REPUBLIC OF TÜRKİYE
ERCİYES UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF AGRICULTRAL SCIENCE AND TECHNOLOGIES

INVESTIGATIONS ON THE GENETICS OF GROWTH AND LINEAR CONFORMATION TRAITS IN AKKARAMAN SHEEP

By
Mehmet KIZILASLAN

Supervisor
Prof. Dr. Mehmet Ulaş ÇINAR
Co-supervisor
Dr. Stephen N. WHITE

PhD Thesis

October 2022
KAYSERİ

REPUBLIC OF TÜRKİYE
ERCİYES UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF AGRICULTRAL SCIENCE AND TECHNOLOGIES

INVESTIGATIONS ON THE GENETICS OF GROWTH AND LINEAR CONFORMATION TRAITS IN AKKARAMAN SHEEP

(PhD Thesis)

By
Mehmet KIZILASLAN

Supervisor
Prof. Dr. Mehmet Ulaş ÇINAR
Co-supervisor
Dr. Stephen N. WHITE

This study was funded by the Scientific Research Projects Department of Erciyes University with Project ID: FDK-2021-10811

October 2022
KAYSERİ

SCIENTIFIC ETHICS CONFORMITY

I fully declare that all the information processed in this study has been obtained in accordance with the academic and ethical rules. Also, as required by these rules and behaviours, I declare that I have fully cited and referenced all the materials and results that are not at the core of this study.

Mehmet KIZILASLAN

Signature

CONCORDANCE WITH THE INSTRUCTION GUIDE

The thesis named **"Investigations on the Genetics of Growth and Linear Conformation Traits in Akkaraman Sheep"** has been prepared in accordance with the Erciyes University Graduate School of Natural and Applied Sciences Thesis Preparation and Writing Guide.

Prepared by	**Supervisor**
Mehmet KIZILASLAN	Prof. Dr. Mehmet Ulaş ÇINAR
signature	signature

Head of Department Agricultural Sciences and Technologies

Doç. Dr. Özhan ŞİMŞEK

signature

ACKNOWLEDGEMENTS

The project established the foundations of the thesis study was funded and financially assisted by the Scientific Research Projects Deparment of Erciyes University with the Project ID: FDK-2021-10811. Therefore, I would like to thank to the Scientific Research Projects Coordination Unit of Erciyes University for the opportunity provided.

It is a complete genuine pleasure for me to thank to my supervisors Prof. Mehmet Ulaş Çınar, and Dr. Stephen N. White for their professional attitude, kind interest, irreplaceable inspiration and strong guidance throughout the PhD journey.

I am also grateful to Asst. Prof. Dr. Sedat Behrem, Dr. Lindsay M.W. Piel and Yunus Arzik for their valuable scientific contribution throughout the process. I further acknowledge the contribution made by the Ankara Sheep and Goat Breeder's Association staff, Mr. Hamit Karakaya, Mr. Cagatay Yıldız and Ms. Simge Tutenk for their time and assistance during the data collection period.

The farms where animals of the study were managed were registered under the National Small Ruminant Breeding Programme. Therefore, my deepest gratitude goes to the farmers who actively contributed in the sampling and data collection period and the General Directorate of Agricultural Research and Policies (Ministry of Agriculture and Forestry) of the Republic of Türkiye, who fund and run the programme. Additionally, I would like to thank Genome and Stem Cell Center (GenKok) for providing laboratory facilities and funding. I am also grateful to the Genetics Laboratory of International Center for Livestock Research and Training (ICLRT), where instruments were used for genotyping.

Finally, I am deeply grateful to my family and friends for their love, support, encouragement and keen interest in my academic achievements.

Mehmet KIZILASLAN

October 2022, KAYSERİ

AKKARAMAN KOYUNLARDA BÜYÜME VE KONFORMASYON ÖZELLİKLERİNİN GENETİĞİ ÜZERİNE ARAŞTIRMALAR

Mehmet KIZILASLAN

Erciyes Üniversitesi, Fen Bilimleri Enstitüsü

Doktora Tezi, Eylül 2022

Danışman: Prof. Dr. Mehmet Ulaş ÇINAR

2. Danışman: Dr. Stephen N. WHITE

ÖZET

İklim değişikliğinin ilerleyişi 1950'lerden bu yana özellikle Türkiye ile kesişen bölgelerde artan tarımsal ve ekolojik kuraklık beklentileri ile en tahmin edilemez düzeylere ulaşmıştır. Yerli hayvan genetik kaynakları tahmin edilemez ekstrem koşullar altında sürdürülebilir üretim ve gıda güvenliğine katkıda bulunacak ve ihtiyaç duyacağımız potansiyel rezervuarları temsil etmektedir. Yetersiz besleme ile yüksek gelişim kabiliyeti ve ekstrem koşullara karşı dayanıklılığı ve yaşama kabiliyeti ile Akkaraman koyunu, değişen iklim ve çevre şartları, tekrarlı pandemiler ve diğer darboğazlara karşı geleceği beslemek için üretilen çözümlerin bir parçası olma açısından büyük potansiyel taşımaktadır. Büyüme ve doğrusal tip özellikleri karkas verimi, kalitesi, yemden yararlanma, üreme, dayanıklılık, hareketlilik, verimlilik ve hayatta kalma gibi özelliklerle olan ilişkisinden dolayı hayvansal üretim sistemlerinin karlılığı, verimliliği ve sürdürülebilirliği üzerinde doğrudan etkiye sahiptir. Bundan dolayı çalışmamız, büyüme ve doğrusal tip özelliklerinin Akkaraman koyun ırkında genetik ve fenotipik parametrelerini hesaplamayı, bu özelliklerin dağılımı ile ilişkili moleküler mekanizmaları düzenleyen QTL tespitini ve ırkın yaygın kullanılan bazı kültür ırkları ve farklı coğrafyalardan yerli koyun ırkları ile genom-boyu SNP paneller vasıtasıyla karşılaştırmalı genomik karakterizasyonunu hedeflemiştir. Genel olarak orta düzey kalıtım dereceleri ve özellikler arasında yüksek düzeyde pozitif genetik ve fenotipik korelasyon tespit edilmiştir. Buna ek olarak yapılan GWAS analizleri sonucunda, özelliklerle ilişkili 2 genom-çapında 18 kromozom-çapında SNP tespit

edilmiştir. Buna göre, *ZNF407, SCD5, SP1, PTGDR, PRDM2, PTPRG, KCND2, ZNF260, CPE*, *GRID2, SPIDR, HCN3, MYOM3, FKBP1A, TLE4, SLC44A1* ve *TMEM50A* genleri muhtemel QTL barındıran genler olarak tavsiye edilmiştir. Diğer taraftan, Akkaraman koyun ırkının diğer ırklarla olan genetik benzerliği ve farklılıkları ortaya konmuştur. Akkaraman ırkındaki mevcut yüksek genetik çeşitlilik, uygulanacak bir kapsamlı seleksiyon programının yakın zamanda soy-içi üreme depresyonuna sebep olacak düzeyde bir tehlike içermediğini desteklemektedir. Dolayısıyla, sonuçlarımız koyun üretim sistemleri için önemli uygulamalar taşımakta ve verimli ve başarılı bir genoma dayalı seleksiyon programı oluşturma konusunda ise pratikte kullanılacak bilgiler sağlamaktadır.

Anahtar Kelimeler: Kalıtım derecesi, genetik korelasyon, büyüme, doğrusal tip sınıflandırması, GWAS, kantitatif özellik lokusları, genetik çeşitlilik, bağlantı dengesizliği, homozigotluk haritalama.

INVESTIGATIONS ON THE GENETICS OF GROWTH AND LINEAR CONFORMATION TRAITS IN AKKARAMAN SHEEP

Mehmet KIZILASLAN

Erciyes University, Graduate School of Natural and Applied Sciences

PhD Thesis, October 2022

Supervisor: Prof. Dr. Mehmet Ulaş ÇINAR

Co-Supervisor: Dr. Stephen N. WHITE

ABSTRACT

The progression of climate change since 1950s is at an unprecedented level, with increased risks of agricultural and ecological drought expectations especially for regions overlapping Türkiye. Domestic animal genetic resources represent potential reservoirs that may contribute sustainable production for food security of those unpredictable times and extreme circumstances. With its hardy nature, ability to thrive on meagre diet and resilience under extreme conditions, Akkaraman sheep appear to have a great potential as part of the solution against feeding the future under changing, recurrent pandemics and other unpredicted bottlenecks. Growth and linear type traits, as indicators of productivity and conformation appraisal, have direct influence on the profitability, efficiency and sustainability of production systems due to their interplay with carcass yield and quality, feed efficiency, reproduction, robustness, mobility, productivity, longevity and survivability traits. Therefore, our study aimed to uncover genetic and phenotypic parameters for growth and linear type traits, discover QTL underlying complex mechanisms determining the phenotypic distribution of those traits and uncover certain population genetic parameters of Akkaraman sheep in comparison with other transboundary and indigenous sheep breeds by using genome-wide dense SNP panels. Moderate genomic heritabilities and overall high and positive genetic/phenotypic correlations were observed among most of the growth and type traits. Additionally, 2 genome- and 18 chromosome-wide significant SNPs were

discovered to be associated with the traits after GWA analyses. Accordingly, *ZNF407, SCD5, SP1, PTGDR, PRDM2, PTPRG, KCND2, ZNF260, CPE*, *GRID2, SPIDR, HCN3, MYOM3, FKBP1A, TLE4, SLC44A1* and *TMEM50A* were proposed as the genes bearing putative QTL for growth and linear type traits. On the other hand, we have displayed genetic similarities and differences of Akkaraman sheep against various breeds. The high genetic diversity found in Akkaraman sheep validates the existence of a large genetic variance to implement a comprehensive selection on the breed, without a great danger of imposing an inbreeding depression in the near future. Therefore, our results carry important implications for sheep production systems and provide practical knowledge that will contribute to designing efficient and successful genome-based selection programmes.

Keywords: Heritability, genetic correlations, growth, linear type traits, GWAS, quantitative trait loci, genetic diversity, linkage disequilibrium, runs of homozygosity.

CONTENTS

INVESTIGATIONS ON THE GENETICS OF GROWTH AND LINEAR CONFORMATION TRAITS IN AKKARAMAN SHEEP

SCIENTIFIC ETHICS CONFORMITY ii

CONCORDANCE WITH THE INSTRUCTION GUIDE iii

APPROVAL iv

ACKNOWLEDGEMENTS v

ÖZET vi

ABSTRACT viii

CONTENTS x

ABBREVIATIONS xii

LIST OF TABLES xv

LIST OF FIGURES xvi

INTRODUCTION 1

CHAPTER 1

LITERATURE REVIEW

1.1. Genetics of Growth and Conformation Traits in Akkaraman Sheep 5

1.2. Genomic Analysis of Akkaraman Sheep 8

1.3. Objectives of the Study 13

CHAPTER 2

MATERIAL AND METHODS

2.1. Ethics Statement 15

2.2. Animals and Phenotyping 15

2.3. Genotyping and Quality Control 17

2.4. Statistical & Bioinformatics Analysis 19

2.4.1 Estimation of Genetic Parameters 19

2.4.2. Genome-wide Association Analysis (GWAS) 21

2.4.3. Functional Annotation of Candidate Genes 22

2.4.4. Genetic Diversity Indices 23

2.4.5. Genetic Distance & Population Structure 23

2.4.6. Linkage Disequilibrium Analysis 23

2.4.7. Genomic Inbreeding and Runs of Homozygosity Analyses 24

CHAPTER 3

RESULTS

3.1. Phenotypic statistics 25

3.2. Genetic parameters 27

3.3. Genome-wide association studies 30

3.4. Annotation of Candidate Genes 32

3.5. Genetic diversity indices 37

3.6. Genetic distance & structure among the breeds 39

3.7. Linkage Disequilibrium (LD) and LD decay 43

3.8. Detection of Runs of homozygosity (ROH) 44

CHAPTER 4

DISCUSSION-CONCLUSIONS AND FUTURE PERSPECTIVES

4.1. Discussion 46

4.1.1. Genetic Parameters 47

4.1.2. Genome-wide Association Studies 48

4.1.3. Comparative Genomic Characterization 52

4.2. Conclusions 57

4.3 Future Perspectives 58

5. REFERENCES 63

6. CURRICULUM VITAE 83

ABBREVIATIONS

QTL: Quantitative Trait Loci

SNP: Single Nucleotide Polymorphism

GWAS: Genome-wide Association Study

FAO: Food and Agriculture Organisation of the United Nations

BLUP: Best Linear Unbiased Prediction

OAR: *Ovis aries* chromosome

EBV: Estimated Breeding Value

gEBV: Genomic Estimated Breeding Value

GLM: Generalised Linear Model

IBS: Identity by State

QC: Quality Control

PCA: Principal Component Analysis

MVN: Multivariate Normal Distribution

GRM: Genomic Relationship Matrix

REML: Restricted Maximum Likelihood

EM: Expectation-Maximization

KEGG: Kyoto Encyclopedia of Genes and Genomes

DAVID: The Database for Annotation, Visualization and Integrated Discovery

HWE: Hardy-Weinberg Equilibrium

H_O: Observed Heterozygosity

H_E: Expected Heterozygosity

F_{ST}: Fixation Index

LD: Linkage Disequilibrium

MAF: Minor Allele Frequency

ROH: Runs of Homozygosity

BW: Birth Weight

WW: Weaning Weight

180DW: 180 Days Weight

preADG: Pre-weaning Average Daily Gain

postADG: Post-weaning Average Daily Gain

180ADG: 180-days Average Daily Gain

BCS: Body Condition Score

TS: Tail Size

RLRV: Rear Legs Rear View

GRRV: Gigot Roundness Rear View

RW: Rump Width

RLW: Rear Legs Width

RLFA: Rear Legs Feet Angle

GRSV: Gigot Roundness Side View

RLSV: Rear Legs Side View

BL: Body Length

CW: Chest Width

LIST OF TABLES

Table 1. Descriptive statistics of growth and linear type traits. 25
Table 2. Heritability estimates, genetic and phenotypic correlations of growth and linear type traits in Akkaraman sheep. 28
Table 3. Significant SNPs from GWAS of growth traits. 30
Table 4. Significant SNPs from GWAS of linear type traits. 31
Table 5. Genetic diversity parameters for different sheep breeds. 37
Table 6. Pairwise FST values between the sheep breeds. 41
Table 7. Descriptive statistics of ROH and LD in different sheep breeds. 45

LIST OF FIGURES

Figure 1. Manhattan plots of growth traits in Akkaraman sheep (−log10 (SNP p-values)). Horizontal red solid line represent genome-wide (5.90) and black dashed lines represent chromosome-wide significance thresholds (4.49).....34
Figure 2. Manhattan plots of type traits in Akkaraman sheep (−log10 (SNP p-values)). Horizontal red solid line represent genome-wide (5.90) and black dashed lines represent chromosome-wide significance thresholds (4.49).....35
Figure 3. Quantile-Quantile (Q-Q) plot of the GWA analyses. 36
Figure 4. HWE deviation plot. Pink lines represent loci violating Hardy-Weinberg Equilibrium within each breed. 39
Figure 5. Principal Components Analysis (PCA) of the breeds.....40
Figure 6. Phylogenetic Neighbor-Joining (NJ) tree of the breeds.....42
Figure 7. Trends in linkage disequilibrium (LD) decay in 500 Kb distance.....43

INTRODUCTION

The livestock domestication symbolize a crucial leap in recent human history. Sheep is suggested to be the first domesticated grazing livestock that is first bred around the Fertile Crescent almost 11,000 years ago (Wei et al., 2015). Currently, more than 1200 extant sheep breeds have been diverged from the ancestors of the modern sheep through natural/artificial selection, mutation, genetic drift as well as migration and genetic isolation (FAO, 2015). However, due to socioeconomic preferences, a few dozens of mainstream transboundary sheep breeds with superior productivity are set to surpass most of the locally adapted landrace sheep during the last few decades.

The progression of climate change since 1950s is at an unprecedented level while already affecting almost every inhabited region across the world, with increased risks of agricultural and ecological drought expectations especially for Mediterranean, West Central Asia and East Central Asia as well as certain other regions (IPCC, 2021). Additionally, increase in the regional hot extremes are already at a hazardous level for the global inhabitation and with special emphasis for the agro-ecological environments. Therefore, it is globally acknowledged that sustainability and resilience of animal production systems and therefore global food security highly depend on the conservation and efficient utilization of animal genetic resources with high adaptive capacity (Boettcher et al., 2014; FAO, 2015; Hoffmann, 2010). Indigenous livestock breeds with different patterns of micro evolutionary histories of local adaptation have novel and mostly unselected genomic variation and various haplotype combinations for fitness related quantitative traits such as fertility and health as well as production traits (Hill, 2014; Kristensen et al., 2015). Therefore, those animals represent potential reservoirs that may contribute sustainable production for food security of those unpredictable times and extreme circumstances. Besides, certain populations of those

indigenous breeds may become core to the future breeding programmes that aim adaptation of livestock production to novel agro-ecological environments (Biscarini et al., 2015; Boettcher et al., 2014).

High-throughput genotyping technologies, genome-wide representative single nucleotide polymorphism (SNP) arrays and specially-tailored statistical & bioinformatics approaches have armed the conservation and animal geneticists with sophisticated tools to characterize genome of farm animals. Those tools have been used to reveal genomic variability and various population parameters, to estimate genetic parameters underlying quantitative traits without any pedigree, to implement genomic selection and to discover loci segregating behind those quantitative traits (QTL) at an ever higher resolution for those indigenous animals (Kizilaslan, et al., 2022a; Mastrangelo et al., 2016; Moradi et al., 2021; Nazari-Ghadikolaei et al., 2018; Visser et al., 2013; Yilmaz et al., 2022). All these studies and many others have potential to pave the way for those breeds for the management and sustainable use of breed's genetic resources with efficient conservation and breeding programmes .

As a native breed with a claimed ancient heritage, Akkaraman sheep is considered a culturally and socioeconomically important sheep breed of Türkiye (Yalcin, 1986; Behrem, 2021b). The breed possibly constitutes a large share of Türkiye's approximately 46 million size of sheep population, of which 42 million have been considered as indigenous animals (Behrem, 2021a; Ozmen et al., 2020; TUIK, 2022). Its hardy nature, ability to thrive on meagre diet and resilience under extreme conditions nominates the breed as a potential contributor to feeding the future against changing climate, recurrent pandemics and other unpredicted bottlenecks. However, despite its socioeconomic importance, studies concerning genetic background of economically important traits as well as the genetic characteristics of the breed are quite scarce in Akkaraman sheep (Bayram et al., 2019; Kirikci, 2022). In fact, no study concerning de-novo QTL discovery of the breed was observed while certain characterisation studies with low density microsatellite markers have been conducted (Ozmen et al., 2020). Importantly, in recent decades Akkaraman sheep have been imposed to certain population size fluctuations due to altered socioeconomic and agricultural circumstances leading to the favouring of transboundary sheep breeds. Therefore, it is of

crucial importance to implement genome-wide characterization and QTL discovery studies for the efficient management and sustainable use of breed's genetic resources.

On the other hand, growth traits and the traits of visual conformation appraisal has always been at the heart of ruminant selection since the domestication (Janssens & Vandepitte, 2004). Growth traits are central to the many sheep genetic improvement programmes due to their direct influence on the profitability, efficiency and sustainability of production systems as a result of their genetic and phenotypic correlations with carcass yield and quality, feed efficiency, reproduction and survivability related traits (Hu et al., 2022). As a method for conformation appraisal, linear type traits has recently been included in the selection programmes of many livestock species, frequently for dairy and beef cattle as well as dairy goats, in many countries, also due to their interplay with robustness, mobility, productivity, feed efficiency, longevity and survivability traits (Bouška et al., 2006; Brotherstone, 1994; Campos et al., 2015; Cromie et al., 2008; Janssens & Vandepitte, 2004; Mellado et al., 2008; Mucha et al., 2018). However, in sheep, studies concerning linear type traits are quite scarce (de la Fuente et al., 2011; Fernández et al., 1997; Gutiérrez-Gil et al., 2011; Pieramati et al., 2011). Akkaraman sheep, as a major domestic genetic resource significantly contributing to the livelihoods of people in Türkiye, critically lack studies concerning estimation of genetic parameters, QTL discovery and population parameters which significantly hampers faster genetic improvement of the breed while wisely managing the current genetic variation that should have been aimed for these traits to improve overall efficiency of sheep production systems and breeding programmes. Inclusion of genomic information into breeding programmes have been proven to remarkably increase genetic gain obtained within a short period of time while carefully managing the potential loss genetic variation due to selection. Methods such as marker-assisted selection, gene-assisted selection and genomic selection have long been proven to significantly contribute profitability, sustainability and efficiency of livestock production systems by remarkably increased genetic gain per generation while limiting loss of genetic variation due to the availability of the genomic information to monitor genetic variability (Dekkers, 2007; Doormaal, 2012; T. Meuwissen et al., 2013; Rupp et al., 2016; Schaeffer, 2006; Shumbusho et al., 2014).

To guarantee the existence and prosperity of the breed for our future generations by designing smart, efficient and sustainable breeding and conservation programmes, inclusion of genomic information to the breeding and conservation programmes that are monitoring Akkaraman sheep is quite fundamental. Therefore, our study aimed to uncover genetic and phenotypic parameters for growth and linear type traits, discover QTL underlying complex mechanisms determining the phenotypic distribution of those traits and uncover certain population genetic parameters of Akkaraman sheep in comparison with other transboundary and indigenous sheep breeds by using genome-wide dense SNP panels. Consequently, our results are expected to guide designing sustainable breeding programmes and conservation of animal genetic resources regarding Akkaraman sheep primarily, while also significantly contributing global sheep production systems.

CHAPTER 1

1 LITERATURE REVIEW

1.1 Genetics of Growth and Conformation Traits in Akkaraman Sheep

Domestic sheep (*Ovis aries*) are considered to be one of the most utilized farm animals since domestication. Robust and highly adaptive nature of the species allowed it to be extensively distributed across various ecosystems and diverse terrains including those arid areas, which also led to the breed divergence. Moreover, the large variety of products provided by sheep established the species as a core component of the worldwide agricultural economy. Native sheep breeds, have a broad spectrum of adaptive traits evolved for many generations with natural and artificial selection, which lay the foundations for sustainable production and global food security. Similarly, Akkaraman sheep, as an extensively dispersed fat-tailed breed of Türkiye, is known for its highly robust nature, survivability and reproductive performance under extreme conditions (Behrem, 2021b; Yalcin, 1986). It has a quite big share of Türkiye's sheep population as a dual-purpose breed, with noticeably low meat productivity (Behrem, 2021b). Economically, sheep maintain the fourth largest share of the worldwide meat production (FAO, 2022b). Therefore, correlated traits, such as growth and linear type traits, have great importance in terms of being used as indicators for breeding practices. In fact, both trait groups have considerably high influence on most of the economically important traits such as disease susceptibility, robustness and fitness, productivity, mobility, reproductive performance, and certain physiological and biochemical parameters (Brotherstone, 1994; de la Fuente et al., 2011). Consequently, improved

selection accuracy for growth and linear type traits would mean additional genetic improvement concerning overall efficiency and profitability of a sheep production system.

As a method for conformation appraisal, linear type traits have widely been utilized for dairy cattle and dairy sheep for over decades. However, studies conducted on meat sheep are very limited (Carta et al., 2009; Meyer et al., 1987). Correspondingly, various scales were suggested for scoring meat sheep for type traits, including those grading based on 1 to 5 as well as 1 to 9 scales (de la Fuente et al., 2011; Gutiérrez-Gil et al., 2011; Janssens & Vandepitte, 2004; Pieramati et al., 2011). However, studies concerning estimation of genetic parameters and investigating the genetic basis of growth and linear type traits in sheep are also quite scarce (de la Fuente et al., 2011; Gutiérrez-Gil et al., 2011; Zhang et al., 2013). Most of the economically important traits in livestock are quantitative in nature. In other words, these traits display a polygenic background where small effects of many loci aggregate together (i.e., Infinitesimal model) to contribute phenotypic observation, while environmental factors also play role on the final observation (Falconer & Mackay, 1996; Fisher, 1919). It is widely known that growth and linear type traits in sheep also display quantitative nature, which indicates that part of the phenotypic variation in the observations are established by the animals' genetic background. Indeed, moderate heritability estimates and a broad range of genetic correlations were observed for those complex traits in different sheep breeds, which demonstrate the existence of a potential for exploiting the genetic determination of growth and linear type traits to aim faster genetic progress regarding productivity, robustness and fitness (Behrem, 2021a; de la Fuente et al., 2011; Janssens & Vandepitte, 2004; Kumar et al., 2019; Safari et al., 2005). Obviously, to estimate genetic parameters accurately, an extensive pedigree is required. Most countries significantly suffer from the absence of pedigree recording schemes. Therefore, the use of genomic information to overcome this problem by obtaining a genomic relationship matrix replacing a pedigree-based average relationship matrix was proposed (Meuwissen et al., 2001; VanRaden, 2008). Furthermore, including genomic information has also been proven to substantially increase the accuracy of selection for genetic improvement programmes and therefore the obtained genetic gain within a short

period of time. Compared to the other existing methods including the infamous BLUP (i.e., Best Linear Unbiased Prediction) and progeny testing, use of genomic information has been proven to boost genetic gain by increased selection accuracy by increasing accuracies of estimated breeding values (EBVs) and selection intensity as well as a remarkably decreased generation interval (Dekkers, 2007; Doormaal, 2012; Meuwissen et al., 2013). However, to implement marker- or gene-assisted selection programmes for economically important traits, detection of the associated markers and quantitative trait loci (QTL) is a major prerequisite. Unfortunately, sheep have a relatively small number of discovered QTL in comparison with those found in cattle. Among those, studies aiming the discovery of QTL related to growth and linear type traits are rather sparse (Hu et al., 2022).

On the other hand, massive progress made by the sequencing technologies and immense efforts for the sequencing of sheep genome also led to a broad spectrum of commercial genome-wide single nucleotide polymorphism (SNP) arrays providing immediate access for the sheep research (Jiang et al., 2014; Kijas et al., 2009). With the availability of those high-density SNP arrays, a major opportunity for the implementation of genomic scans for QTL discovery, genomic characterization and genomic selection (i.e., genome-wide selection, genomic prediction) has emerged for certain traits and parameters sheep. Similarly, genome-wide association studies (GWAS) was suggested as a tool for revealing QTL underlying traits of economic interest in livestock production, which produced valuable information as a method for the detection of markers for marker-assisted selection, to localise genes with causal effects for gene-assisted selection and a target position for the possible genome-editing interventions for certain traits of interest (Goddard & Hayes, 2009). GWAS or in other words linkage disequilibrium mapping, has become one of the most popular approaches for QTL discovery and, on a broad perspective, for understanding the role of genetic variation on the distribution of phenotypic observations over the last decade (Loos, 2020; Uffelmann et al., 2021). The main reasons behind that trend are the supremacy and robustness of its underlying methods in terms of statistical power, precision of the mapping (i.e. confidence intervals lower than ~50 Kb compared to the typical ~10 Mb interval of linkage mapping) and the unnecessity of the family trios, where a population-wide study

is quite informative for GWAS (Cardon & Bell, 2001; Rannala & Slatkin, 2000; Risch & Merikangas, 1996). It has already proven to be a powerful tool to identify effect of the genetic variants associated with a wide range of traits in a high number of farm animals including cattle, goat and sheep (Gebreyesus et al., 2019; Mucha et al., 2018; Zhang et al., 2013). Up until recent years, only a few GWAS conducted on certain growth (Al-Mamun et al., 2015a; Gholizadeh et al., 2015; Lu et al., 2020; Yilmaz et al., 2022; Zhang et al., 2013) and linear type traits (Gutiérrez-Gil et al., 2011) of sheep have suggested putative QTL on OAR1-3, 6, 8, 9, 14-17, 19 and 22 for growth and OAR2, 5, 16, 23, 26 for different linear type traits. Further details regarding the discovered regions and putative QTL can be found on Animal QTL database (Hu et al., 2022).

1.2 Genomic Analysis of Akkaraman Sheep

Historically, captivating wild Asiatic mouflon (*Ovis orientalis*) and breeding those for domestic sheep (*Ovis aries*) has been proposed to be the dawn of domestication for grazing animals in Fertile Crescent around 11,000 years ago (Wei et al., 2015). Since then, sheep have spread throughout the world, even those arid and extreme environments, with over 1200 breeds descended from those early domesticated animals that are bred for variety of resources such as meat, milk and wool (Scherf & Pilling, 2015; Taberlet et al., 2011). World's current animal genetic resources, which also includes thousands of sheep breeds and varieties, are the assurance of feeding the future with sustainable livestock production, among the main pillars of food security and contribute to the livelihoods of almost 1.3 billion people (World Bank, 2021). Dramatically increasing population size, recent breaks in the food production and supply chains due to pandemics and those strongly expected by the latent effects of climate change once again emphasize the compelling importance of the local indigenous animal breeds as the source of uninterrupted food supply and assurance of the security for the future food demands due to their high biological diversity and adaptation capacity (Taberlet et al., 2008). Those indigenous breeds are crucially important due to their natural role as a reservoir of biological diversity, each of which may be an integral part of a solution for future production crises and natural bottlenecks imposed by threats such as climate change and recurrent pandemics. Likewise, many countries in the world

have a variety of indigenous animal genetic resources that have evolved in the relevant ecosystems for many hundreds of generations, which are also mainstay of / central to the sustainable production for those times of uncertainty due to their adaptive capacity (Boettcher et al., 2014). Among those, Türkiye has a wide range of diversification due to its natural position as the historical host to the early domestication. With the 41 indigenous and many outsourced international transboundary sheep breeds (e.g., Merino, Awassi, Lacaune etc.) and a population size of approximately 42 million of animals, it is among the top ten largest sheep producers in the world (FAO, 2022b, 2022a). Akkaraman (i.e., White Karaman), as a fat-tailed indigenous sheep breed of Türkiye, is widespread throughout Central Anatolia. The breed is characterized by its ability to endure extreme environmental conditions, including diseases and parasites, adaptation to nourish on meagre diet and capacity to survive and reproduce under adverse conditions (Ozmen et al., 2020). As a dual-purpose breed, it is mostly raised outdoor based on pasture. Akkaraman has a great share of Türkiye's sheep population with considerably low meat productivity of slaughtered lambs (Behrem, 2021b). Although widespread and highly adaptive, the low productivity of the breed compared to those imported such as Merino hinders potential contribution of the breed to the current economy and productivity of the sheep industry. However it is widely known that most of the economically important and adaptive traits are quantitative in nature, which means that the genetic architecture of the breed, especially its genetic variability is also a major contributor of its sustainable use and productivity (Falconer & Mackay, 1996). On the other hand, one of the initial steps of conserving animal genetic resources with the purpose of food security for future generations is to identify its current genetic architecture including genetic variability and population structure (Scherf & Pilling, 2015). Loss of genetic variability is long known to lead lack of fitness, poor fertility and disease resistance as well as productivity (Charlesworth & Charlesworth, 2010; Falconer & Mackay, 1996; Keller & Waller, 2002). In fact, identification of genetic architecture and especially genetic variability is also laying foundations for the strategic priorities for action (SPAs) by the Global Plan of Action for Animal Genetic Resources of United Nations (FAO, 2007; Scherf & Pilling, 2015). As indicated, absence of proper genomic characterization may lead to the loss of potentially valuable genetic characteristics and breed attributes due to unsupervised mating strategies, crossbreeding and extinction threats. Therefore, sustainable management of the animal genetic

resources also highly depends on characterizing the genomes of the target breeds and exploit these information for faster genetic improvement and conservation purposes. Genetic characterization of those indigenous breeds that are already adapted to a variety of environmental conditions using state of art genomic tools is quite essential to ensure breeding hardy and productive sheep populations for future uncertainties and handicaps (Boettcher et al., 2014;Yang et al., 2016).

On the other hand, identification of various genomic characteristics such as genetic diversity and linkage disequilibrium are very important in terms of designing genome-wide association studies (GWAS) and genomic prediction (Al-Mamun et al., 2015b). The amount of genetic diversity is crucial for the identification of the amount of phenotypic data required for genomic predictions with high accuracy. Likewise, it is excessively important for inferring the actual position of the quantitative trait loci (QTL) associated with a phenotype of interest as a result of GWAS. Although, estimation of genetic diversity is possible with pedigree or low density markers such as microsatellites, currently the utilization of SNP arrays provide more in-depth and accurate estimation than that of previously possible. Certain parameters such as F statistics (e.g., F_{IS} and F_{ST}), Linkage Disequilibrium (LD), runs of homozygosity (ROH), observed (H_o) and expected heterozygosity (H_E) are perceived to be among the indicators of genetic diversity. For instance, F statistics defined by Wright are commonly used to quantify overall population structure, the genetic diversity and amount of differentiation within and between populations (Crispim et al., 2013; Weir & Cockerham, 1984; Wright, 1965). Moreover, LD shows the degree of non-random correlations between loci of a genome. Therefore, information regarding the extent of LD is critical for QTL discovery with GWAS, marker-assisted selection and genomic selection programmes conducted for the genetic improvement of various livestock species. It is influenced by the evolutionary forces such as mutation, selection and migration as well as recombination events and the size of Ne (Charlesworth & Charlesworth, 2010; Nielsen & Slatkin, 2013). Evidently, fewer markers are required to track the genome with the purpose of genome-wide association studies and genome-assisted selection programmes assuming LD remained persistent over long chromosomal stretches than otherwise, when LD decays quickly in short distances.

Heterozygosity, as a measure of variation genetic variation within a population is among the most widely utilized diversity parameters (Toro & Caballero, 2005). Intuitively, high levels of heterozygosity imply for higher genetic variability and vice versa. On the contrary, long runs of homozygosity (ROH) may indicate higher levels of genomic inbreeding as they are consecutive homozygous chromosomal segments, while extremely shorter ROH also suggests loss of genetic diversity due to a recent bottleneck and founder effect. It is also useful for identifying regions of the genome that are under historical selection pressure (Peripolli et al., 2017). Due to their definitive characteristics, comparison of these parameters among populations of different sheep breeds is expected to guide researchers and breeders through more effective conservation and genetic improvement programmes.

Up until now, various molecular markers have been used to understand the genetic characteristics of certain species, breeds and economically important traits so far (Groeneveld et al., 2010; Lenstra et al., 2012; Scherf & Pilling, 2015). Among those, the most commonly used are mitochondrial DNA-based markers, microsatellites and single nucleotide polymorphisms (SNPs). Unlike the others, SNPs are distributed densely throughout the genome and currently the most powerful and robust tool for genome-wide monitoring and representation, fine-tuned diversity and variability estimations, unravelling genomic relationships and linkage disequilibrium properties as well as rendering genome-wide discovery of QTL associated to economically important traits in various farm animals (FAO., 2011; Mucha et al., 2018; Scherf & Pilling, 2015). Massive advances in the last decades led to the high-throughput genotyping technologies to kick-start the genomics era to surround genetic studies of farm animals (Kizilaslan et al., 2022a). Sequencing efforts of farm animals led to the manufacturing of various high density SNP arrays that can simultaneously genotype tens of thousands of loci with tremendously low costs, which revolutionised traditional approach to breeding, characterization and conservation of farm animals. Accordingly, various genome-wide scans were implemented with the purpose of QTL discovery (Al-Mamun et al., 2015a; Kizilaslan et al., 2022b; Yilmaz et al., 2022; Zhang et al., 2013) and genomic characterization in various sheep breeds (Ahbara et al., 2021; Al-Mamun et al., 2015b; Kijas et al., 2012; Rekik et al., 2022). Such characterization studies have

employed various analytical strategies that involve phylogenetic trees, comparison of expected (H_e) and observed (H_O) heterozygosities -per SNP and per individual-, various aspects of allele frequencies such as minor allele frequency, deviations from Hardy-Weinberg Equilibrium, genomic relationships and inbreeding, measures of genetic differentiation (e.g., F_{ST}) and population structure both within and between breeds, while pairwise estimations of linkage disequilibrium (LD) coefficient r^2, decay of LD with distance and runs of homozygosity (ROH) mostly undertaken within breeds (FAO, 2011; Rekik et al., 2022; Scherf and Pilling, 2015). One of the early characterization studies conducted via SNP arrays unravelled the genomic similarities and differences among 74 global sheep breeds (Kijas et al., 2012). Accordingly, high SNP diversity and effective population size were observed for most of the breeds, suggesting no constrained and small genetic base at the early phases of domestication. Additionally, relatively low divergence time between the breeds suggests a persistent genomic exchange while the current breeds were formed. Another study revealed genetic diversities and breed differences among six sheep breeds that are Belclare, Beltex, Charollais, Suffolk, Texel and Vendeen (Purfield et al., 2017). An indigenous fat-tailed sheep breed New Baisary was recently characterized and genomic differences, among certain local sheep breeds were established in Kazakhistan (Narzhan et al., 2022). On the other hand, a genome-wide survey with SNP panels were undertaken to establish differences between British and American Suffolk populations as well as Australian Poll Dorset and American Dorset populations (Kijas et al., 2009; Romanov et al., 2021). Previously, a microsatellite based characterization study for Akkaraman sheep was conducted by (Ozmen et al., 2020). However, no genome-wide characterization has been observed in the relevant literature. Detection of such characterization parameters are crucially important for efficient conservation and breeding programme designs. Thus, it becomes possible to account for the current state of the genomic structure and genetic variability and employ those to carefully dissect QTL for traits of interest as well as breed and conserve those animals that are either under risk of extinction or in need of a properly designed genome-wide selection (i.e., genomic selection) for higher genetic gain in livestock production (Kijas et al., 2009). Although there have been a few studies in various sheep breeds, SNP-based genome-wide studies concerning genomic architecture of indigenous sheep breeds are still quite limited.

1.3 Objectives of the Study

Overall, considering the lack of comprehensive studies in terms of the genetics of growth and conformation traits in sheep research and the crucial need for those parameters for a properly designed breeding programme that is aiming faster genetic improvement in sheep productivity, efficiency, robustness and fertility, one of the main objectives of our study is to unravel the genetic parameters of those traits including their genomic heritabilities as well as genetic and phenotypic correlations in Akkaraman lambs. Being among the very first studies around the world and indeed the first study aiming estimation of genetic parameters of growth and conformation traits in Akkaraman sheep, our results are expected to significantly contribute to the design of sheep breeding programmes in terms of growth and linear type traits, while directly applicable to the Akkaraman breeding programmes.

Furthermore, the obvious scarcity of the discovered QTL, little overlap among them and the complex nature of the genetic background of growth and linear type traits emphasize the need for a deep understanding of the genetic architecture of those traits as well as the need for more QTL discovery studies in sheep. Therefore, the second main objective of the present study is to implement Genome-wide Association Studies for growth and linear type traits to understand the genetic basis of those complex traits with potential implications on the genetic improvement of an indigenous sheep breed. Relevant findings are envisioned to strongly contribute to the genome-assisted breeding programmes of sheep production systems, especially those aiming higher genetic gains in Akkaraman sheep while also uncovering potential targets for potential future genome-editing, gene knockout and expression profiling studies.

Finally, it is extensively established that the identification of characterization parameters such as genetic diversity indices, population structure, linkage disequilibrium, runs of homozygosity as well as inbreeding are crucially important for efficient conservation and breeding programme designs. Knowledge on these parameters facilitate accounting for the current state of the genomic structure and genetic variability and employ those to carefully breed and conserve those animals that

display loss of genetic variability (i.e. inbreeding depression) as well as those either under risk of extinction or in need of a properly designed genome-wide selection (i.e., genomic selection) for higher genetic gain in livestock production (Kijas et al., 2009). Up until now, no genome-wide SNP based study concerning the comparative genomic characterization of Akkaraman sheep, alongside with the other indigenous sheep breeds of Türkiye, against the mainstream transboundary sheep breeds were implemented. Assessing genetic diversity and structure of animal breeds is among the key contributors of deciphering adaptation to environmental extremes and constructing efficient genetic improvement strategies (Boettcher et al., 2014; Groeneveld et al., 2010). Therefore, the final goal of this research is to define genomic characteristics of Akkaraman sheep by revealing the level of genetic differentiation, LD patterns and various levels and structure of genetic diversity in comparison with numerous transboundary, indigenous and wild sheep breeds using genome-wide single nucleotide polymorphisms (SNPs) with the aim of better understanding the nature of the breed's genome. Results of this study are strongly expected to contribute to the design and analytical process of further genome-wide association studies and genomic selection of the traits of economic importance in the production systems of Akkaraman sheep as well as the management and conservation of the world's animal genetic resources.

CHAPTER 2

2 MATERIAL AND METHODS

2.1 Ethics Statement

All the experimental protocols used on animals of the study were approved by the Local Ethics Committee of the Experimental Animals of the Ministry of Agriculture and Forestry in Türkiye, with a file number of 20.November.2020/183. Genomic data for the rest of the animals other than Akkaraman sheep were retrieved from the WIDDE database (Sempéré et al., 2015).

2.2 Animals and Phenotyping

The experimental population of Akkaraman sheep breed was composed of 480 randomly selected lambs (294 female and 186 male lambs) obtained from 3 different large-scale commercial farms located around the outskirts of Ankara/Türkiye. The climatic characteristics of the region is defined by harsh and cold winters, while summers are excessively hot and dry, leading to an insufficient quality grassland and therefore a meagre diet for those grazing animals. Animals were monitored under the

National Small Ruminant Breeding Programme. Therefore, a transgenerational selection based on phenotypic observations of growth rate is applied to the ancestors of the research population. Lambs were born between January and February 2021 and weaned between April and May 2021 with an average weaning age of 3 months. Right after weaning, 372 lambs were kept on with grazing on the pasture till 6 months of age while 114 lambs were subjected to a finishing period of 90 days with 750-1000 g of concentrate mixture (i.e., corn, barley and wheat) per day as well as hay.

Records on the observations of lambs regarding birth weight (**BW**), weaning weight (**WW**) and 180-days weight (**180DW**) were taken. Subsequently, pre-weaning average daily gain (**preADG**), post-weaning average daily gain (**postADG**) and 180-days average daily gain (**180ADG**) were also derived from those records. Here, WW is the adjusted 90-days weight of the lambs, while 180DW is the linearly interpolated 180 days weight. When the lambs were weighed for the 180DW, scoring for 11 traits including chest width (**CW**), rear legs width (**RLW**), body condition score (**BCS**), rear legs feet angle (**RLFA**), body length (**BL**), rump width (**RW**), tail size (**TS**), rear legs rear view (**RLRV**), gigot roundness rear view (**GRRV**), gigot roundness side view (**GRSV**) and rear legs side view (**RLSV**) were implemented following also the descriptions provided by (de la Fuente et al., 2011; Gutiérrez-Gil et al., 2011; Janssens & Vandepitte, 2004; Pieramati et al., 2011). Following the scoring, approximately 6 ml of blood samples were collected from the *V. jugularis* into an EDTA-coated vacutainer to prevent clothing. Samples were immediately transferred within a cold chain to the Genetics Laboratory of the International Center for Livestock Research and Training, Ankara/Türkiye, to implement DNA extraction and genotyping.

Environmental factors to be fitted as fixed effects to determine the final models for GWAS were sex (male and female), birth type (singlet and twin), herd (3 levels), and feed type (pasture and finishing). Preliminary statistics, data pruning and model fitting preceding genetic parameter estimations and GWAS were implemented on R statistical environment (R Core Team, 2021) . Continuous data were checked for normality via Shapiro-Wilk test where observations deviating more than 3 standard deviations ± mean

for each trait were excluded as outliers. To determine the fixed effects to be accounted for in the univariate analyses of each trait, Generalised Linear Models (GLM) were performed preceding genetic parameter estimations. Furthermore, heteroscedasticity of variances were tested with the Breusch-Pagan test (Breusch & Pagan, 1979). Descriptive statistics of the phenotypic observations were given at the **Table 1**.

After the genetic parameter estimations, GWAS and functional annotation analysis, genomic data were reduced to a set of randomly selected 120 animals and used in the rest of the analyses where comparative genomic characterization of the breed was aimed. The reduction of the sample set was done only to avoid extreme differences between the numbers of animals among the study populations, which might eventually lead to the biased estimations. Subsequently, genomic data from 5 mainstream transboundary breeds that are Australian Merino (AUM), Rambouillet (RMB), Suffolk (SUF), Texel (TEX), milk Lacaune (LMI); 4 fat-tailed indigenous sheep breeds that are Karakas (KRS), Cyprus Fat Tail (CFT), Moghani (MOG) and Tibetan (TIB); 2 thin-tailed sheep breeds that are Chios (CHI) and Sakiz (SKZ); and 2 ancestral breeds that are Asiatic Mouflon (OOR= *Ovis orientalis*) and European Mouflon (EMU= *Ovis aries musimon*) were retrieved from the WIDDE database for the comparative genomic characterisation of the breeds (Sempéré et al., 2015). Number of animals belonging to total of 14 breeds were presented on **Table 1**.

2.3 Genotyping and Quality Control

DNA were extracted from blood samples of Akkaraman lambs with the QIAcube HT, an automated high-throughput nucleic acid purification instrument, using a Blood/Tissue DNA purification kit and the manufacturer's protocol (QIAGEN, Hilden, Germany). Following the purification, DNA samples were exposed to a quality control (QC) step where $A_{260/280} > 1.8$; $A_{260/230} > 1.5$ and a sample amount of >20 ng/µL were sought as the minimum genotyping requirements by using a MultiSkan SkyHigh UV/VIS microplate spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA). After the DNA integrity being observed on gel electrophoresis, samples passing the

quality control stage were proceeded to the genotyping which was carried out using the Axiom™ Ovine SNP Genotyping Array (50 K) via GeneTitan™ Multi-Channel Instrument following the manufacturer's guide (i.e., Axiom™ 2.0 Assay 96-Array Format Manual Workflow, ThermoFisher Scientific, Waltham, MA, USA). Briefly, a seven stage procedure was performed for genotyping samples: 1) Amplification of DNA samples; 2) Fragmentation and precipitation; 3) Centrifugation, drying and resuspension; 4) Fragmentation Quality Control; 5) Denaturation and hybridization on GeneTitan Multi-Channel Instrument; 6) Staining and stabilization; 7) Scanning. First four consecutive steps were carried out manually where the last three steps were on an automated genotyping instrument. Following the scanning, '.CEL' files were generated for each sample where intensity calculations and relevant metrics for each feature of the probe array. Those '.CEL' files were finally used for genotype calling on the software package Axiom Analysis Suite, where genotypes for each SNPs of each animal were obtained as 'ATGC' format and allele dosage format (i.e., 0= major allele homozygote, 1= heterozygote and 2= minor allele homozygote).

Preceding further analyses, a quality control (QC) step was performed to ensure lower Type-I and Type-II error rates (Burton et al., 2007; McCarthy et al., 2008; Weale, 2010). Briefly, SNPs with minor allele frequency (MAF) below 0.05 are suggested to be removed to eliminate those rare variants which are remarkably prone to the classification errors during genotyping. Besides, potential power loss due to multiple testing correction is avoided by the removal of SNPs with less likelihood of association. Thresholds for SNP and individual call rates are generally imposed to overcome problems emerging from poor quality genotype clustering and DNA sample handling. Additionally, samples with to high identity by state (IBS) are also proposed to be disposed since they are likely to be duplicates. SNPs on gender chromosomes are also suggested to be eliminated preceding further analysis, since those generate difficulties with kinship estimations and principal components analysis (PCA). Finally, less stringently, thresholds for departures from Hardy-Weinberg Equilibrium (HWE) are also used as indicators for genotype calling issues. Correspondingly, 'GenABEL' R package (Aulchenko et al., 2007) was used to remove SNPs that have a call rate below 95% and those mapped to sex chromosomes, SNPs with minor allele frequency (MAF)

< 0.05 and do not conform with Hardy-Weinberg equilibrium under Bonferroni-corrected significance level (α) equal to 0.05. Furthermore, samples with an individual call rate below 90%, identity by state (IBS) higher than 95% and those with too high heterozygosity (False Discovery Rate < 1%) were also removed from the data, which ended up with 40,439 SNPs from 479 animals passing imposed QC criteria for follow-up statistical analyses.

For the genetic characterization analyses, 50K SNP genotypes for the rest of the breeds were recovered from the WIDDE database. Those retrieved genotypes from the database were originally obtained from the Illumina OvineSNP50v1 and AgResearch OvineHD SNP arrays and were deposited as a result of the studies conducted by (Ciani et al., 2015; Kijas et al., 2012; Rochus et al., 2018). Therefore, RS-ID's of the SNPs on different platforms were used to merge a common SNP set for the analysed breeds. Subsequently, a quality control procedure was applied where SNPs with call rate below 95%; SNPs on the X and Y chromosomes as well as mitochondrial DNA and SNPs with Minor Allele Frequency (MAF) less than 5 % were excluded from the dataset. The dataset were also checked for individual quality control parameters for which animals with a call rate below 90% and IBS > %95 were to be omitted from the analysis, however no animal was found to be violating these thresholds. Consequently, a total of 29,481 SNPs and 610 animals from 14 breeds were retained after QC to perform comparative genomic characterisation analyses. Quality control step of the merged data set was implemented on 'GenABEL' R package (Aulchenko et al., 2007).

2.4 Statistical & Bioinformatics Analysis

2.4.1 Estimation of Genetic Parameters

Genomic heritability estimates of growth and linear type traits as well as pairwise genetic correlations were obtained as a result of univariate and bivariate analyses of linear animal mixed models implemented on 'sommer' R package, general model outlined below with the variance-covariance structure of the estimations (Covarrubias-Pazaran, 2016):

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}$$

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{Z_i G}\sigma^2_{u_i}\boldsymbol{Z'_i} + \boldsymbol{I}\sigma^2_{e_i} & \cdots & \boldsymbol{Z_i G}\sigma_{u_{i,j}}\boldsymbol{Z'_j} + \boldsymbol{I}\sigma_{e_{i,j}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{Z_i G}\sigma_{u_{i,j}}\boldsymbol{Z'_j} + \boldsymbol{I}\sigma_{e_{i,j}} & \cdots & \boldsymbol{Z_j G}\sigma^2_{u_j}\boldsymbol{Z'_j} + \boldsymbol{I}\sigma^2_{e_j} \end{bmatrix}$$

Here, **y** refer to the vector containing phenotypic observations. **X** and **Z** are the design matrices that are respectively mapping fixed environmental effects and random breeding values to the observations. $\boldsymbol{\beta}$ here represents the vector of fixed effects (i.e. fixed environmental effects for genetic parameter estimates and later additional SNP effects for GWA analysis) included in the model, **u** is the vector of random effects for genomic breeding values (gEBVs) and **e** is the vector of random residual effects. Values for u and e are assumed to be drawn from a multivariate normal distribution where u ~ MVN (0, $\mathbf{G}\sigma^2_u$) and e ~ MVN (0, $\mathbf{I}\sigma^2_e$), respectively. Additionally, σ^2_u and σ^2_e are respectively the additive genetic and random residual variances for the trait of interest, where 'i' and 'j' stands for different traits used for bivariate analyses and **I** is the identity matrix. Finally, **G** here is the positive semi-definitive genomic relationship matrix (GRM) obtained by the method proposed by (VanRaden, 2008):

$$\boldsymbol{G} = \frac{\boldsymbol{ZZ'}}{2\sum \boldsymbol{p_i}(\boldsymbol{1} - \boldsymbol{p_i})}$$

When divided by $2\sum p_i(1-p_i)$, the G matrix show properties resembling to that of numerator relationship matrix obtained from a pedigree.

Estimated heritabilities in our study are narrow sense heritability ($h^2 = \frac{\sigma^2_u}{\sigma^2_p}$) where σ^2_u is the additive genetic variance and σ^2_p is the total phenotypic variance that are obtained from the univariate analysis of the interested trait with the model presented above. For decomposing phenotypic (co)variance into its components Newton-Raphson

optimization method with direct inversion (DI) based restricted maximum likelihood (REML) approach was exploited to solve the mixed model above, also using the GRM constructed from the information regarding IBS provided by the SNP markers (Lee & Van der Werf, 2016; G. T. Wilson, 1989). Heritability estimates, genetic and phenotypic correlations are given in **Table 2**. Fixed effects were omitted during the bivariate analyses (i.e., estimation of genetic correlations) to avoid possible convergence problems. Standard errors of the genetic correlations were estimated by using a second order Taylor series expansion on the delta method as explained by (Lynch & Walsh, 1998).

2.4.2 Genome-wide Association Analysis (GWAS)

Single-SNP GWAS for each trait was carried out using the univariate linear mixed model given above and the simple procedure proposed by Chen and Abecasis (2007) was employed on 'GenABEL' R package. For this purpose, the GRM was also included in the models to account for any possible covariance between the observations of animals on the same trait that is attributable to stratification and cryptic relatedness, which otherwise would lead to increased Type I error rate. SNPs with the missing genotypes were imputed with Expectation-Maximization (EM) algorithm where a mean genotype score for the unknown marker was estimated considering the surrounding markers (Chen & Abecasis, 2007). Briefly, the two-step approach suggested for solving the equation above involves a step of accounting for all fixed effects other than SNP effects (i.e., the recorded environmental factors) and estimating the (co)variance components, where the residuals of this primary analysis is later used to estimate the additive genetic effects of the SNPs which are fitted as fixed factors and once at a time.

Quantile- Quantile (Q- Q) plots were used to inspect the inflation of the test statistics due to any possible systematic bias or error such as cryptic relatedness and unaccounted population stratification which might lead higher rates of false positive associations. Quantiles of the expected (null hypothesis of "no association") and observed test statistics are plotted against each other where any mass deviation can be interpreted as a

sign of population structure, genotyping error and unidentified relatedness. Furthermore, genomic control method was used to adjust and remove any possible inflation in the test statistics by setting λ to 1 (Devlin & Roeder, 1999). SNP effects were visualized as -log10 (p-value) of the tested SNPs that are represented on Manhattan plots of each trait, considering their position at the relevant chromosome. To overcome the risk of leveraged false positives imposed by multiple testing, two significance thresholds were applied to the SNPs to identify genome- and chromosome-wide significants based on p-values. Bonferroni correction was utilized to assign both genome- and chromosome-wide significance thresholds.Thus, genome-wide significance threshold was obtained as 1.236×10^{-6} and chromosome-wide as 3.214×10^{-5} which respectively leading to 5.90 and 4.49 on the –log10 (p-value) scale that is on Manhattan plots.

2.4.3 Functional Annotation of Candidate Genes

Genomic positions and the nearby genes related to the associated SNPs were retrieved from the Oar_v4.0 genome assembly on NCBI Genome Data Viewer (Rangwala et al., 2021). Here, the genes that contain the significant SNPs were directly suggested as candidates. However, when the SNP was not within a certain gene, the area on the genome covering ± 250Kbp from the associated SNP was scanned for the nearest candidate genes. Functional enrichment were carried out for the detected genes by using DAVID Bioinformatics Resources-2021 to retrieve biological information and the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways that were involved by those (Huang et al., 2009a, 2009b). Where the sheep genome suffered from the lack of annotation, orthology among species was exploited to further annotate the relevant genes by outsourcing information from other species such as cattle, mice and human. The biological processes of the identified genes were given with their Gene Ontology terms and can be further detailed on QuickGO by EMBL's European Bioinformatics Institute (Binns et al., 2009). Finally, the animal QTL Database is also scanned to identify whether detected SNPs in this study are previously associated with any of the growth and type traits (Hu et al., 2022).

2.4.4 Genetic Diversity Indices

Assessment of the genetic diversity parameters within each population, which are the observed heterozygosity (H_O), unbiased expected heterozygosity (H_E) inbreeding coefficient (F_{IS}) were calculated by the R package "dartR" (Gruber et al., 2018). Inbreeding coefficient was obtained as 1 - (mean (Ho) / mean (He)). Deviations from Hardy-Weinberg equilibrium was tested based on the exact calculations described in (Wigginton et al., 2005) by computing p-values as half the probability of the current sample plus the probabilities of more extreme samples and plotted in the **Figure 4** with the 'dartR' package.

2.4.5 Genetic Distance & Population Structure

To observe the genetic relationship, population structure within and between breeds and the distance among those, Principal Components Analysis (PCA) were performed with the 'adegenet' R package (Jombart, 2008). First two principal components were plotted and presented in Figure 5. On the other hand, unbiased F_{ST} (i.e., fixation index) estimates among the populations were also estimated with the 'dartR' package to obtain pairwise genetic differentiation among the sheep breeds in the dataset as described by (Weir & Cockerham, 1984). Finally, the phylogenetic tree was constructed with the matrix of pairwise F_{ST} values, where the neighbor-joining (NJ) algorithm was implemented on the same R package to observe the genetic similarities and clusters among the sheep breeds.

2.4.6 Linkage Disequilibrium Analysis

Patterns of pairwise linkage disequilibrium (LD) were investigated and reported as the squared Pearson's correlation coefficient (r^2) between the alleles of SNPs to ease the comparison of our results with those existing reports. Pairwise average linkage disequilibrium between adjacent SNPs and the decay of linkage disequilibrium (LD decay) were estimated using the formula suggested by (Hill & Robertson, 1968) that was performed with the 'r2fast' function of GenABEL R package as described below:

$$r^2 = \frac{[f(AB) - f(A)f(B)]^2}{f(A)f(a)f(B)f(b)}$$

Here (A),$f(a)$,$f(B)$,$f(b)$ are the allele frequencies at two different loci where $f(\mathrm{A}B)$ is the haplotype made of the major alleles at the loci of interest. Pairwise LD was measured only on the syntenic loci that are the loci only on the same chromosome.

2.4.7 Genomic Inbreeding and Runs of Homozygosity Analyses

Genomic inbreeding for each animals were evaluated as the genome-wide autozygosity obtained from the genome-wide SNP data using F_{ROH} statistics generated from the identified Runs of Homozygosity (ROH), which was introduced by (McQuillan et al., 2008). Correspondingly, F_{ROH} was obtained as outlined below:

$$F_{ROH} = \frac{\sum L_{ROH}}{L_{AUTO}}$$

Where $\sum L_{ROH}$ is the total lengths of the identified ROH and L_{AUTO} is the length of the autosomal genome represented by the SNPs in the study. For the identification of ROH and F_{ROH} for each individuals of the each breed, sliding window approach provided by the 'detectRUNS' R package was used, where certain parameters were used as the predefined criteria for defining each individual ROH (Biscarini et al., 2018). Accordingly, no LD pruning were implemented, but the minimum length for defining a ROH was set to 1 Mb for excluding those shorter ROH that possibly are the direct consequence of LD. The genome of each animal were scanned with a sliding window of 20 SNPs, where no more than one missing SNP is allowed per window and minimum SNP density was set to 1 SNP per 1000 Kb with a maximum gap between two consecutive SNPs to be 1000 Kb distance. Finally, SNPs were included in the scanning window when p-value < 0.05.

CHAPTER 3

3 RESULTS

3.1 Phenotypic statistics

In total, 17 different traits were evaluated for the estimation of genetic parameters and GWA analysis. Following the linear model fitting, outliers were removed accordingly. Briefly, sex was not significant for any of the traits. For the pre-weaning growth traits, only birth type was used for the adjustment while for the traits recorded after weaning, finishing was also a factor of adjustment preceding GWAS. After the genotyping quality control, the remained number of observations for GWAS of each trait and their relevant descriptive statistics are given in **Table 1**.

Table 1. Descriptive statistics of growth and linear type traits.

Trait	N	Min	Mean	SD	Max
BW (kg)	451	2.96	4.59	0.71	6.54

WW (kg)	449	14.10	25.71	4.29	36.61
180DW (kg)	452	19.33	42.87	7.99	68.05
preADG	450	72	234	46	365
postADG	456	11	168	71	396
180ADG	451	82	212	70	352
BCS	473	2.00	3.66	0.98	5.00
TS	469	2.00	4.78	1.46	8.00
RLRV	470	2	3.68	0.81	5
GRRV	471	4	5.35	0.89	7
RW	473	3	5.38	1.34	8
RLW	473	2	3.39	0.92	5
RLFA	473	2	3.39	0.92	5
GRSV	472	3	5.39	0.87	7
RLSV	469	3	4.41	0.66	5
BL	473	4	4.91	0.82	7
CW	470	3	5.34	1.02	8

Notes: **N**= Number of observations; **Min**= Minimum; **SD**= Standard deviation; **Max**= Maximum; **BW**= Birth weight; **WW**= Weaning weight; **180DW**= 180 days adjusted weight; **preADG**= pre-weaning average daily gain; **postADG**=post-weaning average daily gain; **180ADG**= 180 days average daily gain; **BCS**= Body condition score; **TS**= Tail size; **RLRV**=Rear legs (rear view); **GRRV**= Gigot roundness (rear view); **RW**= Rump width; **RLW**= Rear legs width; **RLFA**= Rear legs foot angle; **GRSV**= Gigot roundness (side view); **RLSV**= Rear legs (side view); **BL**= Body length; **CW**= Chest width.

3.2 Genetic parameters

Estimated genomic heritabilities as well as pairwise phenotypic and genetic correlations between the traits were presented in **Table 2**. Essentially, moderate heritability estimates were observed for growth traits ranging from 0.29 for preADG to 0.52 for 180DW. On the other hand, for the linear type traits, heritability estimates were distributed with a wider range between 0.07 for BL and 0.52 for RLW. In general, positive and high genetic correlations were observed among growth and type traits, which also supports the idea for linear type traits to be used as indicator traits for improving growth. A large proportion of the genetic correlations among the relevant traits were close to unity, indicating an almost perfect linear relationship among the traits. Interestingly, traits RLW and RLFA had pairwise moderate negative between and occasionally almost no correlations with some of the traits analysed. Genetic correlation among the BL and GRRV did not converge to a meaningful number possibly due to the low genetic variance estimates of the traits, therefore excluded from the results. Finally, phenotypic correlations were observed to be within the low to moderate range among the traits, with an exception of RLW and RLFA, which showed no correlation at all with most of the traits in the dataset.

Table 2. Heritability estimates, genetic and phenotypic correlations of growth and linear type traits in Akkaraman sheep.

Trait	BW	WW	180DW	preADG	postADG	sixADG	BCS	TS	RLRV	GRRV	RW	RLW	RLFA	GRSV	RLSV	BL	CW
BW	**.39(.12)**	.32(.04)	.34(.04)	.16(.04)	.30(.04)	.26(.04)	.21(.04)	.28(.04)	.10(.04)	.19(.04)	.13(.04)	-.11(.04)	.03(.04)	.17(.04)	.09(.04)	.22(.04)	.14(.04)
WW	.88(.13)	**.38(.13)**	.85(.04)	.98(.03)	.34(.03)	.84(.00)	.30(.04)	.60(.04)	.18(.04)	.50(.04)	.42(.04)	-.03(.04)	-.10(.04)	.50(.04)	.08(.04)	.53(.04)	.43(.04)
180DW	.96(.13)	.92(.04)	**.52(.13)**	.83(.02)	.75(.03)	.99(.00)	.38(.04)	.67(.03)	.24(.04)	.56(.04)	.48(.04)	.05(.04)	.00(.04)	.56(.04)	.08(.04)	.66(.04)	.53(.04)
preADG	.60(.19)	.99(.01)	.95(.01)	**.29(.12)**	.31(.04)	.84(.02)	.29(.04)	.58(.04)	.17(.04)	.49(.04)	.42(.04)	-.01(.04)	-.10(.04)	.51(.04)	.06(.04)	.52(.04)	.43(.04)
postADG	.71(.14)	.69(.17)	.96(.03)	.77(.14)	**.51(.13)**	.75(.04)	.29(.04)	.47(.04)	.21(.04)	.38(.04)	.33(.04)	.08(.04)	.08(.04)	.39(.04)	.08(.04)	.50(.04)	.39(.04)
sixADG	.83(.16)	.96(.01)	1.00(.00)	.96(.01)	.97(.02)	**.51(.13)**	.38(.04)	.66(.03)	.24(.04)	.56(.04)	.49(.04)	.06(.04)	.00(.04)	.57(.04)	.08(.04)	.66(.04)	.53(.04)
BCS	.62(.18)	.72(.13)	.94(.12)	.74(.14)	.91(.21)	.97(.12)	**.32(.12)**	.45(.04)	.06(.04)	.47(.04)	.45(.04)	.01(.04)	.01(.04)	.35(.04)	.00(.04)	.34(.04)	.46(.04)
TS	.75(.15)	.91(.05)	.91(.04)	.90(.05)	.81(.09)	.91(.04)	.80(.08)	**.50(.13)**	.15(.04)	.71(.04)	.67(.04)	-.02(.04)	-.05(.04)	.60(.04)	.06(.04)	.63(.04)	.63(.04)
RLRV	.30(.17)	.49(.15)	.64(.16)	.51(.17)	.58(.17)	.65(.16)	.21(.19)	.41(.16)	**.26(.11)**	.09(.04)	.03(.04)	.17(0.04)	.13(.04)	.04(.04)	.12(.04)	.15(.04)	.08(.04)
GRRV	.59(.18)	.89(.07)	.91(.06)	.90(.07)	.86(.12)	.92(.06)	.96(.07)	.91(.03)	.33(.19)	**.30(.12)**	.72(.04)	.03(.04)	-.02(.04)	.65(.03)	.04(.04)	.62(.04)	.69(.03)
RW	.41(.17)	.84(.09)	.82(.07)	.88(.09)	.71(.12)	.83(.07)	.85(.08)	.87(.04)	.12(.21)	.83(.14)	**.25(.11)**	.10(.04)	-.03(.04)	.67(.04)	.02(.03)	.60(.03)	.71(.03)
RLW	-.25(.14)	-.02(.12)	.24(.09)	-.03(.14)	.31(.09)	.26(.10)	-.66(.25)	.04(.12)	.44(.12)	.16(.14)	.32(.15)	**.52(.13)**	.03(.04)	.04(.04)	.01(.04)	.19(.04)	.04(.04)
RLFA	.57(.51)	-.30(.20)	.14(.29)	-.33(.21)	.39(.32)	.14(.29)	.14(.36)	-.19(.26)	.65(.33)	-.04(.35)	-.08(.38)	.11(.28)	**.39(.12)**	-.09(.04)	.27(.04)	.00(.04)	-.02(.04)

GRSV	.46(.17)	.88(.07)	.87(.06)	.91(.07)	.81(.11)	.88(.06)	.83(.12)	.88(.05)	.14(.20)	.97(.03)	.96(.03)	.11(.15)	**-**.31(.28)	**.20(.10)**	.10(.04)	.67(.04)	.68(.03)
RLSV	.36(.28)	.31(.28)	.32(.28)	.32(.27)	.36(.32)	.29(.28)	**-**.09(.25)	.22(.24)	.78(.38)	.03(.27)	-.09(.31)	.00(.22)	.96(.14)	.31(.26)	**.16(.09)**	.07(.04)	.07(.04)
BL	.90(.40)	.97(.05)	.92(.04)	.98(.06)	**.**88(.07)	.93(.04)	.94(.13)	.92(.04)	.68(.30)	NA	.99(.04)	.59(.16)	.21(.62)	1.00(.25)	.25(.42)	**.07(.07)**	.62(.04)
CW	.43(.17)	.86(.09)	.81(.06)	.90(.09)	.71(.09)	.82(.07)	.79(.07)	.83(.05)	.31(.20)	.96(.03)	.92(.11)	.15(.15)	.06(.41)	.97(.03)	.23(.26)	1.00(.21)	**.30(.11)**

Notes: Diagonal values represent the genomic heritability of the overlapping trait where phenotypic correlations were given above diagonal and genetic correlations below diagonal. NA= could not be reliably estimated.

3.3 Genome-wide association studies

GWAS for 17 traits were carried out via univariate mixed model analysis of each trait by fitting 40,439 SNPs one at a time recursively, as well as a GRM. Significant SNPs with their descriptive characteristics and associated genes were presented on **Table 3** and **4** for growth and linear type traits, respectively. Q-Q plots for the test statistics (see **Figure 3**) show reasonable patterns of associations and no proper sign of inflation. All traits were conservatively forced to have a lambda (λ) of almost 1 for the test statistics by correcting the p-values with genomic control. The adjusted p-values of the SNPs were visualised with Manhattan plots and given by **Figure 1** and **2** as converted into –log10 (p-value) for each trait respectively. Consequently, 20 different SNPs were found to be genome-wide or chromosome-wide associated with at least one trait at various p-levels. Curiously, 3 SNPs showed multiple associations among the analysed traits.

Table 3. Significant SNPs from GWAS of growth traits.

Trait	SNP Name	Chr	Position (bp)	p-value	Significance	MAF	Candidate gene	Distance
BW	rs414781462	12	51,905,842	1.793×10^{-5}	CW	0.262	*PRDM2*	25 Kb upstream
WW	**rs414279727**	7	42,283,714	1.377×10^{-5}	CW	0.479	*PTGDR*	~50 Kb downstream
preADG	rs423706103	19	39,758,712	1.270×10^{-5}	CW	0.276	*PTPRG*	Intron variant
postADG	rs407771300	4	85,034,136	7.974×10^{-6}	CW	0.486	*KCND2*	Intron variant
postADG	rs404771550	14	46,046,471	9.876×10^{-6}	CW	0.253	*ZNF260*	~100 bp upstream
postADG	rs400100688	17	448,355	1.677×10^{-5}	CW	0.121	*CPE*	Intron variant
postADG	rs429448354	6	30,967,377	1.701×10^{-5}	CW	0.424	*GRID2*	Intron variant
sixADG	**rs414279727**	7	42,283,714	3.031×10^{-5}	CW	0.479	*PTGDR*	~50 Kb downstream

Notes: Chr= Chromosome; MAF= Minor Allele Frequency. Bold SNP names indicate those that are found common among the traits. Here the abbreviations BW, WW, preADG, postADG and sixADG respectively stand for birth weight, weaning weight and pre-weaning average daily gain, post-weaning average daily gain and six months average daily gain. Finally, regarding the significance column, GW stands for genome-wise and CW for chromosome-wise

significance of the SNPs.

Table 4. Significant SNPs from GWAS of linear type traits.

Trait	SNP Name	Chr	Position (bp)	p-value	Significance	MAF	Candidate Gene	Distance
CW	**rs412278842**	6	97,145,365	9.972×10^{-7}	**GW**	0.126	*SCD5*	~100 Kb upstream
RLRV	rs424107094	9	32,023,081	3.510×10^{-7}	**GW**	0.220	*SPIDR*	Intron variant
RLRV	rs399086810	23	3,643,516	2.215×10^{-5}	CW	0.454	*ZNF407*	Intron variant
BCS	rs405451961	1	103,920,255	9.628×10^{-6}	CW	0.399	*HCN3*	Exon variant
TS	rs417737929	2	240,680,051	4.327×10^{-6}	CW	0.050	*TMEM50A*	~100 Kb downstream
TS	rs404544718	2	25,145,453	1.135×10^{-5}	CW	0.347	*LOC105606997*	~15 Kb upstream
GRRV	**rs412278842**	6	97,145,365	7.323×10^{-6}	CW	0.126	*SCD5*	~100 Kb upstream
RW	Affx-282917850	13	58,972,752	1.897×10^{-5}	CW	0.326	*FKBP1A*	~15 Kb
RW	**rs427724890**	2	56,856,646	2.762×10^{-5}	CW	0.120	*TLE4*	~1-2 Kb upstream
RLW	rs417934245	8	81,117,831	1.668×10^{-5}	CW	0.356	*LOC105609002*	~90 Kb downstream
RLW	rs410532183	3	132,823,805	2.189×10^{-5}	CW	0.169	*SP1*	~ 1Kb upstream
RLFA	rs398187690	15	73,532,306	1.912×10^{-5}	CW	0.156	*LOC105602367*	~90 Kb
GRSV	**rs427724890**	2	56,856,646	1.769×10^{-5}	CW	0.120	*TLE4*	~1-2 Kb upstream
RLSV	rs411373597	2	17,728,536	3.243×10^{-6}	CW	0.258	*SLC44A1*	~250 bp upstream
RLSV	rs409929874	2	241,787,646	3.320×10^{-5}	CW	0.450	MYOM3	~30 Kb downstream

Notes: Chr= Chromosome; MAF= Minor Allele Frequency. Bold SNP names indicate those that are found common among the traits. CW, RLRV, BCS, TS, GRRV, RW, RLW, RLFA, GRSV and RLSV are respectively

abbreviations for chest width, rear legs rear view, body condition score, tail size, gigot roundness rear view, rump width, rear legs width, rear legs foot angle, gigot roundness side view and rear legs side view. Finally, regarding the significance column, GW stands for genome-wise and CW for chromosome-wise significance of the SNPs.

Considering in detail, analysed growth traits resulted in 7 distinct SNPs that were observed to be exceeding chromosome-wide significance threshold, which are on OAR4, 6, 7, 12, 14, 17 and 19. Among those, the SNP named 'rs414279727' on OAR7 was common between the traits WW and sixADG as outlined in Table 3. On the other hand, 2 SNPs on OAR6 and OAR9 exceeded the genome-wide significance threshold while 11 SNPs on OAR1, 2, 3, 8, 13, 15 and 23 were reached chromosome-wide significance for the GWAS of linear type traits (Table 4). Specifically, GWAS for chest width (CW) identified the SNP 'rs412278842' ($p = 9.972 \times 10^{-7}$) on OAR6 exceeding genome-wide significance threshold, while the same SNP was also found to be chromosome-wide significant for the GRRV with a p-value of 7.323×10^{-6}. The other genome-wide significant SNP, namely the 'rs424107094' on OAR9, was identified to be associated with RLRV with a p-value of 3.510×10^{-7}. Moreover, one chromosome-wide significant SNP, 'rs427724890' on OAR2, that was found to be associated with RW ($p = 2.762 \times 10^{-5}$) was also identified as chromosome-wide significant for GRSV (1.769×10^{-5}). Finally, it is also worth noting that the 5 of the associated SNPs with 4 different type traits were discovered on OAR2. More detailed information on the associated SNPs can be found on **Table 3** and **4**.

3.4 Annotation of Candidate Genes

Collectively, 17 individual genes together with 3 uncharacterized loci were proposed as candidates that are involving in the genetic control of growth and linear type traits. Oar_v4.0 genome assembly was utilised to annotate genomic regions of the associated SNPs and to detect the nearby genes on NCBI Genome Data Viewer and UCSC Genome Browser. Among those identified genes, *PTGDR, KCND2, PRDM2, ZNF260, CPE, PTPRG* and *GRID2* were associated with growth traits where *PTGDR* was suggested for both WW and sixADG. On the other hand, *ZNF407, SCD5, SPIDR, HCN3, FKBP1A, TLE4, MYOM3, SP1, SLC44A1* and *TMEM50A* genes on various

chromosomes as well as *LOC105606997, LOC105609002* and *LOC105602367* were identified as candidates for linear type traits. The gene *SCD5* was associated with chest width and GRRV and *TLE4* with RW and GRSV simultaneously. Detailed knowledge on the associated SNPs, detected genes and their relative positions were outlined in **Table 3** and **4**.

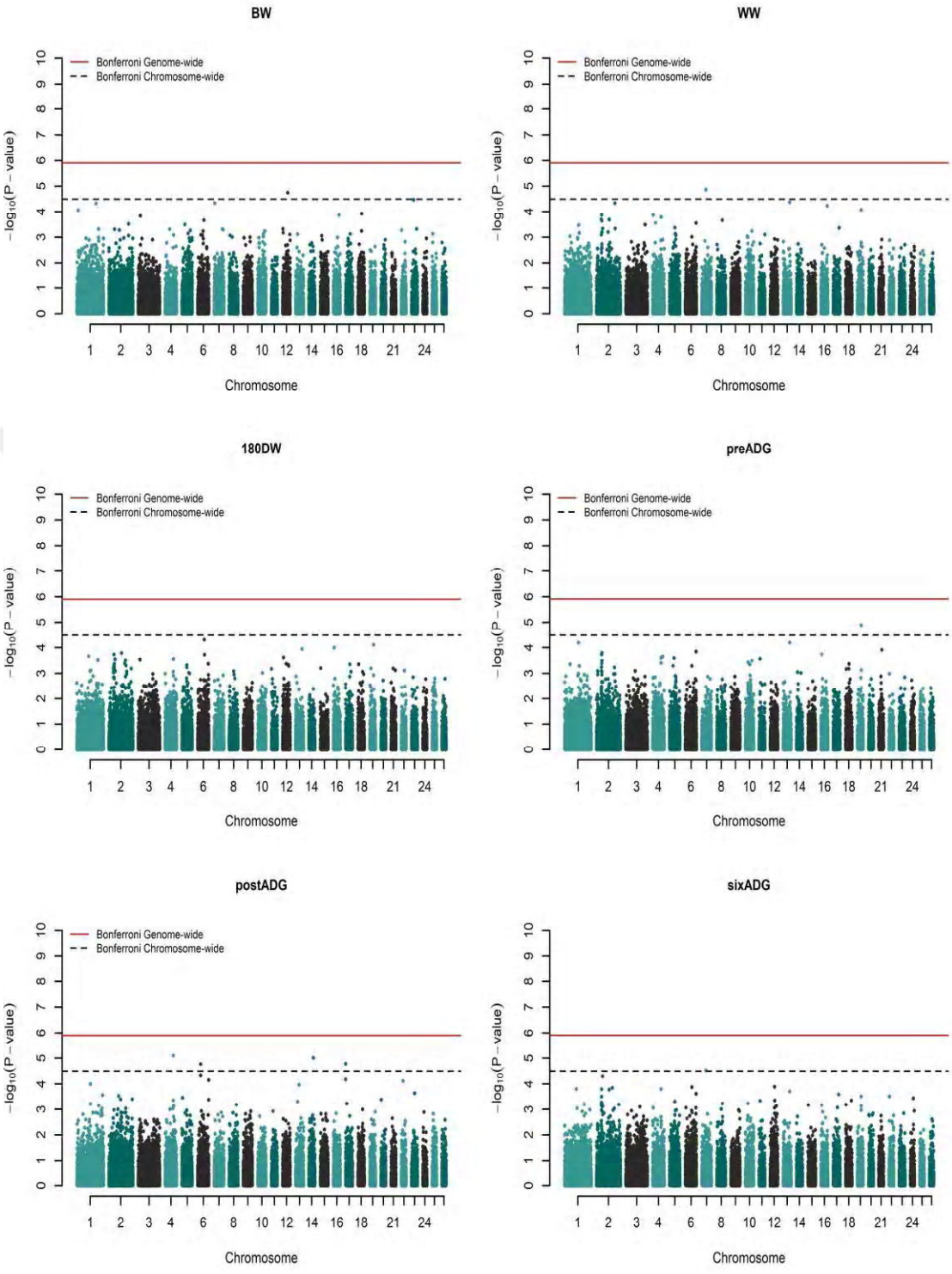


Figure 1. Manhattan plots of growth traits in Akkaraman sheep (−log10 (SNP p-values)). Horizontal red solid line represent genome-wide (5.90) and black dashed lines represent chromosome-wide significance thresholds (4.49).

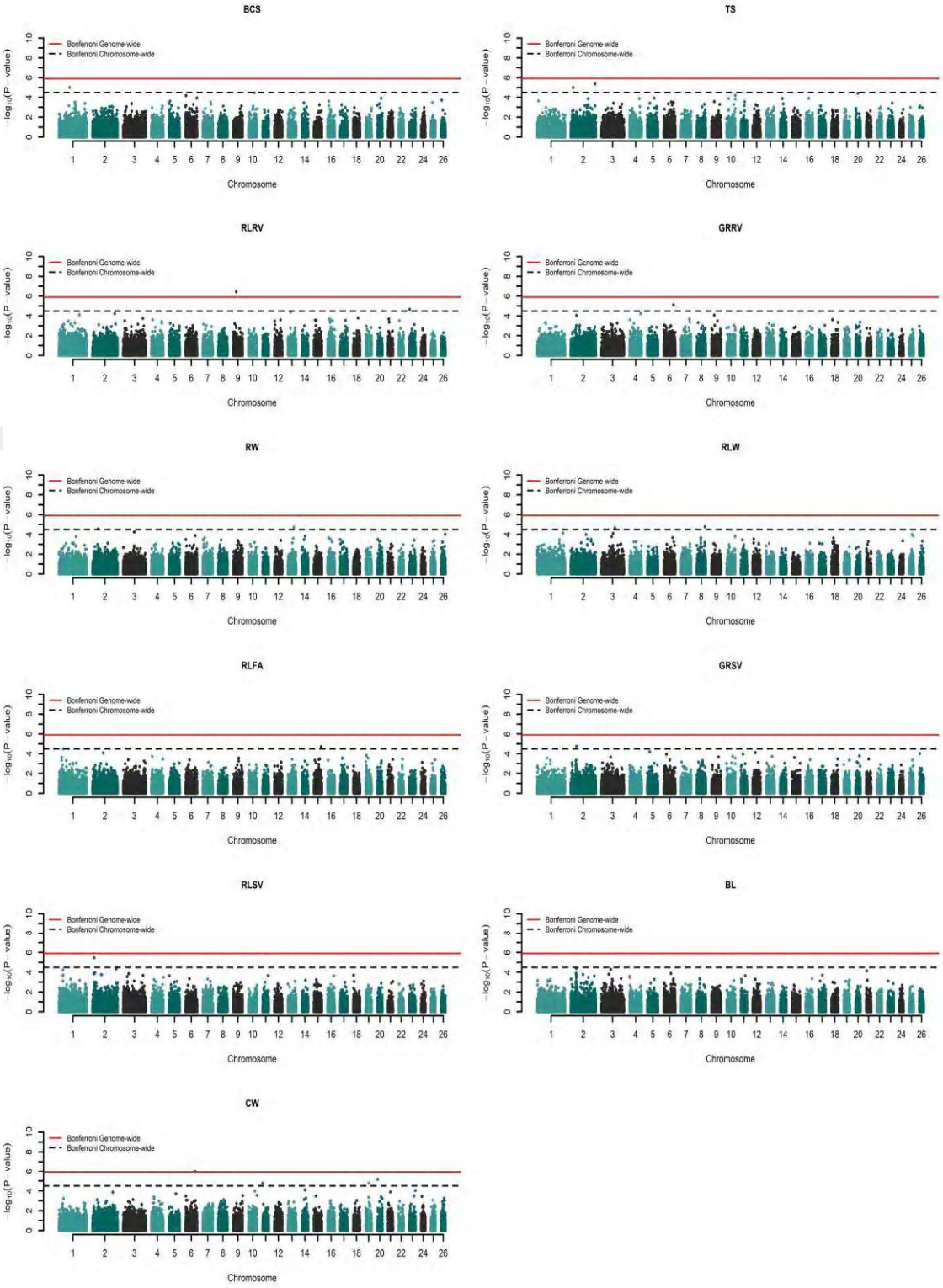


Figure 2. Manhattan plots of linear type traits in Akkaraman sheep (−log10 (SNP p-values)). Horizontal red solid line represent genome-wide (5.90) and black dashed lines represent chromosome-wide significance thresholds (4.49).

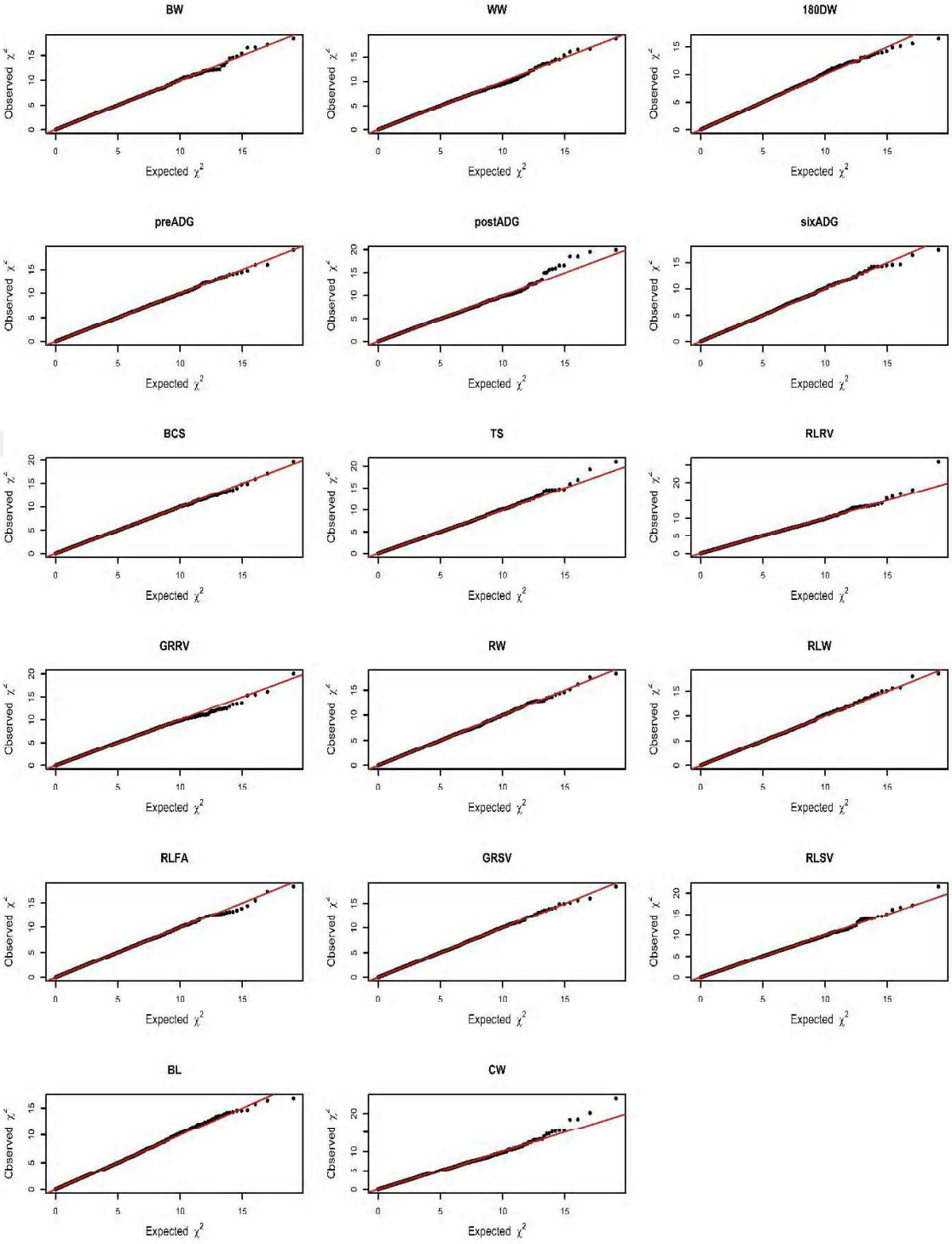


Figure 3. Quantile-Quantile (Q-Q) plot of the GWA analyses.

3.5 Genetic diversity indices

In this study, three common parameters were used to evaluate the levels of genetic diversity within breeds. Observed heterozygosity (H_O) and unbiased expected heterozygosity (H_E), that are averaged across all loci, as well as inbreeding coefficients averaged across individuals of each populations (F_{IS}) were estimated and presented in **Table 5** with the number of animals per breed observed in our study. Subsequently, most of the breeds were observed to have genetic diversities between 0.322 (TIB) and 0.362 (MOG) in terms of observed heterozygosity (H_O) with the exception of Europian Mouflon (EMU) and Asiatic Mouflon (OOR) wild breeds showing the lowest heterozygosities of 0.200 and 0.223, respectively. Again, almost all breeds outperformed those unbiased expected heterozygosities (H_E), which are between 0.183 (EMU) and 0.313 (AUM), estimated from genome-wide SNP markers and averaged across all loci. AKK was revealed to have the fourth largest observed heterozygosity (0.357) and expected heterozygosity (0.305) among the analysed 14 breeds in terms of the genetic variability. Mainly negative values were obtained for F_{IS} except for OOR which was observed to be 0.026. Correspondingly, regarding the absolute values of the estimations, SKZ (-0.201) displayed the largest average inbreeding coefficient within the breed, which is followed by CFT (-0.184), KRS (-0.179), MOG (-0.170) and AKK (-0.168), while the smallest absolute values for estimates were obtained for OOR (0.026) and EMU (0.091) breeds.

Table 5. Genetic diversity parameters for different sheep breeds.

Breed	**N**	**H_O**	**H_E**	**F_{IS}**
Akkaraman (AKK)	120	0.357	0.305	-0.168
Australian Merino (AUM)	50	0.361	0.313	-0.153
Cyprus Fat Tail (CFT)	30	0.324	0.274	-0.184
Chios (CHI)	23	0.322	0.279	-0.151
Europian Mouflon (EMU)	26	0.200	0.183	-0.091

Karakas (KRS)	18	0.352	0.298	-0.179
Lacaune Milk (LMI)	103	0.359	0.308	-0.165
Moghani (MOG)	34	0.362	0.309	-0.170
Asiatic Mouflon (OOR)	2	0.223	0.229	0.026
Ramboulliet (RMB)	102	0.344	0.299	-0.149
Sakiz(SKZ)	22	0.330	0.274	-0.201
Suffolk (SUF)	19	0.334	0.286	-0.165
Texel (TEX)	24	0.336	0.289	-0.162
Tibetan (TIB)	37	0.322	0.283	-0.135
Overall	610	0.324	0.280	-0.168

N= Number of observations; H_O= Observed heterozygosity averaged across loci; H_E= Unbiased expected heterozygosity averaged across loci corrected for sample size; F_{IS}= Inbreeding coefficient averaged across individuals of the population.

Finally, deviations from Hardy-Weinberg Equilibrium (HWE) was evaluated and visualised in **Figure 4** given below.

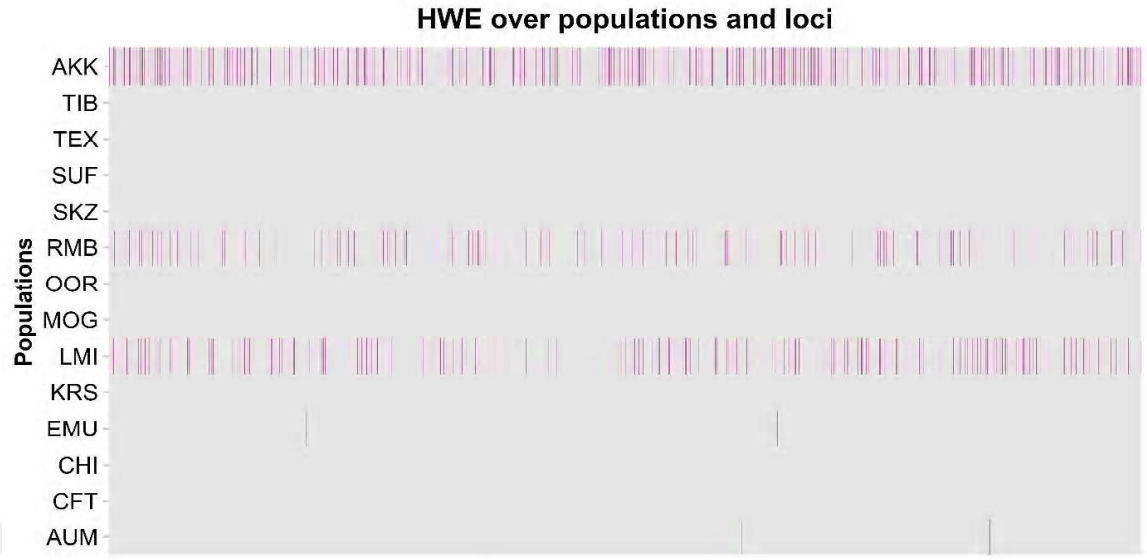


Figure 4. HWE deviation plot. Pink lines represent loci violating Hardy-Weinberg Equilibrium within each breed.

3.6 Genetic distance & structure among the breeds

Genomic relationships between the breeds were examined with the PCA plot which is visualised in **Figure 5.** The analysis resulted in the two major components, namely PC1 and PC2, explaining 14 % and 2 % of the variance among the breeds, respectively. Most of the breeds in our study demonstrated clear separations into visible sub-clusters with notable exceptions of European Mouflon (EMU) and Asiatic Mouflon (OOR), which are the two wild breeds that are speculated to have ancestral relationships with most of the common sheep genetic resources. The most remarkable detail provided by the PCA plot is the alignment of certain breeds into those distinct sub-clusters composed on the display. Interestingly, fat-tailed animals representing Akkaraman (AKK), Karakas (KRS), Moghani (MOG), Cyprus Fat Tail (CFT) and Tibetan (TIB) sheep appear to form a sub-group on the PCA plot, which is quite close to that of generated by Sakiz (SKZ) and Chios (CHI) sheep. Both groups are relatively distant with the group created by Australian Merino (AUM), dairy Lacaune (LMI), Texel (TEX) and Suffolk (SUF) as well as Rambouillet (RMB). European Mouflon (EMU) turned out to be the most genetically differentiated breed among the analysed breeds with the closest relationship

to those individuals of Asiatic Mouflon (OOR) as the second most distant breed to the others in our study.

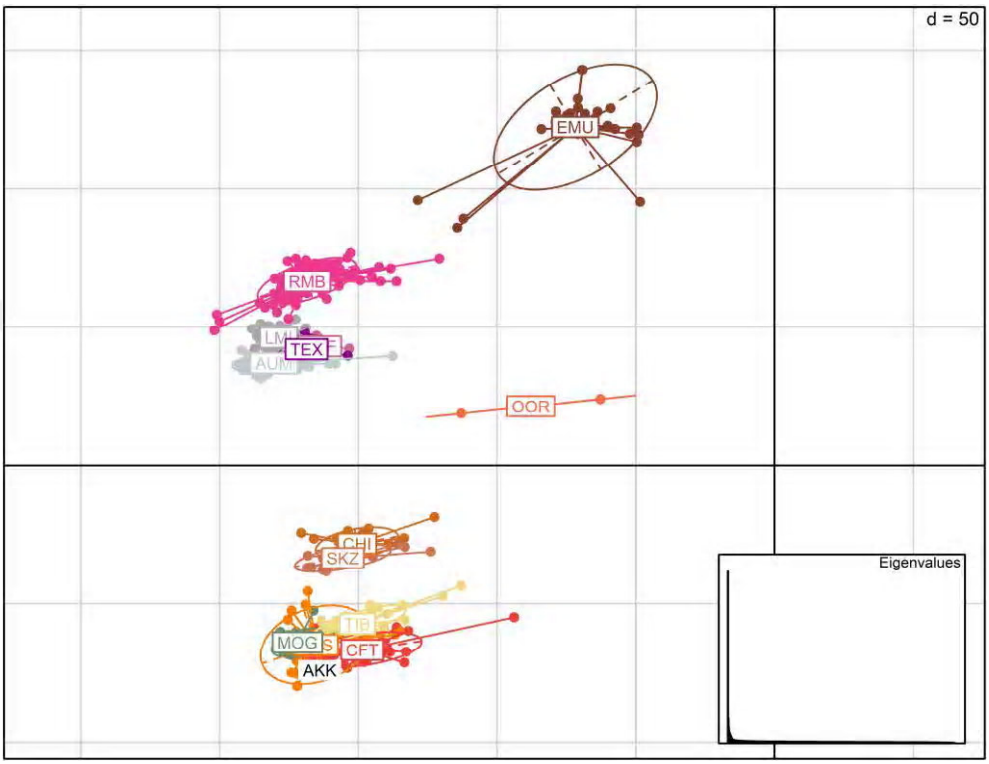


Figure 5. Principal Components Analysis (PCA) of the breeds.

Pairwise F_{ST} values among the breeds were estimated to quantify the genetic divergence among the breeds and provided in **Table 6**. Correspondingly, AKK and MOG were observed to have the lowest pairwise F_{ST} with 0.01, suggesting the least genetic divergence is among these breeds while EMU and OOR appear to have the greatest divergence with an F_{ST} value equal to 0.19. Subsequent breeds that are suggested to be the closest to AKK by pairwise F_{ST} estimations are KRS with 0.02, TIB and AUM with 0.03, CFT and CHI with 0.04, respectively. As also suggested by the PCA plot, EMU appear to be the most divergent breed compared to those in our study. On the other hand, AKK has been observed to have 0.06 and 0.13 with respectively OOR and EMU, which may emphasize the closer relationships with those breeds derived from OOR and

indicate the potential for the breed of origin preceding domestication to be OOR for AKK. The results were partially consistent with the PCA plot.

Table 6. Pairwise FST values between the sheep breeds.

	AKK	AUM	CFT	CHI	EMU	KRS	LMI	MOG	OOR	RMB	SKZ	SUF	TEX	TIB
AKK	NA													
AUM	.03	NA												
CFT	.04	.06	NA											
CHI	.04	.05	.07	NA										
EMU	.13	.12	.18	.17	NA									
KRS	.02	.04	.06	.06	.17	NA								
LMI	.04	.02	.06	.05	.12	.05	NA							
MOG	.01	.03	.05	.04	.14	.02	.04	NA						
OOR	.06	.05	.10	.08	.19	.07	.06	.06	NA					
RMB	.04	.02	.07	.06	.12	.05	.03	.04	.06	NA				
SKZ	.05	.06	.09	.06	.18	.07	.06	.06	.11	.07	NA			
SUF	.06	.05	.10	.08	.16	.07	.05	.06	.09	.06	.10	NA		
TEX	.06	.06	.10	.08	.16	.08	.06	.07	.09	.06	.10	.07	NA	
TIB	.03	.05	.07	.06	.16	.05	.05	.03	.07	.06	.08	.08	.08	NA

AKK=Akkaraman; AUM= Australian Merino; CFT= Cyprus Fat-Tailed; CHI= Chios; EMU= European Mouflon; KRS= Karakas; LMI= dairy Lacaune; MOG= Moghani; OOR= Asiatic Mouflon; RMB= Rambouliet; SKZ= Sakiz; SUF= Suffolk; TEX= Texel; TIB= Tibetan.

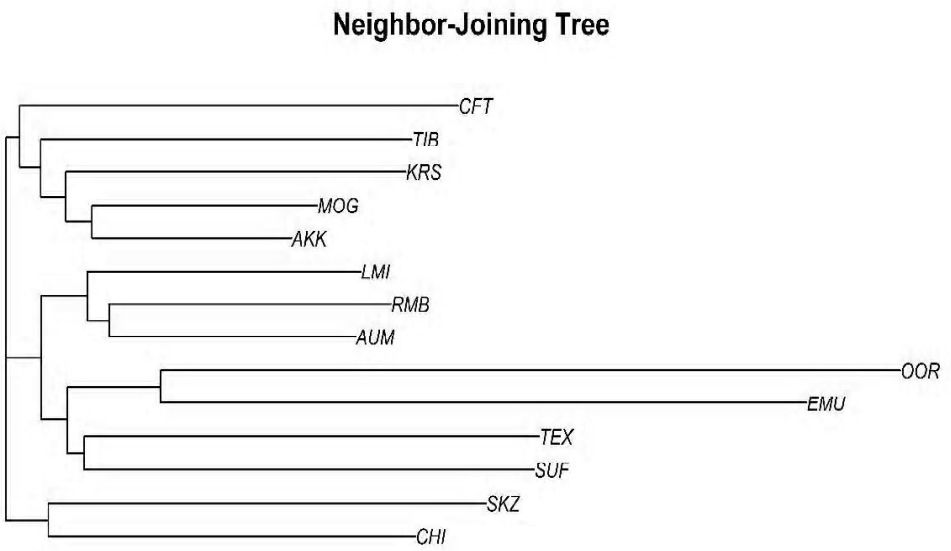


Figure 6. Phylogenetic Neighbor-Joining (NJ) tree of the breeds.

Finally, a phylogenetic tree was constructed using the pairwise F_{ST} estimations among the breeds and the Neighbor-Joining algorithm performed on the 'dartR' package (**Figure 6**). Accordingly, studied populations of the breeds exhibited a strong clustering into three major branch, where the branch is composed of CFT, TIB, KRS, MOG and AKK breeds; the second main lineage is formed by LMI, RMB, AUM, OOR, EMU, TEX and SUF breeds; and the last group by SKZ and CHI breeds. In the first branch, MOG and AKK formed a noticeable subgroup while in the main branch three strong subgroups were observed as LMI, RMB and AUM versus OOR and EMU versus TEX and SUF breeds. Furthest breeds in the tree were OOR and EMU, which is reasonable considering their ancestral role to the all those others in the phylogenetic tree. Finally, it is important to note that the results between PCA graph and the NJ tree were partially consistent, without any conflicts regarding the misclassification of the animals during assignment into breeds.

3.7 Linkage Disequilibrium (LD) and LD decay

Linkage disequilibrium estimations were performed as the r-squared (r^2) with the common 29,481 SNPs among the 14 analysed breeds, which are resulted from combining the datasets from different genotyping platforms and implementing the QC criteria, and presented in **Table 7**. Considering the overall length of the sheep genome as ~2.61 Gb, that leaves an average of 88 Kb gap between two consecutive SNPs. Accordingly, AKK was identified to have an average r-squared of 0.15 which was quite similar to those obtained for MOG, TIB, RMB, LMI and AUM. The lowest r-squared was obtained for OOR with 0.11 whereas LD for TEX, SUF and SKZ were demonstrated to be the highest among those analysed breeds.

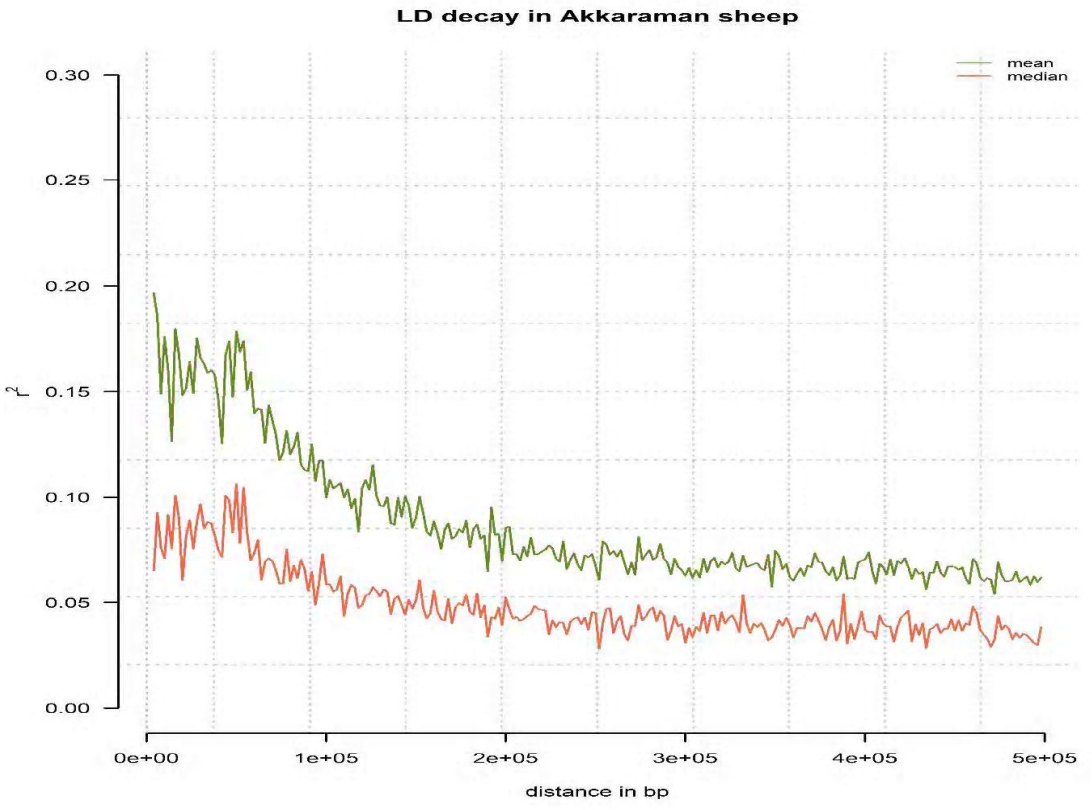


Figure 7. Trends in linkage disequilibrium (LD) decay in 500 Kb distance.

On the other hand, decay of LD with the increased physical marker distance on the genome was revealed for AKK sheep (**Figure 7**). Short distance LD for such as 100 Kb appear to exist in AKK sheep. However, it is clearly observable that LD in AKK quickly decays with physical distance, showing a pattern of continuous decline until just over 0.05 within a 500 Kb range.

3.8 Detection of Runs of homozygosity (ROH)

Various parameters regarding the genomic characteristics of the runs of homozygosity such as the total number of ROH per breed (N_{ROH}), mean number of ROH per individual per breed (MN_{ROH}) and average length of ROH (AL_{ROH}) per breed in Kb as well as the ROH-based genomic inbreeding coefficient (F_{ROH}) with its standard deviation were identified for each breed and also demonstrated on **Table 7**. The highest number of ROH per breed was found in RMB (9,520 ROHs), while the lowest was in OOR (385 ROHs) which was followed by KRS (762 ROHs). On the other hand, the mean number of ROH per individual was found to be the least for AKK (41.10 ROHs), which was interestingly followed by MOG (41.47) and KRS (42.33), whereas the highest number of ROH was found for EMU (202.65 ROHs) and OOR (192.50 ROHs) wild breeds. As well as having the highest number of per animal, EMU has the largest average length of ROH with 5,222 Kb per run, among all studied breeds, while the shortest average ROH length was detected for MOG (2,630 Kb) with AKK having the second smallest ROH length, which is 3,160 Kb per run. The magnitude of ROH-based genomic inbreeding (F_{ROH}) differed significantly among the animals of the 14 analysed breeds in our study. OOR and EMU wild breeds were identified to bear the highest genomic inbreeding evaluated by ROH with the values of 0.30 and 0.23, respectively. In contrast, AKK (0.002) appear to have the lowest average genomic inbreeding, followed by MOG (0.04) among all 14 breeds. The results were partially consistent with the observed heterozygosities (H_O) given above in **Table 5**.

Table 7. Descriptive statistics of ROH and LD in different sheep breeds.

Breed	N_{ROH}	MN_{ROH}	AL_{ROH} (Kb)	F_{ROH}	SD_{FROH}	LD (r^2)
AKK	4,932	41.10	3,160	0.002	0.001	0.15
AUM	3,245	63.62	3,572	0.09	0.04	0.16
CFT	2,260	75.33	4,863	0.15	0.06	0.21
CHI	2,761	120.04	3,780	0.18	0.04	0.21
EMU	5,269	202.65	5,222	0.23	0.08	0.19
KRS	762	42.33	4,154	0.07	0.06	0.19
LMI	6,011	58.36	3,574	0.08	0.01	0.16
MOG	1,410	41.47	2,630	0.04	0.03	0.15
OOR	385	192.5	3,812	0.30	0.20	0.11
RMB	9,520	93.33	3,588	0.13	0.05	0.16
SKZ	2044	92.90	4,063	0.15	0.05	0.22
SUF	2195	115.52	3,494	0.16	0.01	0.22
TEX	2871	119.62	3,575	0.17	0.02	0.22
TIB	3126	84.48	3,556	0.12	0.05	0.15

N_{ROH}=Total number of ROH per breed; MN_{ROH}= Mean number of ROH per individual per breed; AL_{ROH}= Average length of ROH (Kb) in a breed; F_{ROH}= Genomic inbreeding coefficient based on ROH. SD_{FROH}= Standard deviation of F_{ROH}.

CHAPTER 4

4 DISCUSSION-CONCLUSIONS AND FUTURE PERSPECTIVES

4.1 Discussion

Growth and linear type traits are central to the sheep selection programmes due to their high influence on production efficiency and overall profitability. However, indigenous sheep breeds such as Akkaraman suffer from the absence of a comprehensive pedigree, which is the basis for a conventional selection programme. Usage of genomic information in breeding decisions alleviates burdens brought by this problem via introducing genomic relationships, as well as those additional genetic gain provided (Dekkers, 2007; Doormaal, 2012; Meuwissen et al., 2001). The insufficient annotation throughout the sheep genome hampered the genome-based selection methods as well as understanding the complex nature of the traits in the species up until recently. With the advent of genomic technologies and the availability of highly annotated sheep reference genomes, it is now possible to use genome-wide distributed markers to understand the genetic structure of the economically important traits in sheep such as growth and linear type traits. With this purpose, in this study we first estimated heritabilities of 17 growth and linear type traits based on a genomic relationship matrix and obtained their pairwise genetic and phenotypic correlations to understand the nature of the phenotypic variation and the interplay between 17 growth and linear type traits. Subsequently, we performed a series of GWAS for those traits using randomly selected lambs from different herds and revealed 2 genome-wide and 18 chromosome-wide associated SNPs with proposed candidate genes. Finally, certain genomic characterization parameters such as those reflecting genetic diversity, genome structure, population structure and genetic

differentiation of the breed were estimated and displayed for a better understanding of the genetic architecture of the breed which will directly influence management of the breed's genetic resources and design of an efficient and sustainable genomic breeding programme aiming faster genetic improvement.

4.1.1 Genetic Parameters

In livestock production, knowledge on the genetic and phenotypic parameters such as heritability estimates of the quantitative traits that are at the core of the breeding objectives as well as genetic and phenotypic correlations among those are fundamental to implementing a robust and efficient genetic evaluation and designing breeding programmes with comprehensive selection indices. The efficiency of a multi-trait breeding programme is suggested to be directly affected by the accuracy and precision of the genetic parameter estimates exploited for the breeding objective development and construction of the selection indices (Safari et al., 2006). Heritability, as part of the selection response estimations, indicates the extent of genetic contribution to the variation at a particular trait of interest while genetic correlations would explain the potential effects and related consequences of selection on a trait on other traits of interest (Falconer & Mackay, 1996; Simm, 1998). Additionally, studies suggest that selection response in a multi-trait breeding programme is completely sensitive to the magnitude of genetic correlations among the indexed traits (Mortimer et al., 2017; Safari et al., 2006)

Heritability estimates of growth traits were ranged from 0.29 (preADG) to 0.52 (180DW) while for the linear type traits, heritabilities were between 0.07 (BL) and 0.52 (RLW). Both trait groups showed similar trends compared to those estimated in Suffolk by Janssens and Vandepitte (2004) except for body length, which turned out to have almost zero heritability in our study. Estimates for rump width, rear legs rear view and feet angle were only slightly higher than those obtained by de la Fuente et al. (2011) from Spanish Churra breed. Furthermore, obtained estimates for growth traits were significantly higher than those reviewed by (Safari et al., 2005) for various breeds and

by Behrem (2021a) for Central Anatolian Merino. Since it is widely known, heritability estimates are crucial elements of an efficient selection index construction process and prediction of selection response, which emphasizes importance of their accurate estimations for a breeding programme. The fact that moderate and consistent heritability estimates (see Table 2) were found for growth and most of the linear type traits implies a strong potential for faster genetic improvement with a properly designed selection programme.

On the other hand, mostly positive and high genetic and phenotypic correlations were found by our study, unlike those obtained by Janssens and Vandepitte (2004) from Bleu de Maine, Texel and Suffolk breeds. Conversely, relatively similar ranges were found for growth traits in our study and those proposed by (Safari et al., 2005). Although, Behrem (2021a) found negative moderate genetic correlations and again negative low phenotypic correlations between BW/WW and BW/preADG, results of our study suggests an opposite trend for the traits of interest in Akkaraman sheep, with significantly higher positive correlations. Since genetic correlations between traits may potentially arise from pleiotropic effects of certain genes, linkage and an interplay between background genetic mechanisms it is important to reveal and account for those correlations while designing a breeding programme. Otherwise, targeted selection responses may not be achieved, while simultaneously improving certain traits may lead to deterioration of others (Falconer & Mackay, 1996). The overall high and positive genetic/phenotypic correlations between growth and linear type traits in our study may lead to suggest that certain traits may be used as indicators of the others for early and easy breeding decisions, while those negative correlations should be taken into account for faster genetic improvement of growth.

4.1.2 Genome-wide Association Studies

For the GWA studies, on the other hand, GenABEL R package and an animal mixed model based approach exploiting genomic relationship matrix that accounts for the covariance due to cryptic relatedness and population stratification within the data set

was used, which may lead to inflation of the test statistics and eventually spurious associations if not accounted for (Astle & Balding, 2009; Sul et al., 2018). Overall, the Q-Q plots of the study showed no patterns of inflation that would lead deceptive conclusions (**see Figure 3**). Since the p-values of the SNPs were further corrected with genomic control, no inflation was expected nonetheless (Devlin & Roeder, 1999). Significantly, a deflated pattern displayed especially by the Q-Q plot of body length was interpreted as being due to the highly polygenic nature of the trait (Yang et al., 2011). Finally, annotations of the candidate regions were based on sheep *Oar_v*4.0 genome sequence.

One of the most important results of our study is the common SNPs discovered among the growth and linear type traits, which indicates a possible pleiotropy for the molecular mechanisms of the traits. High genetic correlations observed between WW-sixADG (0.96 ± 0.01), RW-GRSV (0.96 + 0.03) and GRRV-CW (0.96+0.03) also support the idea that common significant SNPs can be explained by a possible genetic interplay between the traits. In total, 20 SNPs were significantly associated with the focused traits at various levels, where three SNPs - rs427724890, rs412278842 and rs414279727**-** were found common between the referred traits with high genetic correlations. Further details regarding associated SNPs, their p-values, relevant candidate genes and their physical distance is provided in **Table 1** and **2**, respectively.

GWAS for growth traits have led to the detection of 7 candidate genes on OAR4, 6, 7, 12, 14, 17 and 19, where four of the suggested SNPs were found within the introns of the genes suggested. Additionally, among the suggested candidates, four of the genes were proposed for postADG and one gene, namely *PTGDR* was found to be common between WW and sixADG. As a G-protein coupled membrane receptor on OAR7, *prostaglandin D2 receptor* (*PTGDR*) has a molecular function of transmembrane signalling receptor activity (GO: 0004871) as well as eicosanoid receptor activity (GO: 0004953) in particular. Interestingly, *PTGDR* was also recently observed by a selection signature study conducted on thin- and fat-tailed sheep by (Moradi et al., 2012).

Additionally, the gene itself was also previously associated with body depth and hot carcass weight in different cattle breeds (Ashwell et al., 2005; Casas et al., 2000). On the other hand, the SNP named 'rs423706103' on OAR19, as the intronic variant of the gene *protein tyrosine phosphatase receptor type G* (*PTPRG*) with phosphoric ester hydrolase activity (GO:0042578), was associated with preADG in our study. Recently, *PTPRG* was suggested as a functional candidate for body size in sheep (Kominakis et al., 2017). Another SNP on OAR4, 'rs407771300', as an intronic variant of *potassium voltage-gated channel subfamily D member 2(KCND2)* was found to be associated with postADG. *KCND2* is suggested to be responsible for establishment of localization (GO: 0051234), protein complex assembly (GO: 0065003) and ion transport (GO: 0006811) as a transmembrane protein. Moreover, analysis for KEGG suggests that it also mediates serotonergic synapses. Evidently, another SNP within the gene *KCND2* was formerly associated with withers height in sheep (Tao et al., 2020). Finally, zinc finger protein 260 (*ZNF260*) on OAR14 also worth to mention in more detail. In our study, 'rs404771550' which is located nearly 100 bp upstream of *ZNF260* was specifically associated with postADG. Furthermore, 'rs399086810' on OAR23 was found to be associated with RLRV, as an intron variant of another zinc-finger protein (i.e., *ZNF407*). Zinc-finger proteins (ZNFs) are long known to be involved in the growth and development of various species (Cassandri et al., 2017). Additionally, various members of ZNFs –*ZNF395* and *ZNF641*- were already associated with growth traits in sheep (Erdenee et al., 2020; Yilmaz et al., 2022). Therefore, as validated by our study also, various ZNFs appear to have already proved their worth for studies aiming faster genetic improvement regarding growth in sheep.

The two genome-wide associated SNPs in our study, namely 'rs412278842' on OAR6 and 'rs424107094' on OAR9, were discovered as a result of GWAS for CW and RLRV respectively. The former SNP was also found to be associated with GRRV on chromosome-wide level. It is found to be located near 100 Kb upstream of the *stearoyl-CoA desaturase 5 (SCD5)* gene which has a critical role on fatty acid metabolic process (GO: 0006631) and lipid biosynthetic process (GO: 0008610). KEGG analysis also validated its role in fatty acid metabolism and biosynthesis of unsaturated fatty acids as well as AMPK signalling pathway. A recent study implemented whole-genome tests of

selective sweeps for various economically important traits in sheep and suggested *SCD5* as a strong candidate for body size(Xu et al., 2021). Given that, SCD5 is previously also proposed to involve in the development of body size in grass-fed Brangus steers due to its regulatory role in fatty acid biosynthesis (Baeza et al., 2013), it is also suggested as a strong candidate proposed from our study. The other genome-wide significant SNP, which is also the top associated SNP in our study, was found to be associated with RLRV and lying within the gene *scaffold protein-involved in DNA repair* (*SPIDR*) of OAR9 as an intron variant. As a nucleoplasm protein, *SPIDR* has key functions such the regulation of DNA recombination (GO: 0000018) and repair (GO: 0006281), positive regulation of macromolecule metabolic process (GO: 0010604), response to hydroxyurea (GO: 0072710) and regulation of response to stress (GO: 0080134). Apart from those genome-wide significant SNPs, one SNP named 'rs427724890' on OAR2 showed chromosome-wide association with two type traits that are RW and GRSV. The associated SNP was found to be located only 1-2 Kb upstream from the gene *TLE4* which encodes *Transducin-like enhancer protein 4* that is a transcriptional corepressor. Functional annotation analysis regarding KEGG pathways showed that it involves in Wnt and Notch signalling pathways. As a prominent regulator of the canonical Wnt signalling pathway (GO: 0016055) and cellular biosynthetic process (GO: 0031326) , TLE4 was recently also suggested to regulate muscle satellite cell quiescence and muscle differentiation by repressing Pax7-mediated *Myf5* transcriptional activation in mice (Agarwal et al., 2022). Indeed, another significantly conspicuous finding of our study is the abundance of the SNPs on OAR2 that are associated with various linear type traits. In total, 5 individual SNPs on OAR2 were associated with 4 different linear type traits that are RW, TS, GRSV and RLSV. The candidate genes for those suggested SNPs were *TLE4* for RW and GRSV as mentioned above, *transmembrane protein 50A (TMEM50A)* and *LOC105606997* for tail size (TS) and *myomesin 3 (MYOM3)* and *solute carrier family 44 member 1 (SLC44A1)* for RLSV. Various TMEM's, including *TMEM8B* on OAR2 have been suggested to have relationships with body weight and body length in sheep (Cinar et al., 2020; Tao et al., 2020), while *TMEM50A* was also suggested as a candidate for further investigations as a result of selection signature scan conducted by Fariello et al. (2014) on worldwide sheep populations. On the other hand, *MYOM3* protein is a member of structural proteins that are found within the M-band of the striated muscle sarcomeres with protein (GO: 0005515) and specifically actin

filament binding activity (GO: 0051015). Therefore, considering also the previously associated regions on OAR2 regarding sheep growth and conformation (Cinar et al., 2020; Gutiérrez-Gil et al., 2011; Wang et al., 2016; Zhang et al., 2013), it might be fair to suggest that OAR2 plays an important role in the phenotypic determination of growth and type traits. Finally, SNP 'rs405451961' on the 10th exon of the *hyperpolarization activated cyclic nucleotide gated potassium channel 3* (*HCN3)* on OAR1 was associated with BCS in our study. HCN3 is suggested to regulate biological quality –a measurable attribute of an organism, such as size, mass, shape, colour- (GO: 0065008) and also to involve in the regulation of membrane potential (GO: 0042391). KEGG analysis showed that HCN3 also involves in GnRH secretion.

Functional enrichment of the identified genes considering KEGG pathways also revealed that the gene *SP1* involves in endocrine resistance, cortisol and parathyroid hormone synthesis and secretion and TGF-beta signalling pathway. Additionally, *PRDM2* gene was suggested involve in various metabolic processes including lysine degradation. Finally, the *CPE* gene was suggested to involve in Type I diabetes mellitus whereas *GRID2* is also involved in neuroactive ligand-receptor interaction as is *PTGDR.* All in all, we have found 20 associated SNPs with 20 suggested candidates for which most of their biological importance to the growth and development are well observed. On the other hand, the discovered common SNPs between the traits further support relatively high genetic correlations among them and indicates a shared genetic background. Although sheep genomics had breakthrough improvements recently, as in the publication of various genome assemblies and manufacturing high number of commercial high density SNP arrays, sheep growth and type traits still have only a few studies. Therefore, our results are expected to significantly contribute to the potential knowledge on the genetic background of growth and conformation in sheep which are central to the economic revenue in sheep production systems.

4.1.3 Comparative Genomic Characterization

Indigenous sheep are among the main pillars of food security. Therefore, the characterization of indigenous sheep are quite important for the breeding programmes both based on genomic markers such as genomic selection or traditional selection methods, QTL discovery studies for economically important traits and conservation programmes. In fact assessment of population structure and genetic diversity has been suggested to suggested to define a framework for genomic selection in sheep (Wilson et al., 2022). Recent advances in high-throughput genomic technologies and statistical information processing approaches led to the genomic characterization studies in those indigenous breeds with thousands of genomic markers processed all together. Identification of genetic diversity using genome-wide DNA markers are especially useful when pedigree records of the breeds are not available (Al-Mamun et al., 2015b; Deniskova et al., 2019). In a recent study, a genomic dataset of 74 globally divergent sheep breeds were compiled and 50K SNP genotypes of has been made available to the researchers with relevant diversity and characterization statistics of those breeds (Kijas et al., 2012). Hence, the aim of this study was to unravel within-breed genetic diversity and implement comparative genomic characterisation of Akkaraman sheep against six fat- and thin-tailed indigenous breeds, five mainstream transboundary sheep and two of the main wild breeds from around the world which are collected by the studies conducted by (Ciani et al., 2015; Kijas et al., 2012; Rochus et al., 2018).

Information regarding genetic diversity is crucially important for traditional and genomic selection programmes as well as conservation programmes developed for the sustainable use of animal genetic resources (Narzhan et al., 2022; Scherf & Pilling, 2015)**.** In terms of genetic diversity, expected heterozygosity (H_E), unbiased observed heterozygosity (H_O) and average inbreeding coefficients as Wright's F_{IS} (Wright, 1951) were estimated in this study, since these are the most common parameters in use (FAO, 2011). Overall, high heterozygosities (H_O) were observed for the analysed 12 breeds including Akkaraman sheep, where only the two wild breeds, namely Europian Mouflon (i.e., *Ovis aries musimon* = EMU) and Asiatic Mouflon (i.e., *Ovis orientalis* = OOR), turned out to have quite low genetic diversities. Except for actually having low diversity, one of the main reasons for that could be that these two breeds were not used to develop the genome-wide SNP arrays, which leads to ascertainment bias in the

concluded arrays (Kijas et al., 2012). Studies conducted by Kijas et al. (2012) found slightly higher expected heterozygosities (H_E) for AUM (0.37), TIB (0.34), CFT (0.31), KRS (0.34), MOG (0.36) , SKZ (0.31), CHI (0.32), LMI (0.36), RMB (0.38) breeds. Besides, various other expected heterozygosity measurements were revealed for Australian Suffolk (0.37), Irish Suffolk (0.33), Scottish Texel (0.33), New Zealand Texel (0.34) and German Texel (0.35) by the same study, which was also considerably higher than the estimates in our study. Since the studies share mainly a common sample set excluding Akkaraman sheep that is solely genotyped by our study as well as Suffolk and Texel populations, the differences could be attributed to the SNP marker sizes with 49,034 SNPs used by (Kijas et al., 2012), while in our study overall 29,481 SNPs were used, which are common among the two genotyping platforms and are left after QC of these samples and Akkaraman sheep. For the F_{IS}, excess of negative values indicates abundance of heterozygotes in the analysed 13 breeds which is confirmed by the observed heterozygosities exceeding expected heterozygosities for individual breed evaluations. One of the reasons that OOR showed a positive value for the F_{IS} estimations could be due to the deficit of sample size. On the other hand, obtained heterozygosities and inbreeding coefficients (F_{IS}) indicate an excess of heterozygotes in most of the breeds including Akkaraman sheep which is advantageous in terms of population bottlenecks, future threats of extinction and eventually contribution to the sustainability of sheep production systems (Nielsen, & Slatkin, 2013; Simm, 1998).

According to the PCA plot, analysis of the fourteen breeds showed clear stratification into four separate groups whereas EMU and OOR were remained distinct from all other breeds, of which EMU was the furthest diverged breed and OOR was the second that was located away from those groups, though with a closer position on the plot. Fat-tailed breeds such as Karakas, Cyprus Fat Tail, Moghani and Tibetan sheep, which are breeds that are classified to be Asiatic by Kijas et al. (2012) formed a strong sub-cluster in the PCA plot by overlapping also Akkaraman sheep. Previously, strong evidence suggested that Karakas is a variety of Akkaraman sheep which is also supported by the results of our study (Yalcin, 1986; Ozmen et al., 2020). Furthermore, Sakiz and Chios breeds, which are closest to the sub-group created by those fat tailed Asian breeds, are known to share the same breed origin (i.e., Chios) with recent geographic isolation in

Chios Island and Türkiye as it is also supported by our findings as well as other studies (Kijas et al., 2012; Sönmez, 1962). Interestingly, Cyprus Fat Tail previously was found to be closer to the sub-group formed by Chios and Sakiz rather than the Moghani sheep (Kijas et al., 2012). In our study, F_{ST} values and phylogenetic tree were also supportive of the arguments proposed above, Akkaraman has the lowest differentiation with those other fat tailed breeds while. Furthermore, Suffolk and Texel, as Northern European breeds, were also grouped together previously by a phyologenetic tree created by (Rochus et al., 2018). At a global scale, observations derived from the PCA graph, Fst values and phylogenetic tree display strong genetic divisions for fat-tailed sheep (i.e, Akkaraman, Karakas, Moghani, Tibetan and Cyprus Fat Tail) and sheep with Merino background (i.e., Australian Merino, Milk Lacaune and Rambouillet) as well as those with Eastern Europe origin (Suffolk and Texel) and other thin-tailed sheep (i.e., Chios and Sakiz). Hence based on the findings suggested by those analytical tools it is possible to speculate that those clustered breeds share the same ancestral origin when considered within groups.

Existence of linkage disequilibrium (LD) between any two markers clearly reflect non-random segregation of alleles at those two loci. Estimates of LD lays foundations for Genome-wide Association Studies (GWAS) as well as those marker-assisted selection decisions and genomic selection (i.e. Genome-wide Selection, Genomic Prediction) practices (Al-Mamun et al., 2015b). Effective population size and evolutionary forces such as selection, mutation and admixture as well as recombination events affects the speed of LD decay and the extent of LD (Wang et al., 2005). Therefore, comparison of LD among the breeds of a species allow to speculate about the overall diversity levels as well as to understand the nature of previous selection decisions applied to those breeds. In our study, LD estimations were both implemented for two adjacent SNPs, which on average have 88 Kb pairwise distance, as well as LD decay until 500 Kb distance between pairs of markers. Akkaraman sheep came out to have one of the lowest LD among the analysed breeds right after Asiatic mouflon, which is also an indicator of lower selection pressure, higher effective population size (Ne) and recombination rate in Akkaraman sheep (Qanbari, 2020). Although sheep overall have low LD estimates compared to the other species and although for a smaller physical distance such as 50-

60 Kb, previously lower LD estimates than that of obtained by our study for Akkaraman sheep were observed in Merino, Chinese Merino, Texel, Romney, Valle del Belice, Pinzirita, Comisana and various crossbreds with faster LD decay patterns (Al-Mamun et al., 2015b; Liu et al., 2017; Mastrangelo et al., 2014; Prieur et al., 2017) .

Sheep populations that are exposed inbred mating for over a long period of time have a genomic structure indicating rise of autozygosity, which also give rise to the uninterrupted runs of homozygosity (ROH). Various other forces also lead to the ROHs in the genome, which mainly are intensive selection pressure, population bottleneck and the following founder effect and genetic drift (Peripolli et al., 2017). Akkaraman sheep has been observed to hold the least average number of ROHs (41.10) per individual among the 14 breeds of the study, while still higher than those obtained for Merino and Poll Dorset by (Al-Mamun et al., 2015b), Tunisian sheep by (Ahbara et al., 2021). The average length of ROH observed in Akkaraman was also the second lowest (3,160 Kb) and estimated genomic inbreeding F_{ROH} (0.002) was the lowest compared to the other breeds in the study, which is also supportive of its relatively high diversity observed. The F_{ROH} estimation for Akkaraman sheep was significantly lower than the estimates of many breeds from around the world such as indigenous breeds from Kyrgyzstan and many South African and global sheep populations (Deniskova et al., 2019; Dzomba et al., 2021). Regarding the Akkaraman sheep breed, information provided by our study on the genetic diversity, population structure, and extent of LD and ROH in the genome of the breed will help further GWA studies, conservation programmes and genomic selection programmes to be designed in the most efficient way. In genomic selection especially, the GEBV accuracies mostly depend on the trait heritabilities, genetic architecture and the effective population size of the targeted genome. Our study observed low levels of LD and quite high levels of diversity, which may lead to suggest that the accuracies of GEBVs in this sheep breed could be lower than populations with higher LD exposed under similar scenarios. Therefore, for those studies and further characterization studies larger sample and denser marker sets are suggested for efficient implementation and increased accuracy in Akkaraman sheep.

4.2 Conclusions

This research aimed to uncover the genetic basis of growth and conformation with the genomic analysis of 17 traits observed in Akkaraman sheep. Correspondingly, moderate genomic heritability estimates and overall high and positive genetic/phenotypic correlations were found between growth and linear type traits. This is an indicator of high enough genetic variance for being exploited by the sheep breeders for faster genetic progress of the focused traits, while identified genetic/phenotypic correlations among the traits also can help defining potential of the traits for each other to be an indicator phenotype for early and easy detection of the breeding male/females. Considering the wide influence of growth and linear type traits on other traits of economic importance such as robustness, disease susceptibility, productivity and welfare, the results have prominent indications in terms of designing efficient selection indices. On the other hand, GWA analyses led to the discovery of 2 genome- and 18 chromosome-wide significant SNPs associated with the traits. Consequently, the genes *PRDM2, MYOM3, PTGDR, TLE4, PTPRG, FKBP1A, KCND2, ZNF260, TMEM50A, CPE*, *GRID2, HCN3, SCD5, ZNF407, SPIDR, SP1*, and *SLC44A1* were suggested as bearing putative QTL underlying growth and linear type traits. Among those discovered genes, certain were found to be common between the focused traits, which points out a pleiotropy for the genetic architecture of those traits. Moreover, 4 of the proposed genes were noticed to be located on OAR2, which also emphasizes the importance of the OAR2 for growth and morphological development in sheep, alongside with the previous studies. Therefore, the results carry important implications sheep production systems. Once the results are biologically validated by expression and gene-knockout studies as well as with different populations, detected information can be used in terms of marker-assisted selection programmes, understanding the genetic architecture of growth and linear type traits and pointing out loci for targeted genome-editing studies. To the extent of our knowledge, this study is among the very first studies focused on the genetic parameters and the genetic basis of growth and linear type traits in an indigenous sheep breed.

We also genomically characterized the Akkaraman breed against various world-renowned transboundary sheep and indigenous sheep with fat and thin-tails. We have displayed genetic similarities and differences between those breeds by estimating and comparing various genetic diversity indices, linkage disequilibrium (LD) estimates and fixation index (F_{ST}), runs of homozygosity (ROH) as well as PCA and Neigbor-Joining tree analysis. Akkaraman sheep were found to be clustered with Moghani, Karakas, Tibetan and Cyprus Fat Tail sheep, which are mainly the only fat-tailed sheep in the study, both as a result of PCA and Neigbor-Joining tree analysis. These are also supported by F_{ST} estimations. In an evolutionary perspective, those breeds are also expected to behave genetically similar regarding the traits focused by our study, which is a hypothesis remained to be validated by further studies. In any case, the large genetic diversity found by our study for Akkaraman sheep validates the existence of a large genetic variance to implement a comprehensive selection on the breed, without a great danger of imposing an inbreeding depression in the near future. The fact that Akkaraman sheep has the lowest genomic inbreeding and one of the lowest number of ROHs also reveals that the breed has not been exposed to any historical intensive selection pressure, inbred mating or a massive population bottleneck for a long period of time to leave strong marks of genomic homozygosity. The results of the current study improve our understanding of the genetic diversity in Akkaraman sheep in comparison with certain mainstream sheep breeds as well as those indigenous breeds from around the world. Additionally, our findings will be useful to perform further GWAS studies effectively both considering population structure, diversity and LD patterns observed among the breeds. Finally, results of our study also provide valuable insights into the influence of long term selection applied on these breeds and provide practical knowledge that will contribute to designing efficient and successful genome-based selection programmes.

4.3 Future Perspectives

Growth and linear type traits are of direct economic interest to those sheep production systems aiming sustainability, robustness, efficiency and profitability. High-throughput genomic information in livestock has revolutionized our concept of breeding and

management of animal genetic resources profoundly. Therefore, in this study we aimed to uncover underlying genetic architecture of those traits to render genomic information based selection possible in Akkaraman sheep. Doing so, we also prioritized unravelling the current state of genetic diversity, population structure and genetic differentiation of the breed in comparison with several other sheep breeds. Therefore, parameters concerning an index for optimised selection that combines rates of genetic gain with the track of genetic diversity are now available for the breed's genome-based selection programmes in addition to the full characterization of the breed against several other world-renowned transboundary sheep as well as indigenous ones. Nevertheless, here we would like to point out certain perspectives building around the discoveries and unravelled information by our study for the future studies.

First of all, to our knowledge this is the first study focusing on the genetic parameter estimations and genome-wide discovery of QTL regarding growth and linear type traits in Akkaraman sheep using genome-wide SNP panels. Therefore, to validate and strengthen the reliability of the results, various other Akkaraman populations with larger sample size is suggested to be used for further studies concerning the phenotypes of interest, which will increase statistical power of the associations and potentially help those suggestive SNPs discovered by our study to reach genome-wide significance. Additionally, it is important to note that here heritabilities, genetic correlations and the SNP effects were estimated considering additivity of SNP effects. Therefore, a follow-up study would be to unravel dominance and epistasis effects of those variants. On the other hand, to implement a comprehensive selection programme it is obviously suggested to also discover genetic and phenotypic correlations with other traits such as fleece and/or milk production, disease and parasite susceptibility, feed efficiency, fertility, various physiological parameters such as blood parameters and molecular phenotypes associated with resilience, robustness and efficiency. Akkaraman and most other indigenous sheep breeds known to significantly lack tracking of such data, which brings us to another important issue which is the data recording. The lack of systematic record keeping for sheep breeds such as Akkaraman significantly hinders genetic evaluations and therefore faster genetic improvement of the animals. This problem

should be overcome preceding implementation of any intensive selection programmes. Methods such as Genomic Selection (i.e., Genome-wide Selection) partially relieve the lack of recording schemes for the pedigree part, as genomic relationship matrices can replace average relationship matrix constructed from pedigree (Meuwissen et al., 2001; Rupp et al., 2016; VanRaden, 2008). However, even genomic selection methods initially need mass phenotyping, which is in fact both a technical and economic prerequisite for any genomic breeding programme. The results obtained by the current study would improve the accuracy of genetic evaluations and contribute to the economic outcomes of the Akkaraman production systems. However, the next important step would be to upgrade the selection methods into genomic selection which is known to increase genetic gain par generation up to 40 % in sheep production systems, as well as other advantages such as tracking total genetic variance at once instead of a couple of low-influence loci regarding traits of interest and overcoming the lack of pedigree (Fugeray-Scarbel et al., 2021; Rupp et al., 2016; Shumbusho et al., 2014, 2017). While doing that, an optimized method for resource allocation between phenotyping and genotyping is suggested as implemented for dairy cattle (Obšteter et al., 2021). Nevertheless, considering that the number of QTL discovery studies focused on growth and type traits are very scarce in sheep, it is extremely important to implement also more discovery studies to shed a light on the complex molecular mechanisms behind growth and type traits with the aim of providing the industry with efficient targets for the traits, either for selection or editing.

In our study, genome-wide representative SNP panels were used. It is also shown by our study that linkage disequilibrium between adjacent SNPs are quite low in sheep which hampers the discovery of most variants especially those that are rare variants with large effect. To overcome this problem, use of Next Generation Sequencing (NGS) methods is suggested to generate genomic data for both fine-mapping of the regions discovered by our study and whole genome sequencing. Various genotyping-by-sequencing (GBS) methods and methods such as ‘low-pass sequencing’ followed by imputation are proposed for the reduction of sequencing costs and de-novo variant discovery (Gurgul et al., 2019; Snelling et al., 2020)

It is known that economically important traits such as growth and linear type traits are quantitative in nature, which means genome, environment and their reciprocal interactions are decisive on the phenotypic distribution of those traits. Therefore, it is important to investigate other levels of the biological process towards phenotypic determination with multi-omics data (e.g., transcriptomics, metabolomics, epigenomics, proteomics), which will evolve research on livestock production systems into a systems biology approach (Kizilaslan et al., 2022a). Evidence suggest that combining various sources of high-throughput data contribute discovery of QTL with higher statistical power for many diseases and metabolic conditions. For example, there are studies combining transcriptomics, metabolomics and GWAS for an efficient discovery of molecular mechanisms underlying metabolite levels that have high correlation with certain conditions and disease risks (Yin et al., 2022). Therefore, we suggest that improving mathematical, statistical and biological theory of estimation selection indices with multi-omics data such as those accounting for epigenomic patterns and expression profiles of the genes as well as metabolomics profiling of certain status will also contribute faster genetic progress of growth and linear type traits in sheep. Also bioinformatics pipelines should also be constructed to accommodate multilayer 'omics' data and to process into meaningful results. We also believe that, long-known problem of missing heritability in quantitative traits will also benefit from such approaches. On the other hand, another level of a unifying approach as well as systems biology would be the research with 'One Health' concept, since the final consumers of the livestock products are the human population with extreme diversity.

Finally, genome-wide association studies, selective breeding involving marker-assisted selection with the discovered QTL and genomic selection have undoubtedly had a revolutionary and transformational impact on livestock production systems even with those local ruminant populations and smallholder models (Burrow et al., 2021; Ducrocq et al., 2018). Those are all examples of the so-called 'forward genetics'. Tools such as CRISPR-Cas9 allow researchers to upgrade ruminant production systems through the use of 'reverse genetics' into Livestock 2.0 (Tait-Burkard et al., 2018). Genome-editing

techniques such as CRISPR-Cas9 are tailored for the manipulation of genome to get more productive, robust and healthier farm animals. Once the putative QTL suggested by our study are validated and fine-mapped, those will be open targets for genome-editing techniques to manipulate populations into desired directions. Furthermore, our results report no alarming levels of genetic diversity for Akkaraman sheep compared to the others. Nonetheless, in case of future reductions due to unexpected bottlenecks, it is possible to introduce new genomic variation via genome-editing tools which altogether offers exciting opportunities for the future of the breed and livestock production systems. One important issue and relevant solutions regarding the use of genome-editing techniques in economically important traits of livestock is raised by (Jenko et al., 2015). Since most of the focused traits, including growth and linear type traits in sheep, are quantitative traits which are controlled by many QTL or simply quantitative trait nucleotides, authors of the study suggested an editing approach referred as promotion of alleles by genome editing (PAGE) to be used in combination with genomic selection for substantially higher genetic gains as response to selection. Both short and long term responses were observed to be increased, which is quite promising for livestock breeding programmes aiming faster genetic improvement in a short period of time.

Thus, using all those forward and reverse genetic techniques for the breeding programmes and management of genetic resources for both Akkaraman sheep and other livestock production systems, is expected to help relieve of food security concerns in the face of climate change, recurrent pandemics and other uncertainties awaiting for the ever-growing human population in the near future.

5 REFERENCES

Agarwal, M., Bharadwaj, A., & Mathew, S. J., 2022. TLE4 regulates muscle stem cell quiescence and skeletal muscle differentiation. **Journal of Cell Science**, 135(4): jcs256008. https://doi.org/10.1242/jcs.256008

Ahbara, A. M., Rouatbi, M., Gharbi, M., Rekik, M., Haile, A., Rischkowsky, B., & Mwacharo, J. M., 2021. Genome-wide insights on gastrointestinal nematode resistance in autochthonous Tunisian sheep. **Scientific Reports**, 11(1): 1–16. https://doi.org/10.1038/s41598-021-88501-3

Al-Mamun, Hawlader A., Kwan, P., Clark, S. A., Ferdosi, M. H., Tellam, R., & Gondro, C., 2015a. Genome-wide association study of body weight in Australian Merino sheep reveals an orthologous region on OAR6 to human and bovine genomic regions affecting height and weight. **Genetics Selection Evolution,** 47(1): 1–11. https://doi.org/10.1186/s12711-015-0142-4

Al-Mamun, Hawlader Abdullah, A Clark, S., Kwan, P., & Gondro, C., 2015b. Genome-wide linkage disequilibrium and genetic diversity in five populations of Australian domestic sheep. **Genetics Selection Evolution**, 47(1): 1–14. https://doi.org/10.1186/s12711-015-0169-6

Ashwell, M. S., Heyen, D. W., Weller, J. I., Ron, M., Sonstegard, T. S., Van Tassell, C. P., & Lewin, H. A., 2005. Detection of quantitative trait loci influencing conformation traits and calving ease in Holstein-Friesian cattle. **Journal of Dairy Science,** 88(11): 4111–4119. https://doi.org/10.3168/jds.S0022-0302(05)73095-2

Astle, W., & Balding, D. J., 2009. Population structure and cryptic relatedness in genetic association studies. **Statistical Science**, 24(4): 451–471.

Aulchenko, Y. S., Ripke, S., Isaacs, A., & van Duijn, C. M., 2007. GenABEL: an R library for genome-wide association analysis. **Bioinformatics**, 23(10): 1294–1296. https://doi.org/10.1093/bioinformatics/btm108

Yalcin, B.C., 1986. Fao Animal Production And Protection Paper, Sheep And Goats In Turkey. In The Food and Agriculture Organization (FAO), pp 183. ISBN 92-5-102449-9).

Baeza, M. C., Corva, P. M., Soria, L. A., Pavan, E., Rincon, G., & Medrano, J. F., 2013. Genetic variants in a lipid regulatory pathway as potential tools for improving the nutritional quality of grass-fed beef. **Animal Genetics**, 44(2): 121–129. https://doi.org/10.1111/j.1365-2052.2012.02386.x

Bayram, D., Akyüz, B., Arslan, K., Özdemir, F., Aksel, E. G., & Çinar, M. U., 2019. DGAT1, CAST and IGF-I gene polymorphisms in akkaraman lambs and their effects on live weights up to weaning age. **Kafkas Universitesi Veteriner Fakultesi Dergisi**, 25(1), 9–15. https://doi.org/10.9775/kvfd.2018.20055

Behrem, S., 2021a. Estimation of genetic parameters for pre-weaning growth traits in Central Anatolian Merino sheep. **Small Ruminant Research,** 197: 106319.

Behrem, S., 2021b. Effects of Environmental Factors Growth Traits of Akkaraman Sheep in Çankırı Province. **Livestock Studies**, 61(1): 22–27.

Binns, D., Dimmer, E., Huntley, R., Barrell, D., O'donovan, C., & Apweiler, R.,2009. QuickGO: a web-based tool for Gene Ontology searching. **Bioinformatics**, 25(22): 3045–3046.

Biscarini, F., Cozzi, P., Gaspa, G., & Marras, G., 2018. detectRUNS: Detect runs of homozygosity and runs of heterozygosity in diploid genomes. (**https://CRAN.R-project.org/package=detectRUNS**), (Access date: July, 2022)

Biscarini, F., Nicolazzi, E., Alessandra, S., Boettcher, P., & Gandini, G., 2015. Challenges and opportunities in genetic improvement of local livestock breeds. **Frontiers in Genetics,** 5(1): 1–16. https://doi.org/10.3389/fgene.2015.00033

Boettcher, P. J., Hoffmann, I., Baumung, R., Drucker, A. G., McManus, C., Berg, P., Stella, A., Nilsen, L. B., Moran, D., Naves, M., & Thompson, M., 2014. Genetic resources and genomics for adaptation of livestock to climate change. **Frontiers in Genetics**, 5:461. https://doi.org/10.3389/fgene.2014.00461

Bouška, J., Vacek, M., Štípková, M., & Němec, A., 2006. The relationship between linear type traits and stayability of Czech Fleckvieh cows. **Czech Journal of Animal Science**, 51(7): 299–304. https://doi.org/10.17221/3942-cjas

Breusch, T. S., & Pagan, A. R., 1979. A Simple Test for Heteroscedasticity and Random Coefficient Variation. **Econometrica**, 47(5): 1287–1294. https://doi.org/10.2307/1911963

Brotherstone, S., 1994. Genetic and phenotypic correlations between linear type traits and production traits in Holstein-Friesian dairy cattle. **Animal Science,** 59(2): 183–187.

Burrow, H. M., Mrode, R., Mwai, A. O., Coffey, M. P., & Hayes, B. J., 2021. Challenges and Opportunities in Applying Genomic Selection to Ruminants Owned by Smallholder Farmers. **Agriculture (Switzerland)**, 11(11): 1–17. https://doi.org/10.3390/AGRICULTURE11111172

Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D. P., Green, E. K., Consortium, W. T. C. C., Burton, P. R.,

Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D. P., McCarthy, M. I., Ouwehand, W. H., Samani, N. J., … Compston, A. A., 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. **Nature**, 447(7145): 661. https://doi.org/10.1038/nature05911.Genome-wide

Campos, R. V., Cobuci, J. A., Kern, E. L., Costa, C. N., & McManus, C. M., 2015. Genetic parameters for linear type traits and milk, fat, and protein production in holstein cows in Brazil. **Asian-Australasian Journal of Animal Sciences,** 28(4): 476–484. https://doi.org/10.5713/ajas.14.0288

Cardon, L. R., & Bell, J. I., 2001. Association study designs for complex diseases. **Nature Reviews. Genetics**, 2(2): 91–99. https://doi.org/10.1038/35052543

Carta, A., Casu, S., & Salaris, S., 2009. Current state of genetic improvement in dairy sheep. **Journal of Dairy Science**, 92(12): 5814–5833. https://doi.org/10.3168/jds.2009-2479

Casas, E., Shackelford, S. D., Keele, J. W., Stone, R. T., Kappes, S. M., & Koohmaraie, M., 2000. Quantitative trait loci affecting growth and carcass composition of cattle segregating alternate forms of myostatin. **Journal of Animal Science**, 78(3): 560–569. https://doi.org/10.2527/2000.783560x

Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M., Melino, G., & Raschellà, G., 2017. Zinc-finger proteins in health and disease. **Cell Death Discovery**, 3(1), 17071. https://doi.org/10.1038/cddiscovery.2017.71

Charlesworth, B., & Charlesworth, D. (2010). Elements of Evolutionary Genetics. Roberts Publishers , pp 734..

Chen, W.-M., & Abecasis, G. R., 2007. Family-based association tests for genomewide association scans. **The American Journal of Human Genetics**, 81(5): 913–926.

Ciani, E., Lasagna, E., D'Andrea, M., Alloggio, I., Marroni, F., Ceccobelli, S., Delgado Bermejo, J. V., Sarti, F. M., Kijas, J., Lenstra, J. A., & Pilla, F., 2015. Merino and Merino-derived sheep breeds: A genome-wide intercontinental study. **Genetics**

Selection Evolution, 47(1). https://doi.org/10.1186/s12711-015-0139-z.

Cinar, M. U., Mousel, M. R., Herndon, M. K., Taylor, J. B., & White, S. N., 2020. Association of TMEM8B and SPAG8 with mature weight in sheep. **Animals**, 10(12): 1–8. https://doi.org/10.3390/ani10122391

Covarrubias-Pazaran, G., 2016. Genome-Assisted prediction of quantitative traits using the r package sommer. **PLoS ONE**, 11(6): 1–15. https://doi.org/10.1371/journal.pone.0156744

Crispim, B. A., Grisolia, A. B., Seno, L. O., Egito, A. A., Vargas Junior, F. M., & Souza, M. R., 2013. Genetic diversity of locally adapted sheep from Pantanal region of Mato Grosso do Sul. **Genetics and Molecular Research**, 12(4): 5458–5466. https://doi.org/10.4238/2013.November.11.7

Cromie, A., Wickham, B. W., Coughlan, S., & Burke, M., 2008. The impact of new technologies on performance recording and genetic evaluation of dairy and beef cattle in Ireland. **Proc. ICAR Biennial Conference**, Niagara.

de la Fuente, L. F., Gonzalo, C., Sánchez, J. P., Rodríguez, R., Carriedo, J. A., Primitivo, F. S. F. F. S., Fuente, L., Gonzalo, C., Sanchez, J. P., Rodríguez, R., Carriedo, J. A., Primitivo, F. S. F. F. S., de la Fuente, L. F., Gonzalo, C., Sánchez, J. P., Rodríguez, R., Carriedo, J. A., & Primitivo, F. S. F. F. S., 2011. Genetic parameters of the linear body conformation traits and genetic correlations with udder traits, milk yield and composition, and somatic cell count in dairy ewes. **Canadian Journal of Animal Science**, 91(4): 585–591. https://doi.org/10.4141/cjas2010-031

Dekkers, J. C. M., 2007. Prediction of response to marker-assisted and genomic selection using selection index theory. **Journal of Animal Breeding and Genetics,** 124(6): 331–341. https://doi.org/10.1111/j.1439-0388.2007.00701.x

Deniskova, T., Dotsev, A., Lushihina, E., Shakhin, A., Kunz, E., Medugorac, I., Reyer, H., Wimmers, K., Khayatzadeh, N., Sölkner, J., Sermyagin, A., Zhunushev, A., Brem, G., & Zinovieva, N., 2019. Population Structure and Genetic Diversity of Sheep Breeds in the Kyrgyzstan. **Frontiers in Genetics**, 10(12): 1–15.

https://doi.org/10.3389/fgene.2019.01311

Devlin, B., & Roeder, K., 1999. Genomic control for association studies. **Biometrics**, 55(4): 997–1004.

Doormaal, B. Van., 2012. Increased Rates of Genetic Gain with Genomics. **Canadian Dairy Network**, July, 1–3.

Ducrocq, V., Laloe, D., Swaminathan, M., Rognon, X., Tixier-Boichard, M., & Zerjal, T., 2018. Genomics for ruminants in developing countries: From principles to practice. **Frontiers in Genetics**, 9(7): 1–7. https://doi.org/10.3389/fgene.2018.00251

Dzomba, E. F., Chimonyo, M., Pierneef, R., & Muchadeyi, F. C., 2021. Runs of homozygosity analysis of South African sheep breeds from various production systems investigated using OvineSNP50k data. **BMC Genomics**, 22(1): 1–14. https://doi.org/10.1186/s12864-020-07314-2

Erdenee, S., Li, J., Kang, Z., Xu, H., Zang, R., Cao, X., Yang, J., Cai, Y., & Lan, X., 2020. Sheep zinc finger proteins 395 (ZNF395): insertion/deletion variations, associations with growth traits, and mRNA expression. **Animal Biotechnology**, 31(3): 237–244. https://doi.org/10.1080/10495398.2019.1585865

Falconer, D. S., & Mackay, T. F.C., 1996. In Introduction to Quantitative Genetics (Vol. 4). Essex: Benjamin Cummings, pp 448.

FAO., 2011. Molecular genetic characterization of animal genetic resources. **FAO Anim. Product. Health Guidelines**, 9: 7–25.

FAO., 2015. The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture, edited by B.D. Scherf & D. Pilling. FAO Commission on Genetic Resources for Food and Agriculture Assessments. Rome.

FAO., 2022a. Domestic Animal Diversity Information System (DAD-IS). Rome, Italy: FAO. https://www.fao.org/dad-is/data/en/

FAO., 2022b. FAOSTAT Database. Rome, Italy: FAO.

https://www.fao.org/faostat/en/#data/QCL

FAO Commission on Genetic Resources for Food and Agriculture Assessments., 2007. Global plan of action for animal genetic resources and the Interlaken declaration. **World's Poultry Science Journal**, 286. http://www.journals.cambridge.org/abstract_S0043933909000245

Fariello, M. I., Servin, B., Tosser-Klopp, G., Rupp, R., Moreno, C., Cristobal, M. S., Boitard, S., Arranz, J. J., Banos, G., Barendse, W., El Beltagy, A., Benenwitz, J., Bishop, S., Bunger, L., Calvo, J., Carta, A., Cemal, I., Ciani, E., Cockett, N., … Whan, V., 2014. Selection signatures in worldwide sheep populations. **PLoS ONE,** 9(8). https://doi.org/10.1371/journal.pone.0103813

Fernández, G., Baro, J. A., De La Fuente, L. F., & San Primitivo, F., 1997. Genetic Parameters for Linear Udder Traits of Dairy Ewes. **Journal of Dairy Science,** 80(3): 601–605. https://doi.org/10.3168/jds.S0022-0302(97)75976-9

Fisher, R. A., 1919. XV.—The correlation between relatives on the supposition of Mendelian inheritance. **Earth and Environmental Science Transactions of the Royal Society of Edinburgh**, 52(2): 399–433.

Fugeray-Scarbel, A., Bastien, C., Dupont-Nivet, M., & Lemarié, S., 2021. Why and How to Switch to Genomic Selection: Lessons From Plant and Animal Breeding Experience. **Frontiers in Genetics**, 12(7): 1–9. https://doi.org/10.3389/fgene.2021.629737

Gebreyesus, G., Buitenhuis, A. J., Poulsen, N. A., Visker, M. H. P. W., Zhang, Q., Van Valenberg, H. J. F., Sun, D., & Bovenhuis, H., 2019. Multi-population GWAS and enrichment analyses reveal novel genomic regions and promising candidate genes underlying bovine milk fatty acid composition. **BMC Genomics**, 20(1), 1–16. https://doi.org/10.1186/s12864-019-5573-9

Gholizadeh, M., Rahimi-Mianji, G., & Nejati-Javaremi, A., 2015. Genomewide association study of body weight traits in Baluchi sheep. **Journal of Genetics**, 94(1), 143–146. https://doi.org/10.1007/s12041-015-0469-1

Goddard, M. E., & Hayes, B. J., 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nature Reviews Genetics**, 10(6), 381–391. https://doi.org/10.1038/nrg2575

Groeneveld, L. F., Lenstra, J. A., Eding, H., Toro, M. A., Scherf, B., Pilling, D., Negrini, R., Finlay, E. K., Jianlin, H., Groeneveld, E., & Weigend, S., 2010. Genetic diversity in farm animals - A review. **Animal Genetics**, 41(1): 6–31. https://doi.org/10.1111/j.1365-2052.2010.02038.x

Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A., 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. **Molecular Ecology Resources**, 18(3): 691–699.

Gurgul, A., Miksza-Cybulska, A., Szmatoła, T., Jasielczuk, I., Piestrzyńska-Kajtoch, A., Fornal, A., Semik-Gurgul, E., & Bugno-Poniewierska, M., 2019. Genotyping-by-sequencing performance in selected livestock species. **Genomics,** 111(2): 186–195. https://doi.org/https://doi.org/10.1016/j.ygeno.2018.02.002

Gutiérrez-Gil, B., Alvarez, L., de la Fuente, L. F., Sanchez, J. P., San Primitivo, F., & Arranz, J. J., 2011. A genome scan for quantitative trait loci affecting body conformation traits in Spanish Churra dairy sheep. **Journal of Dairy Science,** 94(8): 4119–4128. https://doi.org/10.3168/jds.2010-4027

Hill, W G, & Robertson, A., 1968. Linkage disequilibrium in finite populations. Theoretical and **Applied Genetics,** 38(6): 226–231. https://doi.org/10.1007/BF01245622

Hill, William G., 2014. Applications of population genetics to animal breeding, from wright, fisher and lush to genomic prediction. **Genetics,** 196(1): 1–16. https://doi.org/10.1534/genetics.112.147850

Hoffmann, I., 2010. Climate change and the characterization, breeding and conservation of animal genetic resources. **Animal Genetics,** 41(1): 32–46. https://doi.org/10.1111/j.1365-2052.2010.02043.x

Hu, Z.-L., Park, C. A., & Reecy, J. M., 2022. Bringing the Animal QTLdb and CorrDB

into the future: meeting new challenges and providing updated services. **Nucleic Acids Research**, 50(1): 956–961. https://doi.org/10.1093/nar/gkab1116

Huang, D. W., Sherman, B. T., & Lempicki, R. A., 2009a. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. **Nature Protocols,** 4(1): 44–57. https://doi.org/10.1038/nprot.2008.211

Huang, D. W., Sherman, B. T., & Lempicki, R. A., 2009b. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. **Nucleic Acids Research,** 37(1): 1–13. https://doi.org/10.1093/nar/gkn923

IPCC., 2021. Climate Change 2021: The Physical Science Basis - Summary for the Policymakers (Working Group I). In Climate Change 2021: The Physical Science Basis.

Janssens, S., & Vandepitte, W., 2004. Genetic parameters for body measurements and linear type trait in Belgian Bleu du Maine, Suffolk and Texel sheep. **Small Ruminant Research**, 54(1–2): 13–24. https://doi.org/10.1016/j.smallrumres.2003.10.008

Jenko, J., Gorjanc, G., Cleveland, M. A., Varshney, R. K., Whitelaw, C. B. A., Woolliams, J. A., & Hickey, J. M., 2015. Potential of promotion of alleles by genome editing to improve quantitative traits in livestock breeding programs. **Genetics Selection Evolution**, 47(1): 1–14. https://doi.org/10.1186/s12711-015-0135-3

Jiang, Y., Xie, M., Chen, W., Talbot, R., Maddox, J. F., Faraut, T., Wu, C., Muzny, D. M., Li, Y., & Zhang, W., 2014. The sheep genome illuminates biology of the rumen and lipid metabolism. **Science,** 344(6188): 1168–1173.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. **Bioinformatics**, 24(11): 1403–1405.

Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. **Trends in Ecology and Evolution**, 17(5): 230–241. https://doi.org/10.1016/S0169-5347(02)02489-8

Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Neto, L. R. P., San, M., Servin, B., Mcculloch, R., Whan, V., Gietzen, K., Paiva, S., Barendse, W., Ciani, E., Raadsma, H., Mcewan, J., & Dalrymple, B., 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. https://doi.org/10.1371/journal.pbio.1001258

Kijas, J. W., Townley, D., Dalrymple, B. P., Heaton, M. P., Maddox, J. F., McGrath, A., Wilson, P., Ingersoll, R. G., McCulloch, R., McWilliam, S., Tang, D., McEwan, J., Cockett, N., Oddy, V. H., Nicholas, F. W., Raadsma, H., Barris, W., Bishop, S. C., Coltman, D., … Schibler, L., 2009. A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. **PLoS One**, 4(3). https://doi.org/10.1371/journal.pone.0004668

Kirikci, K., 2022. Polymorphism of the Calpastatin (CAST) and Growth Differentian Factor 9 (GDF9) genes in Akkaraman Sheep Breed. **Hayvansal Üretim,** 9(1): 21–26. https://doi.org/10.29185/hayuretim.1008768

Kizilaslan, M., Arzik, Y., Cinar, M. U., & Konca, Y., 2022a. Genome-wise engineering of ruminant nutrition – nutrigenomics: applications, challenges, and future perspectives – A review. **Annals of Animal Science**, 22(2): 511–521. https://doi.org/doi:10.2478/aoas-2021-0057

Kizilaslan, M., Arzik, Y., White, S. N., Piel, L. M. W., & Cinar, M. U., 2022b. Genetic Parameters and Genomic Regions Underlying Growth and Linear Type Traits in Akkaraman Sheep. **Genes**, 13(8): 1414. https://doi.org/10.3390/genes13081414

Kominakis, A., Hager-Theodorides, A. L., Zoidis, E., Saridaki, A., Antonakos, G., & Tsiamis, G., 2017. Combined GWAS and 'guilt by association'-based prioritization analysis identifies functional candidate genes for body size in sheep. **Genetics Selection Evolution,** 49(1): 1–16. https://doi.org/10.1186/s12711-017-0316-3

Kristensen, T. N., Hoffmann, A. A., Pertoldi, C., & Stronen, A. V., 2015. What can livestock breeders learn from conservation genetics and vice versa? **Frontiers in**

Genetics, 5(2): 1–12. https://doi.org/10.3389/fgene.2015.00038

Kumar, S., Dahiya, S. P., & Malik, Z. S., 2019. Genetic and phenotypic correlations among linear type traits in Harnali sheep. **The Pharma Innovation Journal**, 8(1): 646–649.

Lee, S. H., & Van der Werf, J. H. J., 2016. MTG2: an efficient algorithm for multivariate linear mixed model analysis based on genomic information. **Bioinformatics**, 32(9): 1420–1422.

Lenstra, J. A., Groeneveld, L. F., Eding, H., Kantanen, J., Williams, J. L., & Taberlet, P., 2012. Molecular tools and analytical approaches for the characterization of farm animal genetic diversity. **Animal Genetics**, 43(5):483-502. https://doi.org/10.1111/j.1365-2052.2011.02309.x

Liu, S., He, S., Chen, L., Li, W., Di, J., & Liu, M., 2017. Estimates of linkage disequilibrium and effective population sizes in Chinese Merino (Xinjiang type) sheep by genome-wide SNPs. **Genes & Genomics**, 39(7): 733–745. https://doi.org/10.1007/s13258-017-0539-2

Loos, R. J. F., 2020. 15 Years of Genome-Wide Association Studies and No Signs of Slowing Down. **Nature Communications,** 11(1): 10–12. https://doi.org/10.1038/s41467-020-19653-5

Lu, Z., Yue, Y., Yuan, C., Liu, J., Chen, Z., Niu, C., Sun, X., Zhu, S., Zhao, H., Guo, T., & Yang, B., 2020. Genome-wide association study of body weight traits in Chinese fine-wool sheep. **Animals,** 10(1): 1–11. https://doi.org/10.3390/ani10010170

Lynch, M., & Walsh, B., 1998. Genetics and Analysis of Quantitative Traits. Oxford University Press, pp 980

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

Mastrangelo, Salvatore, Di Gerlando, R., Tolone, M., Tortorici, L., Sardina, M. T., & Portolano, B., 2014. Genome wide linkage disequilibrium and genetic structure in Sicilian dairy sheep breeds. **BMC Genetics,** 15(1): 1–10. https://doi.org/10.1186/s12863-014-0108-5

McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. A., & Hirschhorn, J. N., 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. **Nature Reviews Genetics**, 9(5), 356–369. https://doi.org/10.1038/nrg2344

McQuillan, R., Leutenegger, A.-L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., Macleod, A. K., Farrington, S. M., Rudan, P., Hayward, C., Vitart, V., Rudan, I., Wild, S. H., Dunlop, M. G., Wright, A. F., … Wilson, J. F., 2008. Runs of homozygosity in European populations. American **Journal of Human Genetics,** 83(3): 359–372. https://doi.org/10.1016/j.ajhg.2008.08.007

Mellado, M., Mellado, J., Valencia-Posadas, M., & Pittroff, W., 2008. The Relationship between Linear Type Traits and Fertility Traits in High-yielding Dairy Goats. **Reproduction in Domestic Animals = Zuchthygiene,** 43: 599–605. https://doi.org/10.1111/j.1439-0531.2007.00962.x

Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E., 2001. Prediction of total genetic value using genome-wide dense marker maps. **Genetics,** 157(4): 1819–1829. https://doi.org/11290733

Meuwissen, T., Hayes, B., & Goddard, M., 2013. Accelerating Improvement of Livestock with Genomic Selection. **Annual Review of Animal Biosciences**, 1: 221–237. https://doi.org/10.1146/annurev-animal-031412-103705

Meyer, K., Brotherstone, S., Hill, W. G., & Edwards, M. R., 1987. Inheritance of linear type traits in dairy cattle and correlations with milk production. **Animal Science**, 44(1): 1–10.

Moradi, M. H., Khaltabadi-Farahani, A. H., Khodaei-Motlagh, M., Kazemi-Bonchenari, M., & Mcewan, J., 2021. Genome-wide selection of discriminant SNP markers for

breed assignment in indigenous sheep breeds. **Annals of Animal Science,** 21(3): 807–831. https://doi.org/10.2478/aoas-2020-0097

Moradi, M. H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K. G., & McEwan, J. C., 2012. Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. **BMC Genetics,** 13. https://doi.org/10.1186/1471-2156-13-10

Mortimer, S. I., Hatcher, S., Fogarty, N. M., van der Werf, J. H. J., Brown, D. J., Swan, A. A., Greeff, J. C., Refshauge, G., Edwards, J. E. H., & Gaunt, G. M., 2017. Genetic parameters for wool traits, live weight, and ultrasound carcass traits in Merino sheep. **Journal of Animal Science**, 95(5): 1879–1891. https://doi.org/10.2527/jas.2016.1234

Mucha, S., Mrode, R., Coffey, M., Kizilaslan, M., Desire, S., & Conington, J., 2018. Genome-wide association study of conformation and milk yield in mixed-breed dairy goats. **Journal of Dairy Science**, 101(3): 2213–2225. https://doi.org/10.3168/jds.2017-12919

Narzhan, Z., Kairat, D., Aigerim, K., Tilek, K., Zhangylsyn, K., & Nurlan, T., 2022. SNP Genotyping Characterizes the Genome Composition of the New Baisary Fat-Tailed Sheep Breed. **Animals**, 12(11):1468. doi: 10.3390/ani12111468.

Nazari-Ghadikolaei, A., Mehrabani-Yeganeh, H., Miarei-Aashtiani, S. R., Staiger, E. A., Rashidi, A., & Huson, H. J., 2018. Genome-wide association studies identify candidate genes for coat color and mohair traits in the Iranian Markhoz Goat. **Frontiers in Genetics**, 9(4): 1–15. https://doi.org/10.3389/fgene.2018.00105

Nielsen, R., & Slatkin, M., 2013. An Introduction to Population Genetics. Sunderland, MA: Sinauer Associates. pp 304.

Obšteter, J., Jenko, J., & Gorjanc, G., 2021. Genomic Selection for Any Dairy Breeding Program via Optimized Investment in Phenotyping and Genotyping. **Frontiers in Genetics,** 12(2): 1–13. https://doi.org/10.3389/fgene.2021.637017

Ozmen, O., Kul, S., & Gok, T., 2020. Determination of genetic diversity of the

Akkaraman sheep breed from Turkey. **Small Ruminant Research**, 182. https://doi.org/10.1016/j.smallrumres.2019.10.009

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. https://doi.org/10.1111/age.12526

Pieramati, C., Lasagna, E., Panella, F., Piro, F., Giontella, A., & Sarti, F. M., 2011. Suitability of linear scoring in meat sheep: The practical case of Merinizzata Italiana breed. **Italian Journal of Animal Science,** 10(1): 51–53. https://doi.org/10.4081/ijas.2011.e11

Prieur, V., Clarke, S. M., Brito, L. F., McEwan, J. C., Lee, M. A., Brauning, R., Dodds, K. G., & Auvray, B., 2017. Estimation of linkage disequilibrium and effective population size in New Zealand sheep using three different methods to create genetic maps. **BMC Genetics,** 18(1): 68. https://doi.org/10.1186/s12863-017-0534-2

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. https://doi.org/10.1371/journal.pone.0176780

Qanbari, S., 2020. On the Extent of Linkage Disequilibrium in the Genome of Farm Animals. **Frontiers in Genetics**, 10(1): 1–11. https://doi.org/10.3389/fgene.2019.01304

R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rangwala, S. H., Kuznetsov, A., Ananiev, V., Asztalos, A., Borodin, E., Evgeniev, V., Joukov, V., Lotov, V., Pannu, R., & Rudnev, D., 2021. Accessing NCBI data using the NCBI Sequence Viewer and Genome Data Viewer (GDV). **Genome Research**, 31(1): 159–169.

Rannala, B., & Slatkin, M., 2000. Methods for multipoint disease mapping using

linkage disequilibrium. **Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society,** 19(S1): S71–S77.

Rekik, E., Ahbara, A. M., Abate, Z., Goshme, S., Getachew, T., Haile, A., Rischkowsky, B., & Mwacharo, J. M., 2022. Genomic analysis of 10 years of artificial selection in community- based breeding programs in two Ethiopian indigenous sheep breeds. **Animal Genetics**, 3: 447–451. https://doi.org/10.1111/age.13190

Risch, N., & Merikangas, K., 1996. The future of genetic studies of complex human diseases. **Science**, 273(5281): 1516–1517. https://doi.org/10.1126/science.273.5281.1516

Rochus, C. M., Tortereau, F., Plisson-Petit, F., Restoux, G., Moreno-Romieux, C., Tosser-Klopp, G., & Servin, B., 2018. Revealing the selection history of adaptive loci using genome-wide scans for selection: an example from domestic sheep. **BMC Genomics,** 19(1): 71. https://doi.org/10.1186/s12864-018-4447-x

Romanov, M. N., Zinovieva, N. A., & Griffin, D. K., 2021. British sheep breeds as a part of world sheep gene pool landscape: Looking into genomic applications. **Animals**, 11(4): 1–13. https://doi.org/10.3390/ani11040994

Rupp, R., Mucha, S., Larroque, H., McEwan, J., & Conington, J., 2016. Genomic application in sheep and goat breeding. **Animal Frontiers**, 6(1): 39–44. https://doi.org/10.2527/af.2016-0006

Safari, E., Fogarty, N. M., & Gilmour, A. R., 2005. A review of genetic parameter estimates for wool , growth , meat and reproduction traits in sheep. **Livestock Production Science**, 92: 271–289. https://doi.org/10.1016/j.livprodsci.2004.09.003.

Safari, E., Fogarty, N. M., & Gilmour, A. R., 2006. Sensitivity of response of multi-trait index selection to changes in genetic correlations between production traits in sheep. **Australian Journal of Experimental Agriculture**, 46(3): 283–290. https://doi.org/10.1071/EA04232.

Schaeffer, L. R., 2006. Strategy for applying genome-wide selection in dairy cattle. **Journal of Animal Breeding and Genetics**, 123(4): 218–223. https://doi.org/https://doi.org/10.1111/j.1439-0388.2006.00595.x

Scherf, B. D., & Pilling, D. (2015). The second report on the state of the world's animal genetic resources for Food and Agriculture.

Sempéré, G., Moazami-Goudarzi, K., Eggen, A., Laloë, D., Gautier, M., & Flori, L., 2015. WIDDE: a Web-Interfaced next generation database for genetic diversity exploration, with a first application in cattle. **BMC Genomics**, 16(1): 1–8.

Shumbusho, F., Raoul, J., Astruc, J. M., Palhiere, I., & Elsen, J. M., 2014. Potential benefits of genomic selection on genetic gain of small ruminant breeding programs. **Journal of Animal Science**, 91(8):3644-57. https://doi.org/10.2527/jas2012-6205

Shumbusho, F., Raoul, J., Astruc, J. M., Palhiere, I., Lemarié, S., & Elsen, J. M., 2017. Economic evaluation of genomic selection in small ruminants : a sheep meat breeding program. **Animal**, 10(6):1033-41. https://doi.org/10.1017/S1751731115002049

Simm, G. (1998). Genetic improvement of cattle and sheep. Farming press. pp 433.

Snelling, W. M., Hoff, J. L., Li, J. H., Kuehn, L. A., Keel, B. N., Lindholm-Perry, A. K., & Pickrell, J. K., 2020. Assessment of Imputation from Low-Pass Sequencing to Predict Merit of Beef Steers. **Genes**, 11(11). https://doi.org/10.3390/genes11111312

Sönmez, R., 1962. Breed characters and different productions in the Sakiz Sheep (Chios Schafe) of Turkey. **Zeitschrift Für Tierzüchtung Und Züchtungsbiologie**, 78(1–4): 281–286. https://doi.org/https://doi.org/10.1111/j.1439-0388.1962.tb01283.x

Sul, J. H., Martin, L. S., & Eskin, E., 2018. Population structure in genetic studies: Confounding factors and mixed models. **PLoS Genetics,** 14(12): e1007309–e1007309. https://doi.org/10.1371/journal.pgen.1007309

Taberlet, P., Coissac, E., Pansu, J., & Pompanon, F., 2011. Conservation genetics of

cattle, sheep, and goats. **Comptes Rendus - Biologies,** 334(3): 247–254. https://doi.org/10.1016/j.crvi.2010.12.007

Taberlet, P., Valentini, A., Rezaei, H. R., Naderi, S., Pompanon, F., Negrini, R., & Ajmone-Marsan, P., 2008. Are cattle, sheep, and goats endangered species? **Molecular Ecology**, 17(1): 275–284. https://doi.org/https://doi.org/10.1111/j.1365-294X.2007.03475.x

Tait-Burkard, C., Doeschl-Wilson, A., McGrew, M. J., Archibald, A. L., Sang, H. M., Houston, R. D., Whitelaw, C. B., & Watson, M., 2018. Livestock 2.0 - Genome editing for fitter, healthier, and more productive farmed animals. **Genome Biology,** 19(1): 1–11. https://doi.org/10.1186/s13059-018-1583-1

Tao, L., He, X. Y., Pan, L. X., Wang, J. W., Gan, S. Q., & Chu, M. X., 2020. Genome-wide association study of body weight and conformation traits in neonatal sheep. **Animal Genetics**, 51(2): 336–340. https://doi.org/10.1111/age.12904

Toro, M. A., & Caballero, A., 2005. Characterization and conservation of genetic diversity in subdivided populations. **Philosophical Transactions of the Royal Society B: Biological Sciences**, 360(1459): 1367–1378.

TUIK., 2022. Hayvansal Üretim İstatistikleri. https://data.tuik.gov.tr/Bulten/Index?p=Hayvansal-Uretim-Istatistikleri-Haziran-2022-45594. (July 2022).

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., & Posthuma, D., 2021. Genome-wide association studies. **Nature Reviews Methods Primers**, 1(1). https://doi.org/10.1038/s43586-021-00056-9

VanRaden, P. M., 2008. Efficient methods to compute genomic predictions. **Journal of Dairy Science**, 91(11): 4414–4423. https://doi.org/10.3168/jds.2007-0980

Visser, C., Marle-köster, E. Van, Snyman, M. A., Bovenhuis, H., & Crooijmans, R. P. M. A., 2013. Quantitative trait loci associated with pre-weaning growth in South African Angora goats. **Small Ruminant Research,** 112(1–3): 15–20.

https://doi.org/10.1016/j.smallrumres.2012.11.035

Wang, J, Zhou, H., Hu, J., Li, S., Luo, Y., & Hickford, J. G. H., 2016. Two single nucleotide polymorphisms in the promoter of the ovine myostatin gene (MSTN) and their effect on growth and carcass muscle traits in New Zealand Romney sheep. **Journal of Animal Breeding and Genetics**, 133(3): 219–226. https://doi.org/10.1111/jbg.12171

Wang, Jinliang., 2005. Estimation of effective population sizes from data on genetic markers. **Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences**, 360(1459): 1395–1409. https://doi.org/10.1098/rstb.2005.1682

Weale, M. E., 2010. Quality control for genome-wide association studies. **Methods in Molecular Biology**, 628: 341–372. https://doi.org/10.1007/978-1-60327-367-1_19

Wei, C., Wang, H., Liu, G., Wu, M., Cao, J., Liu, Z., Liu, R., Zhao, F., Zhang, L., Lu, J., Liu, C., & Du, L., 2015. Genome-wide analysis reveals population structure and selection in Chinese indigenous sheep breeds. **BMC Genomics**, 16(1): 1–12. https://doi.org/10.1186/s12864-015-1384-9

Weir, B. S., & Cockerham, C. C., 1984. Estimating F-Statistics for the Analysis of Population Structure. **Evolution,** 38(6): 1358–1370. https://doi.org/10.2307/2408641

White, S. N., Mousel, M. R., Herrmann-Hoesing, L. M., Reynolds, J. O., Leymaster, K. A., Neibergs, H. L., Lewis, G. S., & Knowles, D. P., 2012. Genome-Wide Association Identifies Multiple Genomic Regions Associated with Susceptibility to and Control of Ovine Lentivirus. **PLoS ONE**, 7(10). https://doi.org/10.1371/journal.pone.0047829

Wigginton, J. E., Cutler, D. J., & Abecasis, G. R., 2005. A note on exact tests of Hardy-Weinberg equilibrium. **The American Journal of Human Genetics,** 76(5): 887–893.

Wilson, C. S., Petersen, J. L., Blackburn, H. D., & Lewis, R. M., 2022. Assessing

Population Structure and Genetic Diversity in US Suffolk Sheep to Define a Framework for Genomic Selection. **Journal of Heredity,** 113(4): 431–443. https://doi.org/10.1093/jhered/esac026

Wilson, G. T., 1989. On the use of marginal likelihood in time series model estimation. **Journal of the Royal Statistical Society: Series B (Methodological)**, 51(1): 15–27.

World Bank., 2021. Opportunities for Climate Finance in the Livestock Sector : Removing Obstacles and Realizing Potential. https://openknowledge.worldbank.org/handle/10986/35495. (Access date: July 2022)

Wright, S., 1965. The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating. **Journal of Evolution**, 19(3): 395–420. http://www.jstor.org/stable/2406450.

Wright, S., 1951. The genetical structure of populations. **Annals of Eugenics,** 15(4): 323–354. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x

Xu, S. S., Gao, L., Shen, M., & Lyu, F., 2021. Whole-Genome Selective Scans Detect Genes Associated With Important Phenotypic Traits in Sheep (Ovis aries). **Frontiers in Genetics,** 12(11): 1–7. https://doi.org/10.3389/fgene.2021.738879.

Yalcin, B. C., 1986. Sheep and Goats in Turkey. Food and Agriculture Organisation of the United Nations, Rome.

Yang, Ji, Li, W. R., Lv, F. H., He, S. G., Tian, S. L., Peng, W. F., Sun, Y. W., Zhao, Y. X., Tu, X. L., Zhang, M., Xie, X. L., Wang, Y. T., Li, J. Q., Liu, Y. G., Shen, Z. Q., Wang, F., Liu, G. J., Lu, H. F., Kantanen, J., … Liu, M. J., 2016. Whole-Genome Sequencing of Native Sheep Provides Insights into Rapid Adaptations to Extreme Environments. **Molecular Biology and Evolution,** 33(10): 2576–2592. https://doi.org/10.1093/molbev/msw129

Yang, Jian, Weedon, M. N., Purcell, S., Lettre, G., Estrada, K., Willer, C. J., Smith, A. V, Ingelsson, E., O'Connell, J. R., Mangino, M., Mägi, R., Madden, P. A., Heath,

A. C., Nyholt, D. R., Martin, N. G., Montgomery, G. W., Frayling, T. M., Hirschhorn, J. N., McCarthy, M. I., … Consortium, G., 2011. Genomic inflation factors under polygenic inheritance. **European Journal of Human Genetics**, 19(7): 807–812. https://doi.org/10.1038/ejhg.2011.39

Yilmaz, O., Kizilaslan, M., Arzik, Y., Behrem, S., Ata, N., Karaca, O., Elmaci, C., & Cemal, I., 2022. Genome-wide association studies of preweaning growth and in vivo carcass composition traits in Esme sheep. **Journal of Animal Breeding and Genetics,** 139(1): 26–39. https://doi.org/10.1111/jbg.12640

Yin, X., Bose, D., Kwon, A., Hanks, S. C., Jackson, A. U., Stringham, H. M., Welch, R., Oravilahti, A., Fernandes Silva, L., Locke, A. E., Fuchsberger, C., Service, S. K., Erdos, M. R., Bonnycastle, L. L., Kuusisto, J., Stitziel, N. O., Hall, I. M., Morrison, J., Ripatti, S., … Wen, X., 2022. Integrating transcriptomics, metabolomics, and GWAS helps reveal molecular mechanisms for metabolite levels and disease risk. **The American Journal of Human Genetics**. 2(22):7. https://doi.org/10.1016/j.ajhg.2022.08.007

Zhang, L., Liu, J., Zhao, F., Ren, H., Xu, L., Lu, J., Zhang, S., Zhang, X., Wei, C., Lu, G., Zheng, Y., & Du, L., 2013. Genome-Wide Association Studies for Growth and Meat Production Traits in Sheep. **PLoS ONE**, 8(6): e66569. https://doi.org/10.1371/journal.pone.0066569

6 CURRICULUM VITAE

PERSONAL INFORMATION

Name Surname: Mehmet KIZILASLAN

Nationality: Türkiye (T.C)

EDUCATION

Degree	Institution	Graduation Date
MSc	The University of Edinburgh / Animal Breeding and Genetics	2016
BSc	Cumhuriyet Üniversity / Molecular Biology and Genetics	2013
High School	Cağaloğlu Anatolian High School / İstanbul	2009

WORK EXPERIENCE

Year	Institution	Time (years)
2017-Current	UHAEMM, Breeding and Genetics Department	6

LANGUAGE

English

PUBLICATIONS

1. Mucha, S., Mrode, R., Coffey, M., **Kizilaslan, M.,** Desire, S., & Conington, J. (2018). Genome-wide association study of conformation and milk yield in mixed-breed dairy goats. ***Journal of Dairy Science***, 101(3), 2213-2225. https://doi.org/10.3168/jds.2017-12919

2. İnanç, M.E., Tekin, K., Akkurt, M.Y., Olgac, K.T., Yılmaz, B., Çil, B., **Kizilaslan, M.,** Taşdemir, U., Tuncer, P.B., Büyükleblebici, S. and Uysal, O. (2018). Genome-wide association of male reproductive traits in Aksaray Malakli

dogs. ***Reproduction in Domestic Animals***, 53(6), pp.1555-1562. https://doi.org/10.1111/rda.13302

3. Yilmaz, O., **Kizilaslan, M.,** Arzik, Y., Behrem, S., Ata, N., Karaca, O., Elmaci, C., & Cemal, I. (2021). Genome-wide association studies of preweaning growth and in vivo carcass composition traits in Esme sheep. ***Journal of Animal Breeding and Genetics***, 00, 1–14. https://doi.org/10.1111/jbg.12640

4. **Kizilaslan, M.,** Arzik, Y., Cinar, M. U. & Konca, Y. (2021). Genome-wise engineering of ruminant nutrition: nutrigenomics: applications, challenges, and future perspectives – a review. ***Annals of Animal Science***, https://doi.org/10.2478/aoas-2021-0057

5. **Kizilaslan, M.,** Arzik, Y., White, S. N., Piel, L. M. W., & Cinar, M. U. (2022). Genetic parameters and genomic regions underlying growth and linear type traits in Akkaraman sheep. ***Genes***, 13(8), 1414. https://doi.org/10.3390/genes13081414

6. Arzik, Y., **Kizilaslan, M.,** White, S.N., Piel, L.M.W., Cinar, M.U. (2022). Estimates of genomic heritability and genome-wide association studies for blood parameters in Akkaraman sheep. ***Nature Scientific Reports*** (Under revision). https://doi.org/10.21203/rs.3.rs-1554456/v1

7. Master Thesis. (2017). Genome-wide association studies of milk production traits and clinical mastitis in UK dairy goats. Scotland's Rural College (SRUC), The Roslin Institute Building, United Kingdom.

8. Behrem, S., Arzik, Y., **Kizilaslan, M.,** Yildiz C., Tutenk, S. (2019). Comparison of different age groups in terms of fleece diameter and length in Anatolian Merino sheep. I. International Livestock Science Congress. 03 Oct – 01 Nov, 2019, Antalya, Turkey, (Poster Sunumu)

9. G.Ünal, İ. Ünal, Ç.M. Sakar, E. Ünay, U. Gül, T. Özüdoğru, **M. Kizilaslan.** (2020). Socioeconomic effect of supports related to in situ conservation of animal genetic resources: The case of Anatolian Black cattle. 2nd World

Conference on Sustainable Life Sciences (WOCOLS 2020 E-CONFERENCE) 8-9 Aralık 2020 (Sözlü Sunum)

10. Book Section, January 2022, Animal Breeding in Farm Animals, Republic of Türkiye Ministry of Agriculture and Forestry, Department of Training and Publication, ISBN NO 978-605-9175-80-7.