). Genetic diversity and population structure in South African, French and Argentinian Angora goats from genome-wide SNP data. *PloS One*, *11*(5), e0154353.

Waineina, R. W., Ngeno, K., Okeno, T. O., & Ilatsia, E. D. (2021, June). Genetic diversity and population structure among indigenous and imported goat breeds in Kenya. In *Genetic Resources*, *2*(3), 25-35.

Walter, W., Sánchez-Cabo, F., & Ricote, M. (2015). GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics*, *31*(17), 2912-2914.

Wang, J., Santiago, E., & Caballero, A. (2016). Prediction and estimation of effective population size. *Heredity*, *117*(4), 193-206.

Weinman, L. R., Solomon, J. W., & Rubenstein, D. R. (2015). A comparison of single nucleotide polymorphism and microsatellite markers for analysis of parentage and kinship in a cooperatively breeding bird. *Molecular Ecology Resources*, *15*(3), 502-511.

Wellmann, R., & Bennewitz, J. (2019). Key genetic parameters for population management. *Frontiers in Genetics*, *10*, 667. doi: 10.3389/fgene.2019.00667

Yadav, A., Sahu, J. D., Dubey, A. D., Gadpayle, R. D., Kiran Barwa, D., Kashyap, K., & Verma, U. (2019). An overview on species hybridization in animals. *International Journal of Fauna and Biological Studies*, *6*, 36-42.

Yakubu, A., Salako, A. E., Imumorin, I. G., Ige, A. O., & Kinyemi, M. O. A. (2010). Discriminant analysis of morphometric differentiation in the West African Dwarf and Red Sokoto goats. *South African Journal of Animal Science*, *40*(4), 381-387.

Yakubu, A. (2018). Computational molecular analysis of deleterious mutations in serum amyloid A3 gene in goats and cattle. *Nigerian Journal of Biotechnology*, *35*(2), 80-90.

Yandell, M., & Ence, D. (2012). A beginner's guide to eukaryotic genome annotation. *Nature Reviews Genetics*, *13*(5), 329-342.

Zergaw, N., Dessie, T., & Kebede, K. (2016). Indigenous breeding practices and selection criteria of goat owners in Konso and Meta-Robi districts, Ethiopia: implications for designing community-based breeding strategy. *Livestock Research for Rural Development*, *28*(7). http://www.lrrd.org/lrrd28/7/zerg28133.html

Zhang, H., Gao, S., Lercher, M. J., Hu, S., & Chen, W. H. (2012). EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Research*, *40*(1), 569-572.

Zhang, Q., Guldbrandtsen, B., Bosse, M., Lund, M. S., & Sahana, G. (2015). Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC Genomics*, *16*(1), 1-16.

Zheng, Z., Wang, X., Li, M., Li, Y., Yang, Z., Wang, X., & Jiang, Y. (2020). The origin of domestication genes in goats. *Science Advances*, *6* (21), eaaz5216. doi: 10.1126/sciadv.aaz5216.

APPENDICES

Appendix 1: Percentage of chromosomes covered by Runs of homozygosities

CHR	GALLA	TOGGENBURG	ALPINE	SAANEN
1	13.35	0.90	1.51	3.29
2	2.86	0.89	5.04	1.40
3	0.00	6.01	0.00	2.00
4	3.26	0.68	1.47	1.91
5	5.00	2.50	0.00	2.23
6	8.03	1.89	2.28	1.13
7	0.00	2.55	7.28	4.51
8	10.26	3.77	1.49	4.24
9	0.00	2.04	16.42	4.28
10	7.90	1.47	0.00	6.47
11	0.00	1.79	1.67	3.20
12	0.00	2.77	4.48	5.87
13	0.00	2.90	0.00	10.20
14	0.00	1.50	9.68	4.88
15	0.00	3.22	7.33	5.25
16	19.65	1.96	8.66	3.82
17	8.90	6.34	0.00	2.86
18	0.00	8.08	0.00	7.79
19	14.41	5.31	0.00	0.00
20	0.00	2.12	6.53	1.89
21	0.00	2.02	8.31	5.39
22	0.00	4.87	7.25	3.56
23	3.52	2.46	15.17	5.68
24	5.72	0.99	1.77	9.68
25	0.00	3.19	9.09	4.40
26	0.00	1.68	4.34	3.78
27	7.10	5.08	5.11	14.97
28	0.00	10.05	4.95	0.00

Appendix 2: Inbreeding coefficients per individual

IID	Total RoHs length of each individual (Mb)	Inbreeding Coefficient per individual
11_G002_11_G002	42.44	0.02
27_G004_27_G004	49.03	0.02
3_G001_3_G001	54.03	0.02
83_G011_83_G011	16.61	0.01
91_G012_91_G012	49.06	0.02
1_H014B_1_H014B	5.58	0.00
10_H003A_10_H003A	33.07	0.01
18_H004A1_18_H004A1	28.7	0.01
2_H005A_2_H005A	17.09	0.01
25_H006A_25_H006A	76.16	0.03
26_H015A_26_H015A	51.92	0.02
33_H003C_33_H003C	20.83	0.01
34_H024A_34_H024A	44.06	0.02
41_H023A_41_H023A	34.75	0.01
42_H010A_42_H010A	11.75	0.00
49_H023B_49_H023B	6.41	0.00
50_H034A_50_H034A	29.04	0.01
57_H010B_57_H010B	58.01	0.02
58_H009A_58_H009A	34.17	0.01
65_H003B_65_H003B	20.13	0.01
66_H018A_66_H018A	24.6	0.01
73_H002B_73_H002B	83.22	0.03
74_H007A_74_H007A	17.97	0.01
81_H002A_81_H002A	43.97	0.02
89_H014A_89_H014A	134.49	0.06
9_H006B_9_H006B	42.05	0.02
90_H004A_90_H004A	28.7	0.01
15_N010A_15_N010A	12.04	0.01
16_N033B_16_N033B	110.76	0.05
24_N020B_24_N020B	20.52	0.01
31_N008A_31_N008A	12.69	0.01
39_N024A_39_N024A	25.05	0.01
47_N008B_47_N008B	28.26	0.01
48_N023B_48_N023B	7.81	0.00
55_N021A_55_N021A	17.02	0.01
56_N035A_56_N035A	7.41	0.00
63_NO35B_63_NO35B	111.64	0.05
64_NO33A_64_NO33A	18.42	0.01
70_N030A_70_N030A	5.55	0.00

71_N010B_71_N010B	24.19	0.01
78_N029B_78_N029B	40.01	0.02
79_N022A_79_N022A	22.47	0.01
8_N007A_8_N007A	9.70	0.00
86_N028A_86_N028A	14.13	0.01
87_N020A_87_N020A	29.51	0.01
95_N034A_95_N034A	16.60	0.01
96_N022B_96_N022B	21.14	0.01
12_M022B_12_M022B	45.13	0.02
13_M038B_13_M038B	54.27	0.02
14_M023B_14_M023B	65.62	0.03
21_M010A_21_M010A	174.68	0.07
22_M037A_22_M037A	52.75	0.02
28_M007B_28_M007B	43.41	0.02
29_M001A_29_M001A	64.47	0.03
30_M021A_30_M021A	53.58	0.02
36_M036B_36_M036B	115.71	0.05
37_M005A_37_M005A	32.31	0.01
4_M008A_4_M008A	49.34	0.02
45_M013A_45_M013A	79.27	0.03
46_M036A_46_M036A	56.87	0.02
5_M031A_5_M031A	68.95	0.03
52_M024A_52_M024A	4.98	0.00
53_M031B_53_M031B	35.07	0.01
54_M007A_54_M007A	89.48	0.04
6_M008B_6_M008B	41.68	0.02
61_M032A_61_M032A	35.05	0.01
68_M11B_68_M11B	129.94	0.05
69_M001B_69_M001B	40.51	0.02
76_M012B_76_M012B	22.91	0.01
77_M012A_77_M012A	11.41	0.00
84_M018A_84_M018A	24.06	0.01
85_M004A_85_M004A	77.36	0.03
92_M005B_92_M005B	75.11	0.03
93_M011A_93_M011A	87.61	0.04

Appendix 3: Genomic regions with high frequency of Runs of homozygosities

IID	CHR	POS1	POS2	KB	NSNP	SNP % in RoHs
GAL	17	50496599	55334347	4837.75	107	2.211772459
GAL	16	62037458	67126223	5088.77	110	2.161624252
TOT	6	82922388	88129197	5206.81	126	2.419907775
TOT	17	23941309	28606126	4664.82	106	2.272328738
TOT	16	12818180	17369133	4550.95	102	2.24128831
TOT	27	600400	10512553	9912.15	222	2.239674646
TOT	24	44726551	52080591	7354.04	162	2.202870503
TOT	21	39316366	46004536	6688.17	147	2.197910311
TOT	2	121194945	127160014	5965.07	131	2.196118403
TOT	13	68573050	79910677	11337.6	248	2.187406396
TOT	21	40245597	47248076	7002.48	153	2.184940193
TOT	13	74766929	79910677	5143.75	112	2.17740018
TOT	28	27378852	32341157	4962.31	108	2.176407501
TOT	8	103241548	108299137	5057.59	110	2.174948938
TOT	24	45026095	51893292	6867.2	149	2.169735021
TOT	2	122540040	127934583	5394.54	117	2.168858017
TOT	24	10510797	16252899	5742.1	124	2.159487561
TOT	16	74304173	79353830	5049.66	109	2.158562025
TOT	2	29975669	35490029	5514.36	119	2.158001625
TOT	16	74251615	79353830	5102.22	110	2.155925974
ALP	28	39181	5758953	5719.77	128	2.237851048
ALP	6	80439756	87432628	6992.87	154	2.202242197
ALP	9	78193023	84279453	6086.43	133	2.18518866
ALP	1	58883659	63643118	4759.46	104	2.185121842
ALP	27	3855715	10045474	6189.76	135	2.181021558
SAA	6	85066602	89164238	4097.64	107	2.611261076
SAA	27	1051338	7735492	6684.16	152	2.274034639
SAA	1	50722301	55967635	5245.34	117	2.230553435
SAA	17	24404597	28958842	4554.25	101	2.217710681
SAA	2	121791284	127464085	5672.8	125	2.203496614
SAA	23	35428109	40516154	5088.05	112	2.201237961
SAA	8	103492446	112591777	9099.33	200	2.197963543
SAA	24	56109867	61291762	5181.9	113	2.180669006
SAA	14	66506418	74399957	7893.54	172	2.178997003

Appendix 4: Estimated ancestral effective population size over past generations

Generations Ago	Alpine	Galla	Saanen	Toggenburg
13	109	49	81	93
15	123	55	90	102
17	139	63	101	111
20	156	72	113	122
23	178	83	127	135
27	205	96	144	150
32	235	113	164	170
38	274	132	189	193
45	322	156	219	221
54	376	187	254	253
66	447	223	302	296
80	532	269	356	351
98	642	325	426	415
121	780	400	512	492
150	937	491	632	601
187	1168	610	766	730
234	1406	751	934	888
293	1685	930	1140	1073
367	1995	1136	1377	1282
454	2382	1326	1589	1543
553	2685	1553	1899	1794
658	3007	1761	2143	2020
758	3333	2036	2320	2317
845	3650	2087	2504	2452
914	3733	2334	2655	2609
959	3706	2296	2679	2632
983	3709	2428	2715	2548

Appendix 5: Genes identified from the genomic regions islands

Genotype	CHR	START	END	GENES	CONSEQUENCES
TOT	2	121194945	127160014	ZSWIM2, FSIP2,	Missense
TOT	2	122540040	127934583	ZSWIM2	Missense
TOT	2	29975669	35490029	ABCA12	Missense
SAA	8	103492446	112591777	MYT1L, MEGF9	Missense
GAL	17	50496599	55334347	NAA15	Synonymous
SAA	17	24404597	28958842	PIWIL1	Missense
TOT	17	23941309	28606126	PIWIL1	Missense
TOT	21	39316366	46004536	EAPP, AKAP6	Missense & Synonymous
TOT	21	40245597	47248076	EAPP, AKAP6	Missense & Synonymous
SAA	23	35428109	40516154	PNPLA1, ZNF76,	Synonymous
SAA	24	56109867	61291762	ATP8B1	Missense
ALP	27	3855715	10045474	RARB, TOP2B	Missense
SAA	27	1051338	7735492	RARB, TOP2B	Missense
TOT	27	600400	10512553	KAT6A, RARB, TOP2B	Missense
ALP	28	39181	5758953	C10orf71	Missense

Appendix 6: Publication
Chapter 2

International Journal of Veterinary Sciences and Animal Husbandry 2022; 7(4): 36-40

ISSN: 2456-2912
VET 2022; 7(4): 36-40

www.veterinarypaper.com
Received: 10-05-2022
Accepted: 12-06-2022

Theodora Chikoko
[a] Department of Animal Sciences, Egerton University, P.O. Box 536-20115, Nakuru, Kenya
[b] Department of Animal Health and Livestock Development, P.O. Box 2096, Lilongwe, Malawi

Ngeno Kiplangat
Department of Animal Sciences, Egerton University, P.O. Box 536-20115, Nakuru, Kenya

Thomas Muasya
Department of Animal Sciences, Egerton University, P.O. Box 536-20115, Nakuru, Kenya

Mutational load and deleterious mutations in goat genome from Kenya

Theodora Chikoko, Ngeno Kiplangat and Thomas Muasya

DOI: https://doi.org/10.22271/veterinary.2022.v7.i4a.429

Abstract
Different domestication processes such as selection and effective population size are among other factors responsible for the accumulation of mutations in an animal. This study investigated the mutation load and accumulation of deleterious mutations in Kenyan goat populations using Single Nucleotide Polymorphism (SNP) data obtained from local (Galla = 12) and exotic (Saanen = 24, Alpine = 28, and Toggenburg = 30) goat genotypes. The SNP annotation was done on the ENSEMBLE goat (*Capra hircus*) using the Variant Effect Predictor (VEP). Sorting Intolerant from Tolerant (SIFT) was done on the annotated file using a score of > 0.01 to describe highly deleterious mutations. Biomart tool in the ENSEMBL goat *C. hircus* was used for Gene Ontology (GO) for the genes identified from highly deleterious mutations. The analysis disclosed no differences in mutation load among the studied genotypes. Overall, the synonymous mutations were in abundance compared to missense mutations, totaling 693 and 258 representing 1.01% and 0.37%, respectively. The calculated mutation load was similar (0.37) suggesting exposure of goats to similar domestication processes that are creating similar effects on the animal's genome. Further analysis of the missense variants revealed 126 deleterious mutations and 132 tolerated mutations. A total of 111 genes were identified from the deleterious mutations and only 5 were among the highly deleterious mutations (SIFT score > 0.01) which include PROS1, EHBP1, LTN1, LRRN4, and FNDC3A. Gene Ontology describes these genes as responsible for different functions associated with animal reproduction, diseases, and memory. This study's results have confirmed the levels of mutation load and the presence of deleterious mutations in Kenyan goats. This can be used as a base to predict the future rate of mutation load to ensure better implementation of conservation programs to avoid an accumulation of deleterious mutations.

Keywords: Annotation, deleterious mutations, goats, genes, mutation load

Appendix 7: Tropentag Conference presentation abstract

Chapter 1

Tropentag 2021, September 15 - 17, hybrid conference, Germany

"Towards shifting paradigms in agriculture for a healthy and sustainable future"

Genetic Analysis of Hybridisation Pattern in Goat Genotypes from East Africa

Theodora Chikoko, Kiplangat Ngeno, Thomas Muasya

Egerton University, Animal Science, Kenya

Abstract

East Africa is home to an estimated 150,667,482 goats in different geographical regions which are exposed to diverse climatic, production and management conditions. This has shaped the goat genome due to adaptation or selection. This study aimed to investigate hybridisation pattern between distantly isolated goat populations in the East African community using single nucleotide polymorphism (SNP) data genotyped using a 50K goat SNP chip. In Uganda, SNP data for 144 goats from six genotypes sampled from five agro-ecological zones was retrieved from the Zenodo genome archive. In Kenya, 94 goats from four genotypes sampled from three regions were used. Quality control procedures were performed in PLINK v 1.9. Data from Kenya (48,303 SNPs) and Uganda (46,105 SNPs) was then merged in Tassel software resulting in 94,408 SNPs available for joint downward analysis. Principle component and phylogenetic analysis was used to visualise relationships between the studied populations. Four well-defined clusters, two from each country were observed with some visible outliers. From the two clusters in each country, genetically mixed genotypes were identified in only one cluster. This might suggest inbreeding within the genotypes due to close proximity between them. The degree of genetic differentiation measured using Fst ranges from 0.191 to 0.324 indicating that all the genotypes within a country are to some extent isolated from each other. Only Boer genotype from Uganda was highly isolated from all genotypes in this study. The diversity shown between Kenya and Uganda goat genotypes might be due to uncommon ancestry, isolation distance or lack of capacity to use reproductive technologies. These results will be useful in the implementation of future genetic conservation, utilisation and improvement programs in the East African community.

Keywords: Genotypes, inbreeding, single nucleotide polymorphism (SNP)

Contact Address: Theodora Chikoko, Egerton University, Animal Science, Egerton Universiy -Main Campus Po Box 536-20115 Egerton Kenya, Nakuru, Kenya, e-mail: theochikoko@gmail.com

Appendix 8: AFMA Conference presentation abstract
Chapter 3

RUNS OF HOMOZYGOSITY AND EFFECTIVE POPULATION SIZE FROM DIFFERENT GOAT GENOTYPES IN KENYA

Theodora chikoko [12], Thomas Muasya [1] and Ngeno Kiplangat [1]

Department of Animal Sciences, Egerton University, P.O. Box 635, Njoro, Kenya.

Department of Animal Health and Livestock Development, P.O.Box 2096, Lilongwe, Malawi.

*Correspondence to: theochikoko@gmail.com or Ngeno.kiplangat@egerton.ac.ke

Abstract

There is limited genetic information available in most African goat populations. However, improvements in genomic technologies have enabled easy and effective analysis of different genetic parameters. This study used Single Nucleotide Polymorphism (SNP) data of four goat genotypes from Kenya; Galla (n = 12), Alpine (n = 28), Saanen (n = 24) and Toggenburg (n = 30). After SNP quality control, 48808 SNPs that were available for analysis. Runs of homozygosities (ROHs) were detected to analyse the distribution and inbreeding coefficients. Across the genotypes, 348 RoHs were detected with the highest number observed in Toggenburg and lowest in Galla. RoHs that described recent inbreeding were observed in Alpine. The highest mean sum of ROH was observed on the long RoHs category (>16 Mb) which have a negative effect on animals performance. Distribution of RoHs per chromosome was breed-specific without a clear pattern across the genotype. Furthermore, 32 genomic regions with the high frequency of RoHs were detected. Sixteen genes (missense and synonymous) were identified to determine their effects on animal performance. High inbreeding coefficient values were observed in all exotic genotypes suggesting continuous use of few breeding bucks. Toggenburg was found to be the most inbred genotype with the highest breeding coefficient of 0.68 compared to other genotypes. To maintain and improve the genetic diversity in Kenya, these findings will be useful for the strategic implementation of genetic improvement and conservation programs to ensure continual contribution to food security.

Keywords: Genotype, Genome, Inbreeding coefficient, Runs of Homozygosity

Appendix 9: Research permit

REPUBLIC OF KENYA

NATIONAL COMMISSION FOR SCIENCE.TECHNOLOGY & INNOVATIO

Ref No:**712062**

Date of Issue:**18/November/2021**

RESEARCH LICENSE

This is to Certify that Miss.. THEODORA CHIKOKO of Egerton University, has been licensed to conduct researc Meru, Nyeri on the topic: GENETIC ANALYSIS OF HYBRIDISATION, MUTATION LOAD AND RUNS OF HOMOZYGOSITY OF GOATS FROM EAST AFRICA for the period ending : 18/November/2022.

License No:**NACOSTI/P/21/14174**

712062

Applicant Identification Number

Director General
NATIONAL COMMISSION FOR SCIENCE.TECHNOLOGY & INNOVATION

Verification OR Code

NOTE: This is a computer generated License. To verify the authenticity of this docu
Scan the QR Code using QR scanner application.